Test-Retest Reliability of Anxiety Symptoms and Diagnoses With the Anxiety Disorders Interview Schedule for *DSM-IV:* Child and Parent Versions

WENDY K. SILVERMAN, PH.D., LISSETTE M. SAAVEDRA, M.S., AND ARMANDO A. PINA, M.S.

ABSTRACT

Objective: To examine the test-retest reliability of the *DSM-IV* anxiety symptoms and disorders in children with the Anxiety Disorders Interview Schedule for *DSM-IV:* Child and Parent Versions (ADIS for *DSM-IV*:C/P). **Method:** Sixty-two children (aged 7–16 years) and their parents underwent two administrations of the ADIS for *DSM-IV*:C/P with a test-retest interval of 7 to 14 days. **Results and Conclusions:** Results revealed that the ADIS for *DSM-IV*:C/P is a reliable instrument for deriving *DSM-IV* anxiety disorder symptoms and diagnoses in children. The ADIS for *DSM-IV*:C/P was found to have excellent reliability in symptom scale scores for separation anxiety disorder, social phobia, specific phobia, and generalized anxiety disorder and good to excellent reliability for deriving combined diagnoses of these disorders, as well as using child-only and parent-only interview information. Reliability coefficients were generally similar and, in most instances, superior to those found in previous ADIS-C/P reliability studies. Limitations and directions for future research are discussed. *J. Am. Acad. Child Adolesc. Psychiatry*, 2001, 40(8):937–944. **Key Words:** reliability, anxiety, Anxiety Disorders Interview Schedule for *DSM-IV:* Child and Parent Versions.

During the past two decades, considerable progress has been made in the area of child psychopathology and its treatment. This progress can in large part be attributed to improvements in the standardized nomenclatures, specifically, in the recent versions of the *DSM* (American Psychiatric Association, 1987, 1994). These improvements have allowed clinical and epidemiological researchers to develop standardized instruments, particularly structured and semistructured interview schedules to assist in the diagnosis of child emotional and behavioral problems. As these interviews have been developed, investigators have conducted studies to examine whether their use leads to improvements in diagnostic reliability. The customary way in which this has been done has been through test-retest designs (Reich, 2000; Silverman, 1991, 1994).

The existing literature on diagnostic reliability of children's problems mainly consists of studies in which interview schedules based on either *DSM-III* or *DSM-III-R* diagnostic criteria were used. Thus, for most of the widely used schedules, such as the Diagnostic Interview Schedule for Children (DISC) (Schwab-Stone et al., 1993; Shaffer et al., 1993), the Diagnostic Interview for Children and Adolescents (DICA) (Herjanic, 1982), and the Child and Adolescent Psychiatric Assessment (CAPA) (Angold and Costello, 1995), studies have appeared that evaluated the reliability of either *DSM-III* or *DSM-III-R* diagnoses. Overall, results from these reliability studies demonstrate that diagnostic reliability is enhanced with the use of these schedules (McClellan and Werry, 2000; Silverman, 1991, 1994).

In a recent special section that appeared in the *Journal of the American Academy of Child and Adolescent Psychiatry* (McClellan and Werry, 2000), an update of the structured interview schedules designed to assess *all* the main types of children's psychiatric disorders was provided. In addition to describing the *DSM-IV* version of the existing schedules, an update of the "latest" research findings was reported for each. Overall, the results reported underscored the value and the utility of these interviews for diagnosing child psychopathology

Accepted February 27, 2001.

From Florida International University, Miami.

This study was funded in part by NIMH grant 54690. The authors thank Candice Alfano, Steven Berman, Janene Bussell, Ximena Franco, Margarita Graeber, Barbara Lopez, and Carl Weems for their assistance in the data collection and organization process of this project, with special thanks to Anne Marie Albano.

Correspondence to Dr. Silverman, Child and Family Psychosocial Research Center, Child Anxiety and Phobia Program, Department of Psychology, Florida International University, Miami, FL 33199; e-mail: Silverw@fiu.edu.

0890-8567/01/4008–0937©2001 by the American Academy of Child and Adolescent Psychiatry.

in general. With respect to diagnosing anxiety disorders, in particular, as Table 1 shows, only one interview (i.e., DISC-IV) (Shaffer et al., 2000) had available test-retest reliability of *DSM-IV* anxiety disorder diagnoses based on child, parent, and combined reports. In light of this, research focusing on the utility of the various child structured interviews for diagnosing *DSM-IV* anxiety disorders in children and adolescents would be worthwhile. This is particularly true in light of the high prevalence of anxiety disorders in youth (see Bernstein, 1996; Silverman and Ginsburg, 1998).

The Anxiety Disorders Interview Schedule for Children (ADIS-C/P) (Silverman and Albano, 1996) has been revised in accordance with *DSM-IV*. Earlier studies on both interrater and test-retest reliability of the *DSM-III* and *DSM-III-R* anxiety diagnoses and specific symptoms using the ADIS-C/P demonstrated satisfactory to excellent levels of reliability (Silverman and Eisen, 1992; Silverman and Nelles, 1988; Silverman and Rabian, 1995). The reliability of the ADIS-C/P also has shown similar satisfactory levels with an Australian sample (Rapee et al., 1994). Since the development of the ADIS-C/P and its evaluation, the interview has gained wide recognition and use in both clinical and research settings and has been the interview schedule used in all the published childhood anxiety clinical trials (e.g., Barrett et al., 1996; Beidel et al., 2000; Kendall, 1994; Kendall et al., 1997; Hayward et al., 2000; Silverman et al., 1999a,b; Spence et al., 2000). Furthermore, the *DSM-IV* version continues to be used in a large number of ongoing clinical trial investigations. In light of this, examining the reliability of *DSM-IV* anxiety symptoms and disorders in children

TABLE 1
Test-Retest Reliability of Anxiety Diagnoses in Children With Structured Interviews

Study	Interview *DSM* Version (Retest Interval)	*N* Age Range (Informant)	κ Coefficients for Diagnoses Reported
Angold and Costello (2000)	CAPA *DSM-III-R* (1–11 days)	77 9–13 yr (Child)	0.74 OAD 0.79 GAD
Ambrosini (2000)	K-SADS *DSM-III-R* (2–38 days)	20 6–18 yr (Child and parent combined)	0.78 OAD (current) 0.78 OAD (lifetime) 0.80 SIP (lifetime)
Reich (2000)	DICA *DSM-III-R* (1–7 days)	40 6–12 yr (Child)	0.55 OAD 0.60 SAD 0.65 SIP
		35 (Adolescent)	0.72 OAD 0.75 SAD (past)
Shaffer et al. (2000)	NIMH DISC-IV (Mean = 6.6 days)	84 (Parent)	0.96 SP 0.54 SOP 0.58 SAD 0.65 GAD
		82 9–17 yr (Child)	0.68 SP 0.25 SOP 0.46 SAD
		82 (Child and parent combined)	0.86 SP 0.48 SOP 0.51 SAD 0.58 GAD

Note: CAPA = Child and Adolescent Psychiatric Assessment; K-SADS = Schedule for Affective Disorders and Schizophrenia for School-Age Children; DICA = Diagnostic Interview for Children and Adolescents; NIMH = National Institute of Mental Health; DISC = Diagnostic Interview Schedule for Children; OAD = overanxious disorder; GAD = generalized anxiety disorder; SIP = simple phobia; SOP = social phobia; SAD = separation anxiety disorder; SP = specific phobia.

using the ADIS for *DSM-IV:*C/P is an important and timely issue.

This study reports on the test-retest reliability of *DSM-IV* anxiety symptoms and diagnoses with the ADIS for *DSM-IV:*C/P (Silverman and Albano, 1996) in a sample of children (aged 7–16 years) referred to a childhood anxiety disorders specialty clinic. Retest reliability was evaluated on three levels. First, retest reliability of *diagnoses* was evaluated. Second, retest reliability of *symptom scale scores* was evaluated (i.e., the total number of symptoms endorsed at time 1 and time 2). Third, the reliability of *child, parent, and clinician severity ratings* was examined. As in previous ADIS-C/P reliability reports, this report focuses on the total sample as well as younger (7–11 years old) and older (12–16 years old) age groups. In addition to examining the reliability of the ADIS for *DSM-IV:*C/P on these three levels, this study reports reliability information of child, parent, and child-parent combined reports.

METHOD

Participants

Participants consisted of 62 children (aged 7–16 years; mean age = 10.15 years) and their parents who presented to the Child Anxiety and Phobia Program, housed within the Child and Family Psychosocial Research Center at Florida International University, Miami. The participants' sociodemographic information is presented in Table 2. The children were referred through school counselors, other mental or medical health professionals, or self-referral. All children were referred to the Center because of difficulties with excessive fear and/or anxiety.

TABLE 2
Sample Characteristics

	n	%
Age		
7–11 yr	39	63.0
12–16 yr	23	37.0
Gender		
Male	26	39.6
Female	36	60.4
Ethnicity		
Euro-American	26	41.8
Hispanic	32	51.8
African American	2	3.2
Other	2	3.2
Socioeconomic status		
$0–11,999	6	9
$12,000–20,999	4	6
$21,000–30,999	12	19
$31,000–40,999	8	13
$41,000–50,999	23	37
>$51,000	9	16

Exclusionary criteria for participation in this study were developmental delays (e.g., Asperger's disorder, mental retardation, autism) or severe psychopathology (e.g., schizophrenia). Screening for these criteria is accomplished through a standardized telephone screen used within the Center. Additional screening sections also are contained on the child and parent interview schedules and are administered as deemed necessary by the clinician. Only one child was excluded from participation in the study, having met criteria for Asperger's disorder in addition to an anxiety disorder (thus bringing the final sample size to 62). All interviews were conducted in English. Bilingual Spanish-speaking participants who present to the Child Anxiety and Phobia Program are given the option of having the interview in either English or Spanish. The majority chose the English version, and in this article we report the results of the English interview. In a subsequent article the Spanish version reliability results will be reported.

Measure

The Anxiety Disorders Interview Schedule for *DSM-IV:* Child and Parent Versions (ADIS for *DSM-IV:*C/P) (Silverman and Albano, 1996) (usually referred to hereafter as ADIS-C/P) are semistructured interviews designed specifically for the diagnosis of anxiety and other related disorders in children and adolescents (hereafter referred to as children). In addition to providing diagnoses based on child report (via the ADIS-C) and parent report (via the ADIS-P), the interviews provide combined diagnoses (via the ADIS-C *and* ADIS-P). In administering the child and parent interviews, *DSM-IV* symptoms are judged by the child and parent as either present ("yes") or absent ("no"). Although on the interview schedules a response option of "other" exists (e.g., if respondents answer "don't know," or "sometimes"), the interviewers in this study were trained to follow up such answers with a prompt in an attempt to obtain a "yes" or "no" response (e.g., "I know it is difficult to say for sure, but would you say [description of symptom] is generally more absent or present, or more yes or no"). In the majority of cases, this additional prompt was sufficient to elicit an elaborated response from the children and parents, allowing for a determination of whether the symptom was either absent or present. However, for the small number of cases for which it was not sufficient, the answer was scored as "other" and was not counted toward a diagnosis.

For both the child and parent interviews, the total number of "yes" responses is calculated to obtain a total symptom scale score *and* to determine whether the total number of symptoms endorsed as "yes" is enough to meet the number of symptoms required to meet *DSM-IV* criteria. If the number of symptoms endorsed as "yes," or as present, is sufficient to meet *DSM-IV* criteria, the child and/or parent is then asked whether those symptoms, taken together, lead to significant clinical interference or impairment. Interference or impairment is described in the interviews as "messing things up in important areas" in that the symptoms are interfering with the child's life in areas that include school, family life, or peers, or perhaps they are leading to internal distress in the child. Impairment ratings are made by the children and parents using a 9-point scale (i.e., 0–8) provided in the form of a "Feelings Thermometer." To warrant a final diagnosis on the child and/or parent interview, the impairment rating for each diagnosis must be 4 or greater (i.e., leads to at least "some" or a moderate degree of impairment).

In deriving combined diagnoses, child *and* parent impairment ratings of each diagnosis are considered. As detailed by Albano and Silverman (1996), in cases in which the child and parent interviews yield the same diagnosis, the child receives the diagnosis and is assigned the higher of the two impairment ratings (i.e., either child or parent). In the case of a discrepancy between the child and parent interviews, if

one or both interviews yield a diagnosis with an impairment rating of 4 or more (on the 9-point scale), the child receives the diagnosis and is assigned the higher of the two impairment ratings. In cases of multiple diagnoses, relative impairment of each specific combined diagnosis is used as the basis for prioritizing the primary combined diagnosis, the secondary combined diagnosis, etc. This includes diagnoses for all disorders, not just anxiety, which can be reliably differentiated using the ADIS-C/P interviews (Silverman and Nelles, 1988).

Diagnosticians

The majority of the interviews were conducted by seven (five female and two male) graduate students in psychology; a few were conducted by the first author (W.K.S.). All of the diagnosticians were thoroughly and extensively trained in handling the different types of circumstances that typically arise during the psychiatric interview of the child, especially in relation to determining the impairment ratings and the priority ranking of diagnoses. Training consisted of having several meetings on how to conduct the interview and issues involved in the diagnosis of *DSM-IV* anxiety disorders, depression, externalizing disorders, and other related childhood disorders. In addition, the diagnostician was required to observe five separate parent–child interviews and match on five consecutive diagnoses with a trained diagnostician. Also, the diagnostician was required to conduct an interview under observation and match the diagnoses. Furthermore, diagnosticians were required to agree not only on the specific diagnoses, but on the order of diagnoses (primary, secondary, tertiary, etc.).

Procedure

Prior to the administration of the interviews, the children and parents were asked to read and sign an informed consent statement that provides a detailed description of the assessment procedures. Upon receiving the participants' signed consent, the ADIS-C and the ADIS-P were administered. The order of administration was randomly determined. The diagnostician was not permitted to discuss with either the child or parent the information or details provided by the other informant. This was done to avoid one informant's biasing the other informant's response. A random spot-check of 25% of the interviews on videotape confirmed that the interviewers did not "cross over" in their interviewing with the child and/or parent. Upon completion of the first interview, the diagnostician derived *DSM-IV* diagnoses from (1) the child interview, (2) the parent interview, and (3) the combined interview data.

For the retest, 7 to 14 days later, the same interviewer readministered the second interview to the children and parents. Across participants, the mean number of days between the test and retest was 11.7. At the time of the retest, the children and parents were reminded that the two interviews are separate and independent and they therefore need not try to remember their answers from the first interview. The second administration was conducted in the same order as the first administration. Upon completion of the second interview, the diagnostician again derived *DSM-IV* diagnoses from (1) the child interview, (2) the parent interview, and (3) the combined interview data.

A few additional comments are warranted regarding the use of the same diagnostician to conduct the child and parent test and retest interviews. Most previous child anxiety disorder clinical trials have used this procedure (e.g., Beidel et al., 2000; Hayward et al., 2000; Silverman et al., 1999a,b; Spence et al., 2000). The use of the same diagnostician is also more akin to what occurs in actual current clinical practice (i.e., it is likely to be the same clinician interviewing the child and parent at the first interview and second interview, if a second interview is conducted). In addition, calls have been made in the child psychiatric research literature regarding the need for research that is sensitive to issues of "effectiveness" (Hoagwood et al., 1995), that is, to conduct research that is feasible for and generalizable to actual clinical settings. Of further note is that the first reliability study on the ADIS-C/P (Silverman and Nelles, 1988) focused exclusively on the issue of interviewer variance (in that the procedure involved the use of different interviewers) and the results revealed high interrater diagnostic reliability. Given the minimal interviewer variance found in Silverman and Nelles (1988), the present reliability study, like past reliability studies (Silverman and Eisen, 1992; Silverman and Rabian, 1995), evaluated variance due to time and/or the informants themselves. It is recognized, however, that this procedure, though preferable from a clinical and external validity perspective, could potentially inflate reliabilities.

Data Analytic Plan

Three main statistics were computed in examining reliability at three levels: κ coefficient, intraclass correlation coefficient (ICC), and Pearson correlation coefficient.

Kappa Coefficient. Kappa coefficients (Cohen, 1960) were used to examine the reliability of the specific anxiety disorder diagnoses and the specific anxiety symptoms. The criteria provided by Landis and Koch (1977) were used to evaluate the obtained κ coefficients. The criteria are as follows: $\kappa > 0.74$ indicates excellent reliability; κ between 0.59 and 0.74 indicates good reliability; κ between 0.40 and 0.58 indicates fair reliability; and $\kappa < 0.40$ indicates poor reliability.

Intraclass Correlation Coefficient. ICCs were used to examine the reliability of the symptom scale scores. More specifically, ICCs were used to evaluate the consistency in the summation of the symptom scale scores for the diagnoses of separation anxiety disorder (SAD), social phobia (SOP), specific phobias (SP), and generalized anxiety disorder (GAD), with the model and guidelines suggested by Maxwell and Pilliner (1968). These disorder subscales were selected because to examine the reliability of symptom subscales, every question within that subscale must be asked. On the ADIS-C/P, all of the questions are asked for these disorder subscales, even if the informant reports "no" to the initial screening question. The same criteria used to evaluate κ coefficients were used to evaluate ICCs (Landis and Koch, 1977).

Pearson Correlation Coefficient. Pearson correlation coefficients were used to examine the reliability of the impairment ratings (with the 9-point scale previously mentioned) endorsed in the child and parent interviews.

RESULTS

Diagnoses

The most prevalent disorder for the total sample (N = 62) was SP for the child interview, followed by GAD and SAD. For the parent interview, the most prevalent disorder for the total sample also was SP, followed by GAD and SAD. For the combined diagnosis, the most prevalent disorder for the total sample was GAD, followed by SAD and SP.

Comorbid diagnoses in this sample were common. Of the 62 participants, 20 (33%) met diagnostic criteria for a primary anxiety disorder diagnosis only; 42 (67%) participants met criteria for two or more disorders. Of these 42 participants, 33 met criteria for two or more *DSM-IV* anxiety disorders; 3 participants met criteria for an addi-

tional anxiety disorder as well as depression. Six (14%) of the 42 participants met criteria for an additional externalizing disorder only.

Kappa Coefficients for Specific Diagnoses

Kappa coefficients were computed for the diagnoses that had a sufficient number of cases meeting criteria for the disorder ($n > 5$), as done in previous studies (e.g., Jensen et al., 1995; Schwab-Stone et al., 1993).

Child Interview. For the total sample using the child interview, the κ coefficients obtained for SAD, SOP, SP, and GAD were in the good to excellent range (κ = 0.63–0.80; see Table 3). For the younger children, κ coefficients for SAD, SOP, SP, and GAD also were excellent, ranging from 0.71 to 0.84. For the older group, κ coefficients for SOP, SP, and GAD were in the fair range, from 0.43 to 0.63.

Parent Interview. For the total sample using the parent interview, the κ coefficients obtained for SAD, SOP, SP, and GAD were in the good to excellent range (κ = 0.65–0.88). For the younger children, κ coefficients for SAD, SOP, SP, and GAD indicated good to excellent reliability, ranging from 0.73 to 0.92. For the older group, κ coefficients for SOP, SP, and GAD indicated a fair to good range of reliability, with coefficients from 0.51 to 0.71.

Combined Diagnoses. Kappa coefficients for SAD, SOP, SP, and GAD were all in the excellent range, with κ values spanning from 0.80 to 0.92. For the younger children, κ coefficients for SAD, SOP, SP, and GAD indicated excellent reliability, with all κ values being 0.84 and 0.85. For the older children, the κ coefficients for SOP, SP, and GAD indicated good to excellent reliability, ranging from 0.70 to 1.00.

Intraclass Correlation Coefficients for the Symptom Scales

Child Interview. Test-retest reliability of the symptom scales for the two administrations of the ADIS-C also are presented in Table 3. For the total sample, ICCs indicated excellent levels of reliability for SAD, SOP, SP, and GAD, with ICCs ranging from 0.78 to 0.95. For the younger

TABLE 3
Kappa and Intraclass Correlation Coefficients for Child, Parent, and Combined Diagnoses With the Anxiety Disorders Interview Schedule for *DSM-IV*: Child and Parent Versions

	7–11 Years (n = 39)		12–16 Years (n = 23)		Total Sample (N = 62)	
	κ	ICC	κ	ICC	κ	ICC
Child diagnoses						
SAD	0.78	0.85	0.63	0.82	0.78	0.87
SOP	0.72	0.89	0.57	0.96	0.71	0.87
SP	0.84	0.92	0.47	0.99	0.80	0.95
GAD	0.71	0.90	0.43	0.81	0.63	0.78
ADHD	—	—	—	—	0.61	—
Parent diagnoses						
SAD	0.84	0.86	—	0.84	0.88	0.89
SOP	0.92	0.89	0.71	0.52	0.86	0.81
SP	0.73	0.99	0.51	0.94	0.65	0.96
GAD	0.78	0.90	0.51	0.84	0.72	0.87
ADHD	0.68	—	—	—	1.00	—
ODD	0.62	—	—	—	0.78	—
Combined diagnoses[a]						
SAD	0.84		—		0.84	
SOP	0.85		1.00		0.92	
SP	0.84		0.71		0.81	
GAD	0.84		0.70		0.80	
ADHD	—		—		1.00	
ODD	—		—		0.62	

Note: — denotes that κ or ICC could not be computed because of an insufficient number of cases. ICC = intraclass correlation coefficient; SAD = separation anxiety disorder; SOP = social phobia; SP = specific phobia; GAD = generalized anxiety disorder; ADHD = attention-deficit/hyperactivity disorder; ODD = oppositional defiant disorder. Analyses for only these disorders were included because they met the necessary sample size.

[a] Because ICCs were used to examine the reliability of symptom scales, these were not computed for the combined diagnoses.

group, ICCs indicated excellent levels of reliability, ranging from 0.85 to 0.92. The older group's ICCs were similarly excellent, ranging from 0.81 to 0.99. Subsequent sign tests revealed no significant age differences in the reliability of symptom scale scores for the two age groups.

Parent Interview. Test-retest reliability of the symptom scales for the ADIS-P also are presented in Table 3. For the total sample, ICCs indicated excellent levels of reliability for SAD, SOP, SP, and GAD, with ICCs ranging from 0.81 to 0.96. For the younger group, ICCs ranged from 0.86 to 0.99, indicating excellent reliability. For the older group, ICCs ranged from 0.52 to 0.94, indicating good to excellent reliability. Subsequent sign tests revealed no significant age differences in the reliability of symptom scale scores.

Pearson Correlation Coefficients for Impairment Ratings

Pearson correlation coefficients were calculated to examine the consistency from time 1 to time 2 on the impairment ratings for each diagnosed disorder as rated by the child and parent. For the child interview, the correlations for the impairment ratings of SAD, SOP, SP, and GAD were 0.60, 0.10, 0.84, and 0.72, respectively. With the exception of SOP, these rates were in the good to excellent range. For the parent interview, consistency on the impairment rates also was in the good to excellent range. For the diagnosis of SAD, SOP, SP, and GAD, correlation coefficients were 0.56, 0.81, 0.78, and 0.84, respectively. The correlations for the clinicians' ratings of the impairment of SAD, SOP, SP, and GAD were 0.80, 0.84, 0.84, and 0.82, respectively.

DISCUSSION

The results of this study indicate that the latest revision of the ADIS-C/P (Silverman and Albano, 1996) yields highly reliable *DSM-IV* anxiety disorder symptoms and diagnoses in children and adolescents. The ADIS for *DSM-IV:* C/P was found to have excellent reliability for deriving combined diagnoses of SAD, SOP, SP, and GAD. The ADIS-C/P was highly reliable for deriving diagnoses with either child or parent information. Specifically, the ADIS-C yielded excellent reliability for the diagnoses of SAD and SP and good reliability for SOP and GAD. The ADIS-P also yielded excellent reliability for the diagnoses of SAD and SOP and good reliability for SP and GAD. The reliability coefficients obtained for the ADIS for *DSM-IV:*C/P were consistent and, in a number of instances, superior to those found in previous ADIS-C/P *DSM-III-R* reliability studies (e.g., Rapee et al., 1994; Silverman and Eisen, 1992) as well as other interview schedules (e.g., Ambrosini, 2000; Angold and Costello, 2000; Reich, 2000; Shaffer et al., 2000).

Also examined in this study was the reliability of the symptom scales of the ADIS for *DSM-IV:*C/P. These also had excellent reliability when either child or parent report was used. In particular, with the ADIS-C, the symptom scales for SAD, SOP, and SP were all in the excellent range of reliability. Similarly, with the ADIS-P, the symptom scale scores of SAD, SOP, SP, and GAD were all in the excellent range.

Overall, reliability of anxiety symptoms and diagnoses with the ADIS for *DSM-IV:*C/P was in the good to excellent range for the anxiety disorders and in the excellent range for all the symptom scale scores that were examined. Although for some disorders reliability of diagnoses appeared to be lower for the older children than the younger children, using either child or parent report (e.g., κ coefficients for SOP and GAD), this was not true of the symptom scale scores. This is probably related to the respective analyses conducted for each level. Because the κ statistic uses a dichotomous measure of reliability as opposed to the ICC, which uses a continuous measure, κ is not as sensitive to slight changes in reports from the first to the second administration (Jensen et al., 1995).

The current study makes an important contribution in that it provides information about the reliability of diagnosing *DSM-IV* anxiety disorders and symptoms in children with the ADIS-C/P, an interview that has wide use among clinical investigators studying childhood anxiety.

Moreover, several strengths of this study are worth mentioning. For example, the current study used a sample from an outpatient childhood anxiety specialty clinic. As in previous studies of the ADIS-C/P, this permitted the examination of relatively less prevalent anxiety disorders and symptoms (e.g., SOP and GAD) that are usually not examined by other diagnostic interviews (e.g., Schedule for Affective Disorders and Schizophrenia for School-Age Children [Ambrosini, 2000], DICA [Reich, 2000]), in which only the most prevalent disorder (usually SAD) is examined and/or the anxiety disorders are grouped together and reliability examined for "*any* anxiety disorder." In addition, a benefit of using an outpatient sample is that inflated reporting or overreporting of symptoms usually characteristic of inpatient samples is probably not present (e.g., Hodges, 1993).

Limitations of the Study and Directions for Research

This study requires replication and extension with a sample that has a higher frequency of some other anxiety disorders. Reliability of disorders such as panic disorder, agoraphobia, posttraumatic stress disorder, and obsessive-compulsive disorder was not examined because of the low number of cases of these disorders. This low presentation rate is consistent with past ADIS-C/P reliability studies (e.g., Silverman and Eisen, 1992) as well as those conducted on other interview schedules (e.g., Ambrosini, 2000; Reich, 2000; Shaffer et al., 2000). Another important direction for future research would be to examine reliability of other disorders covered on the interviews, such as the affective and other externalizing disorders.

As mentioned previously, it is possible that the present study's reliability findings were somewhat inflated because of the decision to use the same diagnostician for the test and retest. On the other hand, as also mentioned, the advantage of conducting an ADIS for *DSM-IV:*C/P reliability study that most closely resembles how the interviews have been used in clinical trials (e.g., Beidel et al., 2000; Hayward et al., 2000; Silverman et al., 1999a,b; Spence et al., 2000) and in clinical practice, likely offsets this potential limitation.

Clinical Implications

The results of this study suggest that children, adolescents, and their parents can provide reliable reports of *DSM-IV* anxiety symptoms and disorders as well as other co-occurring disorders using the ADIS for *DSM-IV:*C/P. Unlike past ADIS-C/P studies (Silverman and Eisen, 1992) as well as studies on other interview schedules (e.g., Edelbrock, 1985) that compared child versus adolescent reports, the current study found that younger children were as reliable as older children. These findings can in part be attributed to the interview's increased clarity and specificity in describing anxiety symptoms as well as modifications to render it more usable by younger children (e.g., calendars, fear thermometers). That the overall findings were generally favorable for the two age groups for both children and parents suggests that researchers and practitioners can use either child or parent reports in situations in which economic and time resources are limited. In contexts in which economic and time resources are not so limited, the use of multiple informants should continue to be encouraged (Edelbrock, 1994; Silverman and Kurtines, 1996).

REFERENCES

Albano AM, Silverman WK (1996), *The Anxiety Disorders Interview Schedule for Children for DSM-IV:* Clinician Manual (Child and Parent Versions). San Antonio, TX: Psychological Corporation

Ambrosini P (2000), Historical development and present status of the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS). *J Am Acad Child Adolesc Psychiatry* 39:49–58

American Psychiatric Association (1987), *Diagnostic and Statistical Manual of Mental Disorders, 3rd edition-revised (DSM-III-R).* Washington, DC: American Psychiatric Association

American Psychiatric Association (1994), *Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV).* Washington, DC: American Psychiatric Association

Angold A, Costello EJ (1995), A test-retest reliability study of child-reported psychiatric symptoms and diagnoses using the Child and Adolescent Psychiatric Assessment (CAPA-C). *Psychol Med* 25:755–762

Angold A, Costello EJ (2000), The Child and Adolescent Psychiatric Assessment (CAPA). *J Am Acad Child Adolesc Psychiatry* 39:39–48

Barrett PM, Dadds MR, Rappee RM (1996), Family treatment of childhood anxiety: a controlled trial. *J Consult Clin Psychol* 64:333–342

Beidel DC, Turner SM, Morris TL (2000), Behavioral treatment of childhood social phobia. *J Consult Clin Psychol* 68:1072–1080

Bernstein G (1996), Anxiety disorders in children and adolescents: a review of the past 10 years. *J Am Acad Child Adolesc Psychiatry* 35:1110–1119

Cohen J (1960), A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20:37–46

Edelbrock C (1994), Assessing child psychopathology in developmental follow-up studies. In: *Developmental Follow-up: Concepts, Domains, and Methods.* San Diego, CA: Academic Press, pp 183–196

Edelbrock D (1985), Age differences in the reliability of the psychiatric interview of the child. *Child Dev* 53:265–275

Hayward C, Varady S, Albano AM, Thienemann M, Henderson L, Schatzberg AF (2000), Cognitive-behavioral group therapy for social phobia in female adolescents: results of a pilot study. *J Am Acad Child Adolesc Psychiatry* 39:721–734

Herjanic B (1982), Development of a structured psychiatric interview for children: agreement between child and parent on individual symptoms. *J Abnorm Child Psychol* 10:307–324

Hoagwood K, Hibbs E, Brent D, Jensen P (1995), Introduction to a special section: efficacy and effectiveness in studies of child and adolescent psychotherapy. *J Consult Clin Psychol* 63:683–687

Hodges K (1993), Structured interviews for assessing children. *J Child Psychol Psychiatry* 34:49–68

Jensen PS, Roper M, Fisher P et al. (1995), Test-retest reliability of the Diagnostic Interview Schedule for Children (DISC 2.1). *Arch Gen Psychiatry* 52:61–71

Kendall PC (1994), Treating anxiety disorders in children: results of a randomized clinical trial. *J Consult Clin Psychol* 62:100–110

Kendall PC, Flannery-Schroder E, Panichelli-Mindel SM, Southam-Gerow M, Henin A, Warman M (1997), Therapy for youths with anxiety disorders: a second randomized clinical trial. *J Consult Clin Psychol* 65:366–380

Landis J, Koch G (1977), The measurement of observer agreement for categorical data. *Biometrics* 33:159–174

Maxwell AE, Pilliner AE (1968), Deriving coefficients of reliability and agreement for ratings. *Br J Math Stat Psychol* 21:105–116

McClellan JM, Werry JS (2000), Research on psychiatric diagnostic interviews for children and adolescents: introduction. *J Am Acad Child Adolesc Psychiatry* 39:19–27

Rapee RM, Barrett PM, Dadds MR, Evans L (1994), Reliability of the *DSM-III-R* childhood anxiety disorders using structured interview: interrater and parent–child agreement. *J Am Acad Child Adolesc Psychiatry* 33:984–992

Reich W (2000), Diagnostic Interview Schedule for Children and Adolescents (DICA). *J Am Acad Child Adolesc Psychiatry* 39:59–67

Schwab-Stone M, Fisher P, Piacentini J, Shaffer D, Davies M, Briggs M (1993), The Diagnostic Interview Schedule for Children-Revised Version (DISC-R), II: test-retest reliability. *J Am Acad Child Adolesc Psychiatry* 32:651–657

Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME (2000), NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *J Am Acad Child Adolesc Psychiatry* 39:28–39

Shaffer D, Schwab-Stone M, Fisher P et al. (1993), The Diagnostic Interview for Children-Revised (DISC-R), I: preparation, field testing, interrater reliability, and acceptability. *J Am Acad Child Adolesc Psychiatry* 32:643–650

Silverman WK (1991), Diagnostic reliability of anxiety disorders in children using structured interviews. *J Anxiety Disord* 5:105–124

Silverman WK (1994), Structured diagnostic interviews. In: *International Handbook of Phobic and Anxiety Disorders in Children and Adolescents: Issues in Clinical Child Psychology,* Ollendick TH, King NJ, eds. New York: Plenum, pp 293–315

Silverman WK, Albano AM (1996), *The Anxiety Disorders Interview Schedule for Children for DSM-IV: (Child and Parent Versions).* San Antonio, TX: Psychological Corporation

Silverman WK, Eisen AR (1992), Age differences in the reliability of parent and child reports of child anxious symptomatology using a structured interview. *J Am Acad Child Adolesc Psychiatry* 31:117–124

Silverman WK, Ginsburg GS (1998), Anxiety disorders. In: *Handbook of Child Psychopathology,* Ollendick TH, Hersen M, eds. New York: Plenum, pp 239–268

Silverman WK, Kurtines WM (1996), *Anxiety and Phobic Disorders: A Pragmatic Approach.* New York: Plenum

Silverman WK, Kurtines WM, Ginsburg GS, Weems CF, Lumpkin PW, Carmichael DH (1999a), Treating anxiety disorders in children with group cognitive-behavioral therapy: a randomized clinical trial. *J Consult Clin Psychol* 67:995–1003

Silverman WK, Kurtines WM, Ginsburg GS, Weems CF, Rabian B, Serafini LT (1999b), Contingency management, self control, and education support in the treatment of childhood phobic disorders: a randomized clinical trial. *J Consult Clin Psychol* 67:675–687

Silverman WK, Nelles WB (1988), The Anxiety Disorders Interview Schedule for Children. *J Am Acad Child Adolesc Psychiatry* 27:772–778

Silverman WK, Rabian B (1995), Test-retest reliability of *DSM-III-R* childhood anxiety disorders symptoms using the Anxiety Disorders Interview Schedule for Children. *J Anxiety Disord* 9:139–150

Spence SH, Donovan C, Brechman-Toussaint M (2000), The treatment of childhood social phobia: the effectiveness of a social skills training-based, cognitive behavioural intervention, with and without parental involvement. *J Child Psychol Psychiatry* 41:713–726