1998

The predictive validity of parent and teacher reports of ADHD symptoms

Thomas J. Power

Brian J. Doherty

Susan Panichelli-Mindel
Philadelphia College of Osteopathic Medicine, susanmi@pcom.edu

James L. Karustis

Ricaro B. Eiraldi

See next page for additional authors

The Predictive Validity of Parent and Teacher Reports of ADHD Symptoms

By: Thomas J. Power, Brian J. Doherty, Susan M. Panichelli-Mindel, James L. Karustis, Ricardo B. Eiraldi, Arthur D. Anastopoulos, and George J. DuPaul

Power, T.J., Doherty, B.J., Panichelli-Mindel, S.M., Karustis, J.L., Eiraldi, R.B., Anastopoulos, A.D., & DuPaul, G.J. (1998). The predictive validity of parent and teacher reports of ADHD symptoms. *Journal of Psychopathology and Behavioral Assessment, 20*, 57-81.

Abstract:

The objectives were to evaluate the ability of the Inattention and Hyperactivity-Impulsivity factors of the ADHD Rating Scale-IV to differentiate children with ADHD from a control group and to discriminate children with different subtypes of ADHD. Also, we sought to determine optimal cutoff scores on the teacher and parent versions of this scale for making diagnostic decisions about ADHD. In a sample of 92 boys and girls 6 to 14 years of age referred to a regional ADHD program, we assessed ADHD diagnostic status using categorical and dimensional approaches as well as parent- and teacher-report measures. Logistic regression analyses showed that the Inattention and Hyperactivity-Impulsivity factors of the ADHD Rating Scale-IV were effective in discriminating children with ADHD from a control group and differentiating children with ADHD, Combined Type from ADHD, Inattentive Type. Although both teacher and parent ratings were significantly predictive of diagnostic status, teacher ratings made a stronger contribution to the prediction of subtype membership. Using symptom utility estimates, optimal cutoff scores on the Inattention and Hyperactivity-Impulsivity scales for predicting subtypes of ADHD were determined.
KEY WORDS: attention deficit disorder; hyperactivity; clinical prediction.

INTRODUCTION

A major concern of researchers in the field of attention-deficit/hyperactivity disorder (ADHD) over the past 20 years has been the identification of core symptom clusters. Based upon factor analytic studies (Bauermeister, Alegria, Bird, Rubio-Stipec, & Canino, 1992; DuPaul, 1991; Healey *et al.*, 1993; Lahey *et al.*, 1988) and field trial research, the Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition (DSM-IV) subdivided symptoms of ADHD into two dimensions: Inattention and Hyperactivity- Impulsivity (American Psychiatric Association, 1994). Symptoms were assigned to each factor on the basis of a thorough review of the literature and extensive field testing involving symptom utility estimates (Frick *et al.*, 1994). Symptom utility estimates were also used to determine the optimal number of symptoms needed on each cluster to predict functional impairment and the presence of an attention-deficit disorder with or without hyperactivity (Lahey *et al.*, 1994). On the basis of these field trials studies, guidelines

were established for diagnosing three subtypes of ADHD: ADHD, predominantly Inattentive type (ADHD/I); ADHD, predominantly Hyperactive-Impulsive type (ADHD/HI); and ADHD, Combined Type (ADHD/COM).

Assessing ADHD with Structured Interviews

Structured interviews are commonly used in research and clinical practice to determine whether individuals meet criteria for ADHD and other psychiatric disorders according to the DSM-IV. Structured interviews employ a categorical approach to assessment: disorders are categorized as present or absent based upon diagnostic criteria derived by a panel of experts after careful consideration of research literature. Benefits of structured interviews based on DSM-IV are that (a) items map directly to DSM-IV criteria, (b) the structured format enables clinicians to diagnose disorders in a highly reliable manner (e.g., see Chen, Faraone, Biederman, & Tsuang, 1994; Eiraldi, Power, & Nezu, 1997), (c) clinicians are directed not only to assess whether specific symptoms are present but also to determine the level of functional impairment resulting from the symptoms, and (d) extensive training is not required for administration.

One potential problem with structured interviews, however, is the lack of guidelines for determining symptom severity as a function of informant and child age and gender. Failure to account for variations in reporting based on these factors is significant, given that levels of inattention and hyperactivity-impulsivity have been shown to vary markedly as a function of child age and gender, and whether the informant is a parent, teacher, or self-reporter (Achenbach, 1991a; Arnold, 1996; Goyette, Conners, & Ulrich, 1978; Trites, Blouin, & LaPrade, 1982).

Clinical Utility of Behavior Rating Scales

Behavior rating scales, in contrast to structured interviews, employ a dimensional method that provides an assessment of behavior and emotional difficulties along a continuum from normal to abnormal, without clear delimitation between the presence or absence of disorder (Achenbach & McConaughy, 1996). Benefits of behavior rating scales are that (a) most are norm-referenced, permitting an evaluation of the severity of a symptom cluster compared to individuals of similar age and gender; (b) many have been developed for multiple informants, including parents, teachers, and self, permitting an assessment of situational variability in symptoms; (c) the leading rating scales, such as the Child Behavior Checklist (Achenbach, 1991a) and the Behavior Assessment System for Children (Reynolds & Kamphaus, 1993) have been demonstrated to have excellent psychometric properties; and (d) they are efficient for clinicians to use. A noteworthy limitation of rating scales, however, is that they are often difficult to use for the purpose of diagnostic assessment. Research is now being conducted to determine optimal cutoff scores on rating scales for determining the presence or absence of a disorder (Biederman *et al.*, 1993; Chen *et al*, 1994), but this information generally is not readily available in rating scale manuals. Another limitation of behavior rating scales is that they are vulnerable to rater bias (Abikoff, Courney, Pelham, & Koplewicz, 1993; Reid & Maag, 1994).

Using Rating Scales to Assess ADHD

Numerous rating scales have been developed to assess children for ADHD. Broad-band rating scales, designed to assess a wide range of externalizing and internalizing problems, usually include subscales useful in assessing ADHD. Narrow-band scales, developed specifically for the purpose of evaluating symptoms related to ADHD**,** are also available. The utility of many of the

broad- and narrow-band measures has diminished with the publication of DSM-IV Other limitations of existing instruments include inappropriate use of factor analytic methods to determine construct validity (Reid, 1995; Taylor & Sandberg, 1984) and, in some instances, insufficient evidence of psychometric properties.

The ADHD Rating Scale-IV (DuPaul, Anastopoulos, Power, Murphy, & Barkley, 1994), which contains a home and school version, was developed to reflect symptoms of ADHD as described in the DSM-IV. Factor analyses of this scale, using principal axis factoring with oblique rotation instead of the standard principal-components analysis with varimax rotation because of the high correlation between factors, have supported the two-factor solution in the DSM-IV (DuPaul *et al.,* 1997; DuPaul *et al.,* 1998). Also, strong evidence has been provided that the home and school versions of this scale have acceptable reliability and criterion-related validity (DuPaul, Power, McGoey, Ikeda, and Anastopoulos, in press). Further an extensive national standardization of this measure provides for normative comparison according to gender, age, and informant. Although the ADHD Rating Scale-IV appears to be a promising measure for the assessment of this disorder, research regarding the ability of this instrument to predict ADHD subtype membership is needed to determine its utility as a diagnostic assessment tool.

Purpose of Study

The purpose of this study was to evaluate the predictive validity and diagnostic utility of parent and teacher reports of ADHD symptoms, as defined by DSM-IV, using the recently developed ADHD Rating Scale-IV in a clinic-based setting. More specifically, we sought to determine the ability of the Inattention factor of the ADHD Rating Scale-IV to differentiate (a) children with ADHD/I from a clinical control group without any subtype of ADHD and (b) children with ADHD/COM from clinical controls. In addition, we investigated the ability of the Hyperactivity—Impulsivity factor to distinguish (a) children with ADHD/COM from clinical controls and (b) children with ADHD/I from those with ADHD/COM. Further, given the lack of research on the diagnostic utility of behavior rating scales, we sought to demonstrate methods for evaluating the utility of rating scales in clinical prediction.

METHOD

Participants

The sample consisted of consecutive referrals to the ADHD Evaluation and Treatment Program of a regional pediatric hospital located in a large metropolitan area in the North East. Children were referred for initial evaluation or reevaluation of ADHD. Participants met the following *inclusion* criteria: (a) completion by parents and teachers of the ADHD Rating Scale-IV and the diagnostic measures described below and (b) an estimated IQ of 80 or above on the Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990). Children were excluded if they presented with evidence of pervasive developmental disorder, a psychotic disorder, or a progressive neurological disorder. Also, children were excluded if they were on psychotropic medication for ADHD or related disorders within 6 months of the time of evaluation.

A total of 92 children, 24 girls (26.1%), participated in the study. Participants ranged in age from 6.0 to 14.9 years (M = 9.0, SD = 2.2). Grade levels ranged from kindergarten through eight; 73% were in grades 1 through 4. The distribution of ethnic groups represented was 21.7% African American, 3.3% Latino/Hispanic, and 75.0% Caucasian. The range of socioeconomic levels as

assessed by the Four Factor Index of Social Status (Hollingshead, 1975) was as follows: 3.2% in Category I (unskilled laborers), 14.2% in Category II (machine operators, semiskilled workers), 25.0% in Category III (skilled craftsman, clerical, sales workers), 40.2% in Category IV (small business owners, technicians), and 17.4% in Category V (major business owners, professionals). On the KBIT, the sample achieved mean scores of 103.1 (SD = 11.9) on the Vocabulary scale, 100.5 (M = 11.7) on Matrices, and 101.9 (SD = 11.1) on the Composite. Twenty-six percent of the sample were enrolled in special education. Twelve children were excluded because they had been on psychotropic medication within 6 months of the evaluation.

Measures

Diagnostic Interview for Children and Adolescents-Revised (DICA-R). The DICA-R (Reich, Shayka, & Taibleson, 1991) is a structured diagnostic interview designed to evaluate symptoms of child psychopathology using criteria from DSM-III-R (American Psychiatric Association, 1987). The DICA-R has been shown to have acceptable internal consistency and discriminant validity. Revisions in the DICA-R were made to reflect changes in the criteria for ADHD as delineated in DSM-IV (Eiraldi *et al.,* 1997). The interrater reliability of this updated version of the DICA-R has been shown to be well above the acceptable range (Eiraldi *et al.,* 1997). The parent version of the DICA-R was used in this study.

Child Behavior Checklist (CBCL). The CBCL (Achenbach, 1991a) is a 112-item, parent-report measure designed to assess multiple domains of children's behavioral and emotional functioning. The Attention Problems subscale, commonly used in the assessment of ADHD, has been shown to demonstrate adequate reliability and criterion-related validity. As indicated, a T-score of 60 on this subscale optimizes the differentiation of children with ADHD versus controls in a clinic-based setting (Chen *et al.,* 1994).

Child Attention Profile (CAP). The CAP is a 12-item teacher-report measure consisting of items from the Teacher Report Form (Achenbach, 1991b) assessing symptoms related to ADHD. A factor analysis of the CAP derived separate factors for Inattention and Overactivity (Edelbrock, 1988, published in Barkley, 1990). Normative data from the large standardization sample of the TRF are available (Barkley, 1990). A cutoff score at the 93rd percentile on the Inattention and Overactivity factors of the CAP has been used in research to differentiate children with attention deficit disorder into those with and without hyperactivity (e.g., see Barkley, DuPaul, & McMurray, 1991).

ADHD Rating Scale-IV. The ADHD Rating Scale-IV (DuPaul *et al.,* 1994) is an 18-item parent and teacher rating scale designed to assess symptoms of inattention and hyperactivity-impulsivity, as reflected in the DSM- IV. Items on this scale were taken directly from DSM-IV, although in many cases they were reworded for clarity and brevity. Each item is rated on a 4-point scale (0 = not at all, rarely; 1 = sometimes; 2 = often; 3 = very often). Factor analyses of both the home and the school versions of the ADHD Rating Scale-IV have shown that the factor structure of this scale confirms the theoretical structure described in DSM-IV. Parent and teacher ratings on this measure were found to be internally consistent, to be stable over a 4-week period, and to correlate significantly with observations of classroom behavior and task accuracy as well as with corresponding subscales on the Conners (1989) Parent and Teacher Rating Scales. Nor-

mative data for both the parent and the teacher versions have been collected in a large national sample stratified according to geographic region and ethnic group (DuPaul *et al.*, 1997, 1998).[1]

Assignment to Diagnostic and Control Groups

Children were assigned to a diagnostic group or clinical control group based on their scores on a multimethod assessment battery including the parent version of the DICA-R, the parent-rated CBCL, and the teacher- rated CAP The use of categorical and dimensional as well as parent and teacher report measures is consistent with recommended practice in the diagnostic assessment of ADHD (Biederman *et al.*, 1995; Loeber, Green, & Lahey, 1990). Children were categorized as having ADHD/I if they demonstrated the following: (a) a DICA-R diagnosis of ADHD/I, *and* (b) a T-score of 60 or above on the Attention Problems factor of the CBCL, *and* (c) a score on the Inattention subscale of the CAP of greater than or equal to the 93rd percentile. Children were diagnosed as having ADHD/HI if they demonstrated (a) a DICA-R diagnosis of ADHD/HI, *and* (b) a T-score of 60 or above on the Attention Problems factor of the CBCL, *and* (c) a score on the Overactivity subscale of the CAP of greater than or equal to the 93rd percentile. Children were diagnosed as having ADHD/COM if they demonstrated (a) a DICA-R diagnosis of ADHD/COM, *and* (b) a T-score of 60 or above on the Attention Problems factor of the CBCL, *and* (c) a score on the Overactivity and Inattention subscales of the CAP of greater than or equal to the 93rd percentile. Children were assigned to a clinical control group if they did not meet criteria for any subtype of ADHD, as defined by the above criteria.

Based on these criteria, 30 children were classified as having ADHD/I, 2 had ADHD/HI, and 25 had ADHD/COM. Thirty-five children in the sample did not meet criteria for any subtype of ADHD. Given that only two children had ADHD/HI, children belonging to this subtype were not included in statistical analyses.

The clinical groups and control group were compared to determine if there were differences among them with regard to gender, age, and pattern of comorbidity. The presence of comorbid diagnoses was determined based on the DICA-R administered to the parents. Table I provides information about age, gender, comorbid diagnoses, and special education placement for each of the groups. Chi-square analyses revealed that the groups differed with regard to the presence of Conduct Disorders ($x^2 = 8.07$, $p < .001$); the ADHD/COM group had three children with Conduct Disorder, and each of the other two groups had no children with this disorder. There were no differences between the groups with regard to gender, other comorbid psychiatric diagnosis, or special education placement. The ratio of males to females was 2:1 for ADHD/I, 5:1 for ADHD/COM, and 2.5:1 for the control group. A one-way analysis of variance failed to detect differences among the groups with regard to age.

Procedures

Parents and teachers were requested to complete and return via mail the CBCL, TRF, and home and school versions of the ADHD Rating Scale- IV prior to the initial clinic visit. In cases where the parents and/or teachers did not do so, they were requested to complete the measures before a fol-

[1] For scoring and normative information about the ADHD Rating Scale-IV, the reader is referred to DuPaul *et al.* (1997, 1998).

Table I. Composition of the Clinical and Control Groups

	ADHD, inattentive	ADHD, combined	Controls
Gender (% male)	67%	84%	71%
Age[a]	9.5 (2.4)	8.4 (2.0)	8.9 (1.9)
Oppositional defiant[b]	6 (20%)	5 (20%)	6 (17%)
Conduct disorder	0 (0%)	3 (12%)	0 (0%)
Anxiety disorders	3 (10%)	1 (4%)	1 (3%)
Mood disorders	2 (7%)	2 (7%)	0 (0%)
Special ed. placement	6 (20%)	4 (16%)	10 (29%)

[a]Age is presented in years with standard deviations in parentheses.
[b]The frequency of children with comorbid diagnoses or special education placement in each group is presented, with the corresponding percentage value in parentheses.

low-up meeting. The DICA-R was conducted by a doctoral-level psychology clinician or an advanced doctoral candidate in psychology at the initial clinic visit. Administrations of the DICA-R were audiotaped; 23% of the audiotapes were selected at random and reviewed by another clinician to establish interscorer reliability. Kappa coefficients were .87 for a diagnosis of ADHD/COM, .81 for ADHD/I, .91 for oppositional defiant disorder, .78 for conduct disorder, .84 for anxiety disorder, and .65 for mood disorders, all of which are above the established limits of acceptability (Hartman, 1982).

Statistical Analyses

For each of the data analytic procedures used, raw scores were transformed to percentile scores based on normative data derived from the national standardization of the ADHD Rating Scale-IV (DuPaul *et al.*, 1997, 1998) to adjust for variations in Inattention and Hyperactivity—Impulsivity due to age and gender. Transformation to T-scores was deemed inappropriate given that the distributions of raw scores on these factors in the normative sample were highly skewed.

Receiver operating characteristic (ROC) analyses and logistical regression methods were used to evaluate (a) whether parent and teacher ratings used separately were able to differentiate clinical groups from the control group and clinical groups from each other and (b) whether a strategy cornbining parent and teacher ratings achieved a higher level of diagnostic accuracy than a single informant approach. The ROC analysis yields an index of area under the curve (AUC), which varies from .50, referring to a scale that is useless for prediction, to 1.00, a scale with a perfect rate of prediction (Chen *et al.*, 1994). A ROC analysis was conducted for parent and teacher ratings separately and in combination to determine which scale(s) generated the highest AUC and therefore had the best discriminating power.

Logistic regression generates a correlation-type statistic (R) reflecting the unique association between each predictor variable and the criterion variable (i.e., diagnostic group membership) and a chi-square statistic indicating whether the logistic model results in a significant level of prediction and whether each additional variable entered into the model results in an improvement

in prediction. Finally, the Hosmer—Lemeshow (1989) test was used to determine whether the logistic model adequately fit the data.

Symptom utility estimates were used to determine optimal cutoff scores on the Inattention factor of the ADHD Rating Scale-IV for differentiating (a) children with ADHD/I from a clinical control group without any subtype of ADHD and (b) children with ADHD/COM from clinical controls. Symptom utility estimates were also used to determine the optimal cutoff scores on the Hyperactivity—Impulsivity factor for differentiating (a) children with ADHD/COM from clinical controls and (b) children with ADHD/I from those with ADHD/COM.

In determining the predictive validity of measures used to make diagnostic decisions, sensitivity and specificity are important, but it is usually more useful to examine positive predictive power (PPP) and negative predictive power (NPP) (Chen *et al.*, 1994; Widiger, Hurt, Frances, Clarkin, & Gilmore, 1984). PPP refers to the probability that a child has a disorder given the presence of a symptom or score above a designated cutoff point on a diagnostic measure. NPP refers to the probability that a child does not have a disorder given the absence of a behavior or a score below a selected cutoff point on a diagnostic measure.

A problem with using PPP and NPP is that these statistics are highly sensitive to base rates of symptoms (cutoffs) and diagnoses in a sample (Verhulst & Koot, 1992). Thus, to make comparisons about the clinical utility of one cutoff score versus another, each of which may have a different base rate in the sample, it is necessary to employ a kappa statistic that corrects for the number of accurate predictions based on chance alone. Formulas for the kappa statistics used in this study to correct PPP (cPPP) and NPP (cNPP) are reported by Frick *et al.* (1994).

RESULTS

Predictive Validity of the Inattention Scale

Teacher ratings on the ADHD Rating Scale-IV were entered alone as a predictor and the resultant model was able to differentiate children with ADHD/I from those in the control group [$x^2(1) = 18.68$, $p < .0001$] and correctly classified 73.9% of the cases. The association between teacher ratings of Inattention and diagnostic status was moderate ($R = .38$). When parent ratings on the ADHD Rating Scale-IV were entered alone in a logistic regression analysis, the model generated was able to predict the presence of ADHD/I [$x^2(1) = 10.72$, $p < .005$] and correctly classified 67.7% of the cases. The association between parent ratings of Inattention and diagnostic status was low ($R = .26$).

In a forward stepwise logistic regression analysis, teacher ratings entered the equation first. The addition of parent ratings resulted in a significant improvement of the logistic model ($x^2 = 6.85$, $p < .01$). In the model created with both sets of ratings entered jointly, the unique association between teacher ratings and diagnostic status was low to moderate ($R = .34$), and the unique relation between parent ratings and diagnosis was low ($R = .20$). The combination of teacher and parent ratings in the model correctly classified 72.3% of the cases, slightly less than the percentage predicted by teacher ratings alone. The Hosmer-Lemeshow goodnessof-fit test with both sets of ratings included indicated that the logistic model provided an adequate fit with the data [$x^2(7) = 2.84$, $p = .90$].

The results of the ROC analysis when the Inattention scale was used to differentiate children with ADHD/I from the control group revealed an AUC of .80 for teacher ratings and .70 for parent ratings. The AUC with both sets of ratings entered together increased to .84.

Next, logistic regression analyses were used to evaluate the ability of the Inattention scale to differentiate children with ADHD/COM from those in the control group. Teacher ratings were entered alone as a predictor and the resultant model was able to differentiate children with AD/HD/COM from those in the control group [$_x$²(1) = 24.11, $p < .0001$] and correctly classified 80.0% of the cases. The association between teacher ratings of Inattention and diagnostic status was moderate *(R* = .44). When parent ratings were entered alone in a logistic regression analysis, the model generated was able to predict the presence of ADHD/COM [$x^2(1) = 8.08$, $p < .005$] and correctly classified 61.7% of the cases. The association between parent ratings of Inattention and diagnostic status was low *(R* = .23). In a forward stepwise logistic regression analysis, teacher ratings only entered the equation; the addition of parent ratings did not result in a significant improvement of the logistic model. The Hosmer-Lemeshow goodness-of-fit test with teacher ratings alone entered indicated that the logistic model provided an adequate fit with the data [$x^2(4) = 2.01$, $p = .73$].

The results of the ROC analysis when the Inattention scale was used to differentiate children with ADHD/COM from the control group revealed an AUC of .84 for teacher ratings. Parent ratings were not added to the model given that they did not make a significant contribution to the prediction of diagnostic status.

Predictive Validity of the Hyperactivity-Impulsivity Scale

Logistic regression analyses were used to evaluate the ability of the Hyperactivity-Impulsivity scale to differentiate children with ADHD/COM from those in the control group. When teacher ratings were entered alone in a logistic regression analysis, the resultant model was able to differentiate children with ADHD/COM from those in the control group [$x^2(1) = 12.70$, $p < .001$] and correctly classified 65.0% of the cases. The association between teacher ratings of Hyperactivity-Impulsivity and diagnostic status was low to moderate *(R* = .32). When parent ratings were entered alone in a logistic regression analysis, the model generated was able to predict the presence of ADHD/COM [$x^2(1) = 6.95$, $p < .01$] and correctly classified 60.0% of the cases. The association between parent ratings of Hyperactivity-Impulsivity and diagnostic status was low *(R* = .22).

In a forward stepwise logistic regression analysis, teacher ratings entered the equation first. The addition of parent ratings resulted in significant improvement of the logistic model [$x^2(1) = 7.18$, $p < .01$]. In the model created with both sets of ratings entered jointly, the unique association between teacher ratings and diagnostic status was low to moderate *(R* = .32), and the unique relation between parent ratings and diagnosis was low *(R* = .22). The combination of teacher and parent ratings in the model correctly classified 75.0% of the cases. The Hosmer-Lemeshow goodness-of-fit test with both sets of ratings included indicated that the logistic model provided a good fit with the data [$x^2(7) = 8.79$, $p = .27$].

The results of the ROC analysis when the Hyperactivity-Impulsivity scale was used to differentiate children with ADHD/COM from the control group revealed an AUC of .74 for

teacher ratings and .66 for parent ratings. The AUC with both sets of ratings entered together increased to .81.

Next, logistic regression analyses were used to evaluate the ability of the Hyperactivity-Impulsivity scale to differentiate children with ADHD/COM from those with ADHD/I. When teacher ratings were entered alone in the analysis, the resultant model was able to differentiate children with ADHD/COM from those with ADHD/I [$x^2(1) = 30.14$, $p < .0001$] and correctly classified 83.6% of the cases. The association between teacher ratings of Hyperactivity-Impulsivity and diagnostic status was moderate *(R* = .46). When parent ratings were entered alone in a logistic regression analysis, the model generated was able to predict the presence of ADHD/COM [$x^2(1) = 7.69$, $p < .01$] and correctly classified 63.6% of the cases. The association between parent ratings of Hyperactivity-Impulsivity and diagnostic status was low *(R* = .24). In a forward stepwise logistic regression analysis, teacher ratings only entered the equation; the addition of parent ratings resulted in significant improvement of the logistic model. The Hosmer-Lemeshow goodness-of-fit test with both sets of ratings included indicated that the logistic model provided a good fit with the data [$x^2(4) = 1.02$, $p = .90$].

The results of the ROC analysis when the Hyperactivity-Impulsivity scale was used to differentiate children with ADHD/COM from ADHD/I revealed an AUC of .89 for teacher ratings. Parent ratings were not added to the model given that they did not make a significant contribution to the prediction of diagnostic status.

Clinical Utility of the Inattention Scale

Symptom utility estimates associated with a series of possible cutoff scores on the Inattention factor of the ADHD Rating Scale-IV, as rated by teachers and parents separately, are provided in Table II. This table indicates base rates, sensitivity, specificity, PPP, cPPP, NPP, and cNPP statistics associated with each cutoff score investigated in this study **(**80th, 85th, 90th, 93rd, and 98th). This table presents data regarding the ability of the Inattention scale, as rated by teachers and parents, to differentiate children with ADHD/I from controls as well as to differentiate children with ADHD/COM from the control group. The strategy for determining optimal thresholds was to identify the cutoff score associated with the highest level of accurate prediction (cPPP or cNPP) that resulted in a reasonable level of sensitivity or specificity. For the purposes of this study, a cutoff score was considered clinically useful if cPPP or cNPP was greater than or equal to .65 and the sensitivity or specificity was approximately .50 or greater.

The optimal cutoff score on the teacher rating scale for differentiating children with ADHD/I from controls (98th percentile) was only marginally useful: 85% of the children scoring at or above this cutoff had ADHD/I (cPPP = .73), but only 20% of children known to have ADHD/I scored at or above this level (sensitivity). The optimal cutoff on the teacher scale for ruling out ADHD/I (80th percentile) was also marginally useful; 84%

Table II. *Symptom Utility Estimates Associated with Cutoff Scores on the Inattention Factor*

Cutoff score	Base rate	Sensitivity	Specificity	PPP (cPPP)	NPP (cNPP)
			Differentiating ADHD/I from controls—teacher ratings		
≥ 98	.11	.20	.97	.85 (.73)	.59 (.10)
≥ 93	.37	.63	.86	.79 (.61)	.73 (.42)
≥ 90	.42	.67	.80	.74 (.52)	.74 (.43)
≥ 85	.54	.77	.66	.66 (.36)	.77 (.49)
≥ 80	.71	.90	.46	.59 (.23)	.84 (.66)
			Differentiating ADHD/COM from controls—teacher ratings		
≥ 98	.10	.20	.97	.83 (.71)	.63 (.11)
≥ 93	.38	.72	.86	.78 (.63)	.81 (.54)
≥ 90	.45	.80	.80	.74 (.56)	.85 (.64)
≥ 85	.55	.84	.66	.64 (.38)	.85 (.64)
≥ 80	.70	.92	.46	.55 (.22)	.89 (.73)
			Differentiating ADHD/I from controls—parent ratings		
≥ 98	.35	.47	.74	.61 (.27)	.62 (.17)
≥ 93	.66	.83	.49	.58 (.22)	.77 (.51)
≥ 90	.75	.93	.40	.57 (.20)	.88 (.73)
≥ 85	.77	.93	.37	.56 (.18)	.87 (.71)
≥ 80	.83	.93	.26	.52 (.11)	.82 (.61)
			Differentiating ADHD/COM from controls—parent ratings		
≥ 98	.27	.28	.74	.44 (.04)	.59 (.02)
≥ 93	.65	.84	.49	.54 (.21)	.81 (.54)
≥ 90	.73	.92	.40	.52 (.18)	.88 (.70)
≥ 85	.77	.96	.37	.52 (.18)	.93 (.83)
≥ 80	.85	1.00	.26	.49 (.13)	1.00 (1.00)

Note: Cutoff scores are percentile values. Base rate refers to the proportion of children in the clinical and control groups scoring at or above the specified cutoff score. PPP refers to positive predictive power and NPP refers to negative predictive power; cPPP and cNPP are kappa statistics used to correct PPP and NPP.

of children scoring below this cutoff had ADHD/I (cNPP = .66), and 46% of children who did not have ADHD scored below this cutoff (specificity).

Teacher ratings on the Inattention scale were also marginally useful in predicting children who had ADHD/COM. The optimal cutoff for ruling in ADHD/COM was the 98th percentile; the optimal cutoff for ruling out this subtype was the 80th percentile. As was the case when teacher ratings were used to differentiate ADHD/I from controls, cPPP and cNPP values were relatively high but sensitivity and specificity values were quite low.

Parent ratings of Inattention were not useful in predicting or ruling in a diagnosis of ADHD/I or ADHD/COM. For ruling out ADHD/I, the 90th percentile was most useful: cNPP was relatively high (.73), but only 40% of children who did not have ADHD fell below this cutoff. For ruling out ADHD/COM, the 85th percentile had some utility: cNPP was high (.83), but only 37% of children who did not have ADHD fell below this cutoff.

Given that the results of logistic regression analyses demonstrated that the integration of teacher and parent ratings on the Inattention scale was better than single informant cutoff scores in predicting group membership, at least for children who met criteria for ADHD/I, symptom utility

estimates for all possible combinations of parent and teacher ratings were computed. The combination cutoffs for the Inattention scale that had the highest level of utility for predicting ADHD/I and ADHD/COM are presented in Table III. In general, the combination cutoffs were better than the single scale thresholds in predicting or ruling in ADHD/I and ADHD/COM. For example, teacher ratings greater than or equal to the 90th percentile on the Inattention factor were associated with a diagnosis of ADHD/I in 74% of the cases (cPPP = .52). However, when teacher ratings greater than or equal to the 90th percentile were combined with parent ratings greater than or equal to the 93rd percentile, the rate of prediction improved to 85% (cPPP = .72) and sensitivity was still reasonably high (.57). The combination of parent and teacher ratings of Inattention was marginally useful in ruling out a diagnosis of ADHD/I; with teacher ratings below the 80th percentile and parent ratings below the 85th percentile, cNPP was moderate (.63) and specificity was .69.

With regard to differentiating children with ADHD/COM from controls, combination cutoffs involving teacher ratings at or above the 90th percentile and parent ratings at or above the 93rd percentile once again were the optimal combination. For instance, this combination resulted in a high level of cPPP (.77) and a sensitivity to 76% of the cases who met the criteria for ADHD/COM. For ruling out ADHD/COM, teacher ratings below the 80th percentile and parent ratings below the 85th percentile had a high degree of utility: cNPP was .73 and 69% of children without ADHD scored below these levels.

Clinical Utility of the Hyperactivity-Impulsivity Scale

Symptom utility estimates associated with a series of possible cutoff scores on the Hyperactivity—Impulsivity factor of the ADHD Rating Scale- IV, as rated by teachers and parents separately, are provided in Table IV This table presents data regarding the ability of the Hyperactivity—Impulsivity scale to differentiate children with ADHD/COM from controls as well as to differentiate children with ADHD/COM from those with ADHD/I.

Table III. Symptom Utility Estimates Associated with the Combination of Teacher (T) and Parent (P) Ratings on the Inattention Factor

Cutoff score	Base rate	Sensitivity	Specificity	PPP (cPPP)	NPP (cNPP)
Differentiating ADHD/I from controls					
$T \geq 80$, $P \geq 80$	.60	.83	.60	.64 (.33)	.81 (.58)
$T \geq 80$, $P \geq 85$	.55	.83	.69	.69 (.43)	.83 (.63)
$T \geq 80$, $P \geq 90$	.55	.83	.69	.69 (.43)	.83 (.63)
$T \geq 80$, $P \geq 93$	.49	.77	.74	.72 (.48)	.79 (.54)
$T \geq 80$, $P \geq 98$	.29	.43	.82	.68 (.41)	.63 (.20)
$T \geq 85$, $P \geq 80$	.46	.70	.74	.70 (.44)	.74 (.44)
$T \geq 85$, $P \geq 85$	.43	.70	.80	.75 (.54)	.76 (.47)
$T \geq 85$, $P \geq 90$	.43	.70	.80	.75 (.54)	.76 (.47)
$T \geq 85$, $P \geq 93$	.38	.67	.86	.80 (.63)	.75 (.46)
$T \geq 85$, $P \geq 98$	.23	.37	.89	.73 (.50)	.62 (.18)
$T \geq 90$, $P \geq 80$	.37	.60	.83	.75 (.54)	.71 (.37)
$T \geq 90$, $P \geq 85$	.35	.60	.86	.78 (.60)	.71 (.38)
$T \geq 90$, $P \geq 90$	.35	.60	.86	.78 (.60)	.71 (.38)
$T \geq 90$, $P \geq 93$	.31	.57	.91	.85 (.72)	.62 (.18)
$T \geq 90$, $P \geq 98$	.18	.33	.94	.83 (.69)	.62 (.18)
Differentiating ADHD/COM from controls					
$T \geq 80$, $P \geq 80$	.62	.92	.60	.62 (.35)	.91 (.79)
$T \geq 80$, $P \geq 85$	.55	.88	.69	.67 (.43)	.89 (.73)
$T \geq 80$, $P \geq 90$	.53	.84	.69	.66 (.41)	.86 (.66)
$T \geq 80$, $P \geq 93$	.50	.84	.74	.70 (.49)	.87 (.68)
$T \geq 80$, $P \geq 98$	.22	.28	.83	.54 (.21)	.62 (.08)
$T \geq 85$, $P \geq 80$	.50	.84	.74	.70 (.49)	.87 (.68)
$T \geq 85$, $P \geq 85$	.47	.84	.80	.75 (.57)	.88 (.70)
$T \geq 85$, $P \geq 90$	.45	.80	.80	.74 (.56)	.85 (.64)
$T \geq 85$, $P \geq 93$	.42	.80	.86	.80 (.66)	.86 (.66)
$T \geq 85$, $P \geq 98$	.17	.24	.89	.60 (.31)	.62 (.09)
$T \geq 90$, $P \geq 80$	.43	.80	.83	.77 (.60)	.85 (.65)
$T \geq 90$, $P \geq 85$	.42	.80	.86	.80 (.66)	.86 (.66)
$T \geq 90$, $P \geq 90$	.40	.76	.86	.79 (.64)	.83 (.60)
$T \geq 90$, $P \geq 93$	.37	.76	.91	.86 (.77)	.84 (.62)
$T \geq 90$, $P \geq 98$	.13	.24	.94	.75 (.57)	.63 (.12)

Teacher ratings of Hyperactivity-Impulsivity generally were not useful in predicting children who had ADHD/COM when this subtype was compared to controls, however, this scale was useful in ruling out a diagnosis of ADHD/COM. For instance, 90% of the children with teacher ratings below the 85th percentile did not have ADHD/COM (cNPP = .76), and 51% of the children who did not have ADHD/COM scored below this cutoff (specificity).

Teacher ratings on the Hyperactivity-Impulsivity scale were much more useful in differentiating children with ADHD/COM from those with

Table IV. Symptom Utility Estimates Associated with Cutoff Scores on the Hyperactivity–Impulsivity Factor

Cutoff score	Base rate	Sensitivity	Specificity	PPP (cPPP)	NPP (cNPP)
			Differentiating ADHD/COM from controls—teacher ratings		
≥ 98	.33	.48	.77	.60 (.31)	.68 (.22)
≥ 93	.47	.64	.66	.57 (.27)	.72 (.33)
≥ 90	.60	.88	.60	.61 (.33)	.88 (.70)
≥ 85	.67	.92	.51	.58 (.27)	.90 (.76)
≥ 80	.80	.96	.31	.50 (.14)	.92 (.80)
			Differentiating ADHD/COM from ADHD/I—teacher ratings		
≥ 98	.24	.48	.97	.92 (.86)	.69 (.32)
≥ 93	.35	.64	.90	.84 (.71)	.75 (.45)
≥ 90	.51	.88	.80	.79 (.61)	.89 (.76)
≥ 85	.60	.92	.67	.70 (.44)	.91 (.90)
≥ 80	.73	.96	.47	.60 (.27)	.93 (.85)
			Differentiating ADHD/COM from controls—parent ratings		
≥ 98	.45	.16	.63	.27 (−.34)	.47 (−.15)
≥ 93	.67	.53	.46	.46 (−.01)	.47 (−.01)
≥ 90	.68	.60	.43	.47 (.02)	.56 (.04)
≥ 85	.73	.67	.40	.49 (.05)	.58 (.10)
≥ 80	.78	.73	.26	.48 (.31)	.58 (.09)
			Differentiating ADHD/COM from ADHD/I—parent ratings		
≥ 98	.35	.56	.83	.74 (.52)	.69 (.33)
≥ 93	.67	.84	.47	.57 (.21)	.78 (.51)
≥ 90	.71	.84	.40	.54 (.15)	.75 (.45)
≥ 85	.78	.92	.33	.53 (.15)	.83 (.63)
≥ 80	.82	.92	.27	.51 (.10)	.80 (.56)

ADHD/I. The optimal cutoff score for predicting or ruling in ADHD/COM was the 98th percentile: 92% of children scoring at or above this level had ADHD/COM (cPPP = .86), and 48% of children with this subtype scored at or above the 98th percentile. The Hyperactivity—Impulsivity scale was also useful in ruling out ADHD/COM when children with the two subtypes of ADHD were compared. For instance, teacher ratings below the 85th percentile resulted in a cNPP of .80 and a specificity of .67. Parent ratings of Hyperactivity—Impulsivity were not useful in ruling in or ruling out a diagnosis of ADHD/COM when children with the two subtypes of ADHD were compared and when children with ADHD/COM were compared to controls.

Given that the integration of teacher and parent ratings on the Hyperactivity—Impulsivity scale was better than single informant cutoff scores in predicting children who met criteria for the ADHD, Combined Type, at least when children with ADHD/COM were contrasted with the control group, symptom utility estimates for all possible combinations of parent and teacher ratings were computed. The combination cutoffs with the most utility are presented in Table V. As expected based on the results of logistic regression analyses, the combination cutoffs were better than teacher ratings alone in differentiating ADHD/COM from the control group, but not in contrasting ADHD/COM from ADHD/I. When ADHD/COM was compared with the control group, the optimal combination was teacher ratings at or above the 90th percentile and parent ratings at or above the 98th

Table V. Symptom Utility Estimates Associated with the Combination of Teacher (T) and Parent (P) Ratings on the Hyperactivity–Impulsivity Factor

Cutoff score	Base rate	Sensitivity	Specificity	PPP (cPPP)	NPP (cNPP)
		Differentiating ADHD/COM from controls			
$T \geq 80$, $P \geq 80$	.63	.88	.54	.58 (.28)	.86 (.67)
$T \geq 80$, $P \geq 85$	.62	.88	.57	.59 (.31)	.87 (.69)
$T \geq 80$, $P \geq 90$	.57	.80	.60	.59 (.29)	.81 (.54)
$T \geq 80$, $P \geq 93$	.55	.80	.63	.61 (.32)	.81 (.56)
$T \geq 80$, $P \geq 98$	.37	.56	.77	.64 (.38)	.71 (.31)
$T \geq 85$, $P \geq 80$	.53	.84	.69	.66 (.41)	.86 (.66)
$T \geq 85$, $P \geq 85$	.52	.84	.71	.68 (.45)	.86 (.67)
$T \geq 85$, $P \geq 90$	.47	.76	.74	.68 (.45)	.81 (.55)
$T \geq 85$, $P \geq 93$	.45	.76	.77	.70 (.49)	.82 (.56)
$T \geq 85$, $P \geq 98$	.28	.52	.88	.76 (.60)	.72 (.33)
$T \geq 90$, $P \geq 80$	.48	.80	.74	.68 (.47)	.84 (.61)
$T \geq 90$, $P \geq 85$	.47	.80	.77	.71 (.51)	.84 (.63)
$T \geq 90$, $P \geq 90$	.42	.72	.80	.72 (.52)	.80 (.52)
$T \geq 90$, $P \geq 93$	.40	.72	.83	.75 (.57)	.81 (.53)
$T \geq 90$, $P \geq 98$	.25	.48	.91	.80 (.66)	.71 (.31)
		Differentiating ADHD/COM from ADHD/I			
$T \geq 80$, $P \geq 80$	.67	.88	.50	.59 (.26)	.83 (.63)
$T \geq 80$, $P \geq 85$	.67	.88	.50	.59 (.26)	.83 (.63)
$T \geq 80$, $P \geq 90$	.64	.80	.50	.57 (.21)	.75 (.45)
$T \geq 80$, $P \geq 93$	.60	.80	.57	.61 (.28)	.77 (.50)
$T \geq 80$, $P \geq 98$	.35	.56	.83	.74 (.52)	.69 (.33)
$T \geq 85$, $P \geq 80$	.55	.84	.70	.70 (.45)	.84 (.65)
$T \geq 85$, $P \geq 85$	.55	.84	.70	.70 (.45)	.84 (.65)
$T \geq 85$, $P \geq 90$	.51	.76	.70	.68 (.41)	.78 (.51)
$T \geq 85$, $P \geq 93$	.47	.76	.77	.73 (.51)	.79 (.54)
$T \geq 85$, $P \geq 98$	.27	.52	.93	.87 (.76)	.70 (.34)
$T \geq 90$, $P \geq 80$	.45	.80	.83	.80 (.63)	.83 (.63)
$T \geq 90$, $P \geq 85$	.45	.80	.83	.80 (.63)	.83 (.63)
$T \geq 90$, $P \geq 90$	.42	.72	.83	.78 (.60)	.78 (.52)
$T \geq 90$, $P \geq 93$	.40	.72	.87	.82 (.67)	.79 (.53)
$T \geq 90$, $P \geq 98$	.24	.48	.97	.92 (.86)	.69 (.32)

percentile; cPPP was moderate to high (.66) and 48% of the cases with ADHD/COM were identified. For ruling out ADHD/COM, both when ADHD/COM was compared to the control group and when ADHD/COM was compared with ADHD/I, the combination cutoffs were not as accurate as teacher ratings alone.

DISCUSSION

Predictive Validity of ADHD Rating Scale-IV

The results indicated that the Inattention and Hyperactivity-Impulsivity factors of the ADHD Rating Scale-IV were accurate and useful in predicting membership in the ADHD/COM and ADHD/I subtypes. Logistic regression analyses demonstrated that the Inattention scale was able to differentiate children with ADHD/I from a control group and children with ADHD/COM from controls. Also, logistic regression analyses indicated that the Hyperactivity-Impulsivity scale successfully differentiated children with ADHD/COM from controls and children with ADHD/COM from those with ADHD/I. Furthermore, the AUC values obtained when using the combination of teacher and parent ratings or teacher ratings alone were high and consistent with the results of other studies investigating the clinical utility of behavior ratings in predicting ADHD (Chen *et al.*, 1994; Rey, Morris-Yates, & Stainslaw, 1992).

The results indicated that teacher ratings on the ADHD Rating Scale- IV are extremely important in predicting subtype membership. Although parent ratings were also significantly predictive of diagnostic status, teacher ratings were better at predicting ADHD subtypes than parent ratings. As compared to parent ratings, teacher ratings were associated with (a) higher AUC indices in ROC analyses, (b) stronger relationships with diagnostic status in logistic regression analyses, and (c) higher cPPP values in symptom utility analyses.

The results suggested that an approach to diagnostic prediction that includes both teachers and parents may be superior to a single informant approach, although this finding was not demonstrated consistently. The results of logistic regression analyses indicated that the combination of teacher and parent ratings of Inattention was better than either informant alone in differentiating ADHD/I from controls, but not ADHD/COM from the control group. Logistic regression analyses also demonstrated that combining teacher and parent ratings of Hyperactivity-Impulsivity was better than each informant alone in differentiating ADHD/COM from controls, but not ADHD/COM from ADHD/I.

The findings of this study provide a method for using teacher and parent report information when making diagnostic decisions about ADHD. Optimal thresholds were chosen (a) to maximize the rate of accurate prediction (cPPP or cNPP) and (b) to maximize the number of children known to have ADHD/I and ADHD/COM who were accurately identified (sensitivity) or maximize the number of children known to be in the control group who were not identified as having ADHD (specificity). The optimal approach for using the Inattention scale to predict the presence of ADHD/I and ADHD/COM was one that combined teacher and parent ratings. The most useful combination appeared to be teacher ratings at or above the 90th percentile and parent ratings at or above the 93rd percentile. The optimal approach to using the Inattention scale to rule out the presence of ADHD was also a multiinformant one: teacher ratings less than the 80th percentile and parent ratings less than the 85th percentile had the most utility.

The Hyperactivity-Impulsivity scale had limited utility in differentiating ADHD/COM from the control group. However, this scale was useful in differentiating ADHD/COM from ADHD/I. The optimal approach to ruling in ADHD/COM when ADHD/COM and ADHD/I were compared was a single informant strategy using teacher ratings at or above the 98th percentile. The most useful approach to ruling out ADHD/COM when the two subtypes were compared also appeared to be a single informant approach using teacher ratings below the 85th percentile.

The optimal thresholds selected for purposes of diagnostic assessment ensure that the prevalence rate for ADHD in the general population is less than 10%. In actual practice, the selected thresholds would probably identify less than 5% of children, the prevalence of this disorder as estimated in the DSM-IV. For instance, the optimal threshold for using the Inattention scale to predict ADHD/I and ADHD/COM (teacher ratings at or above the 90th percentile, parent ratings at or above the 93rd percentile) by definition would identify less than 7% of the population. Taking into consideration that clinic-referred children are a preselected population who are likely to have higher levels of functional impairment than their peers with comparable scores who do not get referred, it is probable that the optimal cutoffs on the Inattention scale would identify considerably fewer than 7% of the population as having ADHD/I or ADHD/COM. Similarly, the optimal threshold on the Hyperactivity-Impulsivity factor for predicting ADHD/COM when the

two subtypes were compared (teacher ratings at or above the 98th percentile) would identify less than 2% of children as having ADHD/COM.

Limitations

The results of this study are generalizable to clinic-based settings serving children in the age range from 6 to 14 years. The findings may be less applicable in other settings, including school-based programs, or clinics serving preschool children or older adolescents. Use of the kappa correction statistics for PPP and **NPP** mitigates variations in predictability across settings that have differential base rates for symptoms and disorders (Chen *et al.,* 1994; Frick *et al.,* 1994). However, fundamental differences among settings with regard to referral agent (parent, physician, teacher) and type of functional impairment (family disruption, academic underachievement, classroom noncompliance) may result in cross-situational inconsistencies in prediction that are not entirely corrected by kappa statistics. For instance, in clinic-based programs where referral for services is strongly determined by parents' perception of a functional impairment and their willingness to pursue treatment, the optimal threshold for predicting a diagnosis of ADHD may be relatively high for parents and relatively low for teachers. In contrast, in schools where access to care is mediated largely by determination on the part of a teacher that there is functional impairment, the optimal threshold may be relatively high for teachers and relatively low for parents. Future research is needed to identify the optimal strategy for integrating parent and teacher report data in school-based settings and clinic-based programs serving a different population from the one in this study.

Relatedly, this study was conducted with a relatively homogeneous sample of children who were referred for problems with inattention and/or hyperactivity/impulsivity. A lack of diversity in the sample may have reduced differences between the clinical and control groups. Nonetheless, a relatively high degree of diagnostic accuracy was achieved in this sample.

The rate of externalizing disorders in the clinical groups, in particular the group with ADHD/COM, is low in comparison to other studies using DSM-IV to classify children (Eiraldi *et al.,* 1997; Morgan, Hynd, Riccio, & Hall, 1996). The fact that the Inattention scale performed equally well in predicting diagnoses of ADHD/COM and ADHD/I in this study as well as a related study with a more typical rate of comorbid externalizing disorders (Power *et al.,* 1998) suggests that the results of this study are generalizable to clinic-based populations of children with ADHD. Nonetheless, future research is needed to determine the potential impact of demographic characteristics and rates of comorbidity on the predictability of behavior rating scales.

A strength of this study is that it included a relatively high percentage of girls (26%) and children of minority status, in particular African Americans (22%), groups that are often excluded from studies of this kind. However, the inclusion of members of specific subgroups in a sample does not ensure the applicability of the findings to each subgroup (Reid, 1995). Given that the sample consisted primarily of Caucasian boys, the applicability of the findings to girls and children of specific ethnic minority groups needs to be investigated. The number of girls and African American children participating in the study did not permit a separate analysis of the predictive validity of the ADHD Rating Scale-IV for each subgroup of children.

These findings are also limited by the validity of the gold standard used to determine diagnostic status. The strategy utilized to assign participants to diagnostic groups was guided by the emerging consensus that a comprehensive assessment of ADHD should include categorical and dimensional measures as well as teacher and parent reports of behavior. Selection of different measures or cutoff scores on the CBCL and CAP may have altered somewhat the results of this study. Regardless, the criteria used in this study to determine subtype membership were highly stringent. The advantage of this approach is that children assigned to each diagnostic group in all likelihood were assigned correctly. The disadvantage is that a certain percentage of children with ADHD may have been incorrectly assigned to the control group.

Clinical Implications

Although stringent research criteria were used in this study to validate the ADHD Rating Scale-IV, teacher and parent ratings on this measure appear to have considerable utility in differential diagnosis of children from 6 to 14 years of age referred to clinic-based settings. Given that the Inattention scale generally was more useful than the Hyperactivity-Impulsivity scale in differentiating children with ADHD from those in the control group, a helpful strategy is to decide first whether a referred child has a disorder of inattention, that is either ADHD/I or ADHD/COM, using the Inattention scale. The optimal cutoffs found in this study should be helpful in making this determination. Next, if the clinician determines that the child has ADHD/I or ADHD/COM, the Hyperactivity-Impulsivity scale can be very useful in identifying the most appropriate subtype.

It is important to note that 24% of children who actually had ADHD/COM and 43% of children who had ADHD/I were missed in this study using the optimal cutoffs on the Inattention scale determined in this study. Also, in 52% of the cases the Hyperactivity-Impulsivity scale failed to sort out who had ADHD/COM versus ADHD/I when the optimal cutoff on the teacher rating scale was used. Thus, clinicians need to be cautious in the use of teacher and parent ratings of DSM-IV symptoms in determining whether ADHD exists and what subtype is the most descriptive of the child. At this point, a strategy that incorporates teacher and parent scales along with other information derived from informants is recommended in the diagnostic assessment of ADHD.

Unfortunately, it is difficult to compare the clinical utility of this measure with that of other scales designed to assess ADHD because most researchers have not subjected their measure to the same degree of scrutiny with regard to evaluating predictive validity. In the future, researchers involved in developing behavior rating scales are encouraged to evaluate the clinical utility of their measure by examining its ability to predict diagnostic status as defined by classification systems such as DSM-IV.

Given that there were so few children with ADHD/HI in this sample, it was not possible to investigate the ability of the Hyperactivity-Impulsivity scale to differentiate children with ADHD/HI from controls. In the future, research with a much larger group of children who have ADHD/HI is needed to determine the optimal method of using the ADHD Rating Scale- IV to assist with diagnostic assessment of this subtype.

Conclusions

In this study teacher and parent ratings of DSM-IV symptoms related to ADHD, as assessed by the ADHD Rating Scale-IV, were examined to determine their ability to predict the presence and absence of ADHD/COM and ADHD/I, as diagnosed on the basis of a multimethod battery consisting of categorical and dimensional as well as teacher- and parent-report measures. The results demonstrated that the ADHD Rating Scale-IV was able to differentiate (a) children with ADHD/COM from a control group without ADHD, (b) children with ADHD/I from controls, and (c) children with ADHD/COM from those with ADHD/I. Although teacher and parent ratings were both able to predict subtype membership, teacher ratings generally were superior to parent ratings. The results suggested that a diagnostic strategy combining teacher and parent ratings may be superior to a single informant approach. This study was conducted in a clinic-based program for children with problems related to ADHD. Additional research is needed to establish the predictive validity of teacher and parent ratings of ADHD symptoms in other settings, most notably schools.

REFERENCES

Abikoff, H., Courtney, M,. Pelham, W. E., Jr., & Koplewicz, H. S. (1993). Teachers' ratings of disruptive behaviors: the influence of halo effects. *Journal of Abnormal Child Psychology, 21(5),* 519-533.

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile.* Burlington: University of Vermont Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Teacher's Report Form and 1991 Profile.* Burlington: University of Vermont Department of Psychiatry.

Achenbach, T. M., & McConaughy, S. H. (1996). Relations between DSM-IV and empirically-based assessment. *School Psychology Review, 25,* 329-341.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arnold, L. E. (1996). Sex differences in AD/HD: Conference summary. *Journal of Abnormal Child Psychology, 24,* 555-570.

Barkley, R. A. (1990). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment.* New York: Guilford.

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1991). Attention deficit disorder with and without hyperactivity: Clinical response to three dose levels of methylphenidate. *Pediatrics, 87,* 519-531.

Bauermeister, J. J., Alegria, M., Bird, H. R., Rubio-Stipec, M., & Canino, G. (1992). Are attentional-hyperactivity deficits unidimensional or multidimensional syndromes? Empirical findings from a community survey. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 423-431.

Biederman, J., Faraone, S. V., Doyle, A., Lehman, B. K, Kraus, I., Perrin, J., & Tsuang, M. T. (1993). Convergence of the Child Behavior Checklist with structured interview-based psychiatric diagnoses of AD/HD children with and without comorbidity. *Journal of Child Psychology and Psychiatry, 34,* 1241-1251.

Biederman, J., Wozniak, J., Kiely, K., Ablon, S., Faraone, S., Mick, E., Mundy, E., & Kraus, I. (1995). CBCL clinical scales discriminate prepubertal children with structured interview-derived

diagnosis of mania from those with ADlHD. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 464-471.
Chen, W. J., Faraone, S. V., Biederman, J., & Tsuang, M. T. (1994). Diagnostic accuracy of the Child Behavior Checklist Scales for attention-deficit hyperactivity disorder: A receiver operating characteristic analysis. *Journal of Consulting and Clinical Psychology, 62,* 1017-1025.
Conners, C. K. (1989). *Conners' Rating Scales manual.* North Tonawanda, NY: Multi-Health Systems.
DuPaul, G. J. (1991). Parent and teacher ratings of ADlHD symptoms: Psychometric properties in a community-based sample. *Journal of Clinical Child Psychology, 20,* 245-253.
DuPaul, G. J., & Stoner, G. (1994). *ADHD in the schools: Assessment and intervention strategies.* New York: Guilford.
DuPaul, G. J., Anastopoulos, A. D., Power, T. J., Murphy, K., & Barkley, R. A. (1994). *ADHD Rating Scale-IV,* Unpublished rating scale.
DuPaul, G. J., Power, T. J., Anastopoulos, A. D., Reid, R., McGoey, K., & Ikeda, M. J. (1997). Teacher ratings of attention-deficitlhyperactivity disorder symptoms: Factor structure, normative data, and psychometric properties. *Psychological Assessment,* 9, 436-444.
DuPaul, G. J., Anastopolous, A. D., Power, T. J., *Reid,* R., McGoey, K. E., Ikeda, M. J., & McGoey, E. (1998). Parent ratings of attention-deficitlhyperactivity disorder symptoms: Factor structure and normative *data. Journal of Psychopathology and Behavioral Assessment, 20,* 83-102.
DuPaul, G. J., Power, T. J., McGoey, K. E., Ikeda, M. J., & Anastopoulos, A. D. (in press). Reliability and validity of parent and teacher ratings of attention-deficit/hyperactivity disorder symptoms. *Journal of Psychoeducational Assessment.*
Eiraldi, R. B., Power, T. J., & Nezu, C. M. (1997). Patterns of comorbidity associated with subtypes of attention-deficit/hyperactivity disorder among 6-12 year old children. *Journal of the American Academy of Child and Adolescent Psychiatry,* 36, 503-514.
Frick, P. J., Lahey, B. B., Applegate, B., Kerdyck, L., 011endick, T., Hynd, G. W., Garfinkel, B., Greenhill, L., Biederman, J., Barkley, R. A., McBurnett, K., Newcorn, J., & Waldman, I. (1994). DSM-IV field trials for the disruptive behavior disorders: Symptom utility estimates. *Journal of the American Academy of Child and Adolescent Psychiatry,* 33, 529-539.
Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data on revised Conners parent and teacher rating scales. *Journal of Abnormal Child Psychology, 6,* 221-236.
Hartman, D. P. (1982). *Using observers to study behavior: New directions for methodology of social and behavioral science.* San Francisco, CA: Jossey-Bass.
Healey, J. M., Newcorn, J. H., Halperin, J. M., Wolf, L. E., Pascualvaca, D. M., Schmeidler, J., & O'Brien, J. D. (1993). The factor structure of AD/HD items in DSM-III-R: Internal consistency and external validation. *Journal of Abnormal Child Psychology, 21,* 441-453.
Hollingshead, A. B. (1975). *Four factor index of social status.* New Haven, CT: Yale University.
Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression.* Ncw York: Wiley.
Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test.* Circle Pines, MN: American Guidance Service.
Lahey, B. B., Pelham, W. E., Schaughency, E. A., Atkins, M. S., Murphy, H. A., Hynd, G., Russo, M., Hartdagan, S., & Lorys-Vernon, A. (1988). Dimensions and types of attention deficit disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 2Z* 330-335.

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, B. W., Barkley, R. A., Newcorn, J., Jensen, P., Richters, J., Garfinkel, B., Kerdyk, L., Frick, P. J., 011endick, T., Perez, D., Hart, E. L., Waldman, I., & Schaffer, D. (1994). DSM-IV field trials for attention-deficit hyperactivity disorder in children and adolescents. *American Journal of Psychiatry, 151,* 1673-1685.

Loeber, R., Green, S. M., & Lahey, B. B. (1990). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology, 1,* 317-337.

Morgan, A. E., Hynd, G. W., Riccio, C. A., & Hall, J. (1996). Validity of DSM-IV ADHD predominantly Inattentive and Combined types: Relationship to previous DSM diagnoses/subtype differences. *Journal of the American Academy of Child and Adolescent Psychiatry,* 35, 325-333.

Power, T. J., Andrews, T. J., Eiraldi, R. B., Doherty, B. J., Ikeda, M. J., DuPaul, G. J., & Landau, S. (1998). *Evaluating ADHD using multiple informants: The incremental utility of combining teacher with parent reports* (manuscript submitted for publication),

Reich, W., Shayka, M. A., & Taibleson, C. (1991). *Diagnostic Interview for Children and Adolescents-DSM-III-R Version (Parent Form).* St. Louis, MO: Washington University Division of Child Psychiatry.

Reid, R. (1995). Assessment of AD/HD with culturally different groups: The use of behavior rating scales. *School Psychology Review, 24,* 537-560.

Reid, R., & Maag, J. W. (1994). How many fidgets in a pretty much: A critique of behavior rating scales for identifying students with ADHD. *Journal of School Psychology, 32,* 339-354.

Rey, J. M., Morris-Yates, A., & Stainslaw, H. (1992). Measuring the accuracy of diagnostic tests using receiver operating characteristics (ROC) analysis. *International Journal of Methods in Psychiatric Research, 2,* 39-50.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Manual for the Behavior Assessment System for Children.* Circle Pines, MN: American Guidance Service.

Taylor, E., & Sandberg, S. (1984). Hyperactive behavior in English school children: A questionnaire survey. *Journal of Abnormal Child Psychology, 12,* 143-156.

Trites, R. L., Blouin, A. G., & LaPrade, K. (1982). Factor analysis of the Conners Teacher Rating Scale based on a large normative sample. *Journal of Consulting and Clinical Psychology, 50,* 615-623.

Verhulst, F. C., & Koot, H. M. (1992). *Child psychiatric epidemiology: Concepts, methods, and findings.* Newbury Park, CA: Sage.

Widiger, T. A., Hurt, S. W., Frances, A., Clarkin, J. F., & *Gilmore, M.* (1984). *Diagnostic* efficiency and DSM-III. *Archives of General Psychiatry, 41,* 1005-1012.