

2006

ADHD Rating Scale-IV: Checklists, Norms, and Clinical Interpretation

Danielle N. Pappas

General Description

The *ADHD Rating Scale-IV: Checklist, norms, and clinical interpretation* (ADHD Rating Scale-IV; DuPaul, Power, Anastopoulos, & Reid, 1998) is a norm-referenced checklist that measures the symptoms of attention deficit/hyperactivity disorder (ADHD) according to the diagnostic criteria of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV;* American Psychiatric Association, 1994). The diagnostic criteria changed from the *DSM-III* to the *DSM-IV*. These changes reflect the factor analyses of teacher ratings of ADHD symptoms. This includes two separate factors of Inattention and Hyperactivity-Impulsivity. The ADHD Rating Scale-IV was designed to incorporate these changes. The purpose of this scale is to provide clinicians with a means of gathering information regarding the frequency of certain behaviors from parents and teachers. This information may be used by clinicians to screen, diagnose, or evaluate treatment of ADHD. There is a home and school version of the ADHD Rating Scale-IV. The home version of the scale is also provided in Spanish. Each version is completed independently by the parent or teacher, who reports the frequency of the symptoms over the past 6 months (or the beginning of the school year if the teacher has not known the student for 6 months) on a 4-point Likert scale. The ADHD Rating Scale-IV is an 18-item questionnaire that takes approximately 5 minutes to complete. Normative data are provided for ages 5 to 18. The clinical trials on the normative data were conducted with 6- to 14-year-old participants.

Two subscales are distinguished on the ADHD Rating Scale-IV: Inattention and Hyperactivity-Impulsivity. The odd-numbered items represent the symptoms on the Inattention subscale, and the even-numbered items represent the symptoms on the Hyperactive-Impulsive subscale. Results are described in terms of the Inattention and Hyperactivity-Impulsivity subscales, the total score, and percentile ranks for each score. Materials consist of the examiner's manual, rating scales, and scoring sheets.

Specific Description

Each statement on the ADHD Rating Scale-IV is rated on a 4-point Likert scale using *never or rarely, sometimes, often, or very often* as responses. The Inattention subscale is scored by summing the responses on all the odd-numbered items (Items 1, 3, 5, 7, 9, 11, 13, 15, 17). The Hyperactivity-Impulsivity subscale is scored by summing the responses on all the even-numbered items (Items 2, 4, 6, 8, 10, 12, 14, 16, 18). The Total Scale raw score is the total of the Inattention and Hyperactivity-Impulsivity subscale scores. Using the child's gender and age (and the appendix), raw scores are converted to percentile scores. The raw scores are then circled in the body of the scoring profile and can be converted to percentile scores on the outer columns of the profile. If the child's raw score is associated with more than one percentile rank, the examiner should circle and report the lowest one. If an item on the rating scale is missed, the respondent should be directed to provide an answer. If the respondent indicates a lack of knowledge on that particular behavior, the item should be omitted in deter-

mining the raw score. If 3 or more items are skipped by the respondent, extreme caution should be used in interpreting the scale for any of the intended purposes.

Inattention. As noted previously, this subscale consists of the odd-numbered items (total = 9). These items are designed to measure the child's attention level on tasks or play activities. Respondents are to rank to what degree they observed behaviors such as "does not seem to listen when spoken to directly" or "has difficulty organizing tasks and activities." These items correspond to the *DSM-IV-TR* description of inattention: fails to give close attention to details, makes careless mistakes, has difficulty sustaining attention, loses things necessary for tasks or activities, often easily distracted by extraneous stimuli, and often forgetful in daily activities (American Psychiatric Association, 2000).

Hyperactivity-Impulsivity. This subscale is composed of the even-numbered items (total = 9). These items measure the child's hyperactivity level and impulsivity level. Respondents are asked to what degree they observed behaviors such as "fidgets with hands or feet or squirms in seat," "talks excessively," or "has difficulty awaiting turn." These items correspond to the *DSM-IV-TR* description of hyperactivity: leaves seat often in the classroom or in other situations in which remaining seated is expected, runs about or climbs excessively, has difficulty playing or engaging in leisure activities quietly, and often "on the go" or "driven by a motor" (American Psychiatric Association, 2000).

Technical Adequacy

Standardization of home version. Data were collected on 2,000 randomly selected 4- to 20-year-olds from the overall sample used for factor analyses. Participants attended kindergarten through 12th grade. This sample approximated the U.S census data from 1990 in terms of ethnic group and region. The parent or guardian respondents ranged in age from 19 to 80 years. The respondents were predominantly self-reported as Caucasian (73.5%), with African American (14.2%), Latino (4.7%), Asian American (4.3%), Native American (0.7%), other (2.4%), and those with unspecified ethnic background (0.3%) also reported. Most of the respondents were mothers (85.6%), with fathers (11.3%), grandparents (1.2%), guardians (1.1%), and unspecified (1.0%) also responding. Most of the respondents were in the middle-class socioeconomic range with a median Hollingshead Index of 60.

Standardization of school version. Data were collected on 2,000 randomly selected students ranging in age from 4 to 19 years from the larger factor-analytic sample group. Participants attended kindergarten through 12th grade. Most students attended general education ($n = 1,816$) rather than special education ($n = 161$). The sample approximated the 1990 U.S. census data in terms of ethnic group and region of representation. There were 1,001 teacher respondents that were predominantly Caucasian (90.2%), with African American (6.1%), Latino (1.3%), Asian American (0.3%), Native American (0.1%), other (0.6%), and those with unspecified ethnic background (1.5%) also responding.

African American children were rated to exhibit more frequent ADHD symptoms by both parents and teachers as compared to Caucasian and Latino students. When socioeconomic status was statistically removed, this finding was maintained. Due to inefficient cell sizes

however, the manual does not provide separate ethnic group data. Therefore, clinicians are cautioned regarding the use of the ADHD Rating Scale-IV with African American children so as to avoid overidentification of this population.

Reliability. Test-retest reliability, internal consistency, and criterion-related validity were examined using a portion of the same sample of 71 students ages 5 to 17 years. Participants were randomly selected from two suburban school districts in eastern Pennsylvania and southern New Jersey. The sample was predominantly Caucasian, with African American, Latino, and Asian American children also participating. Only students in general education were included in the sample.

For internal consistency, coefficient alphas on the school version were calculated at .94 for total score, .96 for inattention, and .88 for hyperactivity-impulsivity. Coefficient alphas on the home version were calculated at .92 for total score, .86 for inattention, and .88 for hyperactivity-impulsivity.

Parents and teachers were asked to complete the ADHD Rating Scale-IV on two occasions 4 weeks apart to assess test-retest reliability. On the school version of the scale ($n = 52$), the Pearson product-moment correlation coefficients were .90 for total score, .90 for inattention, and .88 for hyperactivity-impulsivity. On the home version of the scale ($n = 43$), the Pearson product-moment correlation coefficients were .85 for total score, .78 for inattention, and .86 for hyperactivity-impulsivity.

Interrater agreement between parents and teachers were calculated on a sample of 62 students ages 5 to 17, kindergarten to 12th grade. Interrater agreement coefficients were .41 for total score, .45 for inattention, and .40 for hyperactivity-impulsivity. All the coefficients were in the moderate range, but these relatively low coefficients suggest that the behaviors characteristic of ADHD might be different across the home and school environment.

Interobserver agreement was conducted in the kindergarten through eighth-grade classrooms with a research assistant completing a behavioral observation using an adaptation of the ADHD Behavior Code (Barkley, 1990) on 3 separate days for a total of 45 minutes. A second observer was present for 30% of the sessions. Interobserver agreement averaged 88% across the two behavioral categories. These observations were used in assessing validity.

Validity. Criterion-referenced validity was addressed by determining the relationship between the ADHD Rating Scale-IV to the Conners Teacher Rating Scale-39 (CTRS-39), Conners Parent Rating Scale-48 (CPRS-48), behavior observations for off-task and fidgety behavior, and task accuracy. The school version of the ADHD Rating Scale-IV was correlated the strongest with the CTRS-39 Hyperactivity and Hyperactivity Index scores. After a Bonferroni correction was applied to control for Type I error rate, 20 out of 30 correlations remained statistically significant ($p < .05$). The school version Total Score and the Inattention subscale score was significantly correlated with direct behavior observations of off-task and fidgety behavior. The school version Total Score, Inattentive subscale score, and the Hyperactivity-Impulsivity subscale score were all significantly negatively correlated with accuracy on academic tasks. It was noted that significance of correlations between the school version of the ADHD Rating Scale-IV and the classroom observation measures was lower than the correlations with the CTRS-39.

In addition, 12 out of 18 correlation coefficients were statistically significant ($p < .05$) between the home version of the ADHD Rating Scale-IV and the CPRS-48 after a Bonferroni

correction was applied. Parent ratings were significantly correlated with teacher ratings on the CTRS-39 Hyperactivity and Conduct Problems factors and the Hyperactivity Index ($p < .05$ or better). Parent Inattention subscale scores and the Total Scores were correlated with the CTRS-39 Daydreams-Attention Problem factor. Parent ratings were not significantly correlated with classroom observations for fidgety or off-task behavior ($p < .05$). However, the Inattention subscale score and the Total score from the home version was significantly negatively correlated with academic accuracy ($p < .05$).

Discriminant validity was assessed using a sample of 92 children referred to the ADHD Evaluation and Treatment Program of the Children's Seashore House in Philadelphia. Upon initial assessment, the participants were placed in one of three groups: (a) having ADHD, predominantly inattentive subtype (ADHD/I; $n = 30$); (b) having ADHD, combined subtype (ADHD/COM; $n = 25$); and (c) a clinical control group of children not meeting the criteria for any ADHD subtypes ($n = 35$). The parent and teacher of each participant completed an ADHD Rating Scale-IV when the child was initially assessed. The difference in mean ratings for each group was statistically significant for the parent Inattention subscale score ($p < .001$), the parent Hyperactivity-Impulsivity subscale score ($p < .01$), the teacher Inattention subscale score ($p < .0001$), and the teacher Hyperactivity-Impulsivity subscale score ($p < .0001$). Parent and teacher inattention ratings were significantly higher for students in the ADHD/I group and the ADHD/COM group compared to the control group. Parent and teacher hyperactivity-impulsivity ratings were significantly higher for students in the ADHD/COM group compared to the control and ADHD/I groups.

Predictive validity in a clinical setting was assessed using the same sample described previously for the discriminant validity study. Logistical regression models were used to evaluate the ability of the parent and teacher ratings to differentiate clinical groups from the control group and clinical groups from each other. Also evaluated was whether the parent and teacher ratings combined achieved a higher level of accuracy in diagnosing ADHD than from a single-informant approach. Teacher Inattention subscale scores accurately predicted inclusion in the ADHD/I group compared to the control group 74% of the time. Parent Inattention subscale scores accurately predicted inclusion in the ADHD/I group compared to the control group 68% of the time. Teacher and parent Inattention rating scales correctly predicted placement in the ADHD/COM group compared to the control group 80% and 62% of the time, respectively. Combining the Inattention subscale scores (parent and teacher scores) did not improve prediction accuracy. Teacher and parent Hyperactivity-Impulsivity subscale rating scores correctly predicted placement in the ADHD/COM group compared to the control group 65% and 60% of the time, respectively. When the ratings were combined, the scores correctly classified 75% of the students for the ADHD/COM group versus the control group. Teacher and parent Hyperactivity-Impulsivity subscale rating scores correctly predicted placement in the ADHD/COM group versus the ADHD/I group 85% and 64% of the time, respectively. The combination of the parent and teacher rating scores on the Hyperactivity-Impulsivity subscale did not increase the predictive accuracy.

Predictive validity in a school setting was assessed with a subgroup of students referred from two school districts to the Pupil Assistance Committee for academic and/or behavior problems. In a multistep screening process, 128 students were classified into the following three groups: (a) ADHD/I, (b) ADHD/COM, and (c) control (no students qualified for ADHD/HI). The parent and teacher Inattention subscale ratings when entered separately were predictive of membership in the ADHD/I group versus the control group 75% and 76%

of the time, respectively. When the scores were combined, they accurately predicted inclusion in the ADHD/I group versus the control group 78% of the time. The parent and teacher Inattention subscale ratings when entered separately were predictive of membership in the ADHD/COM group versus the control group 78% and 80% of the time, respectively. The combination of the parent and teacher Inattention subscale scores accurately predicted inclusion in the ADHD/COM group versus the control group 83% of the time. The parent and teacher Hyperactivity-Impulsivity subscale differentiated inclusion in the ADHD/COM group versus the control group 78% and 83% of the time, respectively. Combining the teacher and parent subscale scores correctly predicted placement in the ADHD/COM group versus the control group 83% of the time. The parent and teacher Hyperactivity-Impulsivity subscale differentiated the inclusion in the ADHD/COM group versus the ADHD/I group 62% and 79% of the time, respectively. Combining the teacher and parent subscale scores did not improve the differentiation of placement in the ADHD/COM group versus the ADHD/I group.

The usefulness of the ADHD Rating Scale-IV in diagnosing ADHD and screening for ADHD was discussed in terms of its sensitivity, positive predictive power (PPP), negative predictive power (NPP), and specificity. Sensitivity was investigated by assessing the probability that children known to have ADHD are rated at or above a certain cutoff score. PPP refers to the probability of the presence of certain behavior or a disorder concurrent with a certain score on the scale. Specificity refers to the probability that children known not to have ADHD would score below a certain cutoff on the scale. NPP refers to the probability that a child does not have ADHD if he or she scores below a certain cutoff score on the scale. Studies were conducted with a clinical trial and a school-based trial to determine optimal cutoff scores for diagnosing ADHD and ruling out ADHD using sensitivity, specificity, PPP, and NPP. The results are presented in a variety of tables describing the subscale score cutoffs for either the parent or the teacher version or a combination of scores for ruling out ADHD or diagnosing ADHD. A large percentage of children were inaccurately predicted for inclusion in different subtypes, and therefore it is recommended that this scale be used in conjunction with other measures of ADHD for diagnoses purposes. The ADHD Rating Scale-IV is useful as a screening measure using both the parent and the teacher version. A comprehensive evaluation including both the parent and teacher version of the ADHD Rating Scale-IV and other measures of ADHD is recommended.

In addressing the ability of the ADHD Rating Scale-IV to evaluate treatment outcome, a discussion on the Reliable Change Index (RCI) by Jacobsen and Truax (1991) is provided. To calculate the RCI score for a posttreatment evaluation, a clinician subtracts the pretreatment rating scale score from the posttreatment score and then divides the difference by the standard error of difference for that scale. Two tables provide the standard error of difference needed for the RCI formula evaluation. Standard error differences are provided by gender, age, subscale or total score, and rating scale version. Because the test-retest reliability ratings are adequate for the ADHD Rating Scale-IV and the RCI formula compares the same child on two separate assessment occasions, the ADHD Rating Scale-IV is well suited for evaluating treatment outcome using the RCI formula and the tables provided in the manual.

Summary

The ADHD Rating Scale-IV attempts to measure behaviors associated with ADHD and its *DSM-IV* subtypes. The scale was developed to closely approximate the *DSM-IV* diagnostic criteria for ADHD primarily inattentive subtype, ADHD primarily hyperactive-impulsivity subtype, and ADHD combined subtype. This is an 18-item scale to be completed by the parent and/or guardian or teacher of a 4- to 20-year-old student. There are two separate versions of this scale: the home (also available in Spanish) and the school version. The administration of both versions is preferable, particularly for subtype specification.

The ADHD Rating Scale-IV has adequate reliability and validity, although some of the technical aspects suggest caution when using this scale. The standardization sample closely matches the 1990 census data on ethnic group and region. However, limited and generalized information is provided on the socioeconomic status of the participants. No information is provided on the urban/rural residence status and the parent education levels of the norm population. Limited information is provided on the education classroom for the normative participants (i.e., special education vs. general education). The only information provided is for the school version normative group (91% general education and 8% special education), and then there is no discussion of how this breakdown compares to the school district or the population in general. The limited cell size for ethnic groups restricts interpretation of the ADHD Rating Scale-IV with minority populations. Data were not provided separately for different ethnic groups. The authors caution a possible overidentification of African Americans. When sample participants were grouped by ADHD subtype, no individuals qualified for the ADHD primarily hyperactive-impulsive subtype. This limits the use of the ADHD Rating Scale-IV as a measure of this subtype. Considering the low specificity ratings for the ADHD primarily inattentive and ADHD combined subtypes, this scale should be used with caution and in conjunction with other measures for subtype specific diagnoses.

The ADHD Rating Scale-IV has other limitations as well. As mentioned, the ADHD Rating Scale-IV results are limited for use with ethnic minority groups, particularly African Americans. In general, more research is needed to determine the best use of the ADHD Rating Scale-IV as it differs by setting, version (home or school), referral source, and subtype functioning. Nonetheless, many clinicians will find the ADHD Rating Scale-IV a useful screening tool. It is user friendly and time efficient for the parent, teacher, and clinician. It can also be a valuable tool as a treatment outcome measure when employing the RCI formula, assuming a pretreatment rating is obtained. The ADHD Rating Scale-IV is a good tool for gathering essential information in an initial ADHD referral despite its limitations.

Danielle Pappas
University of Tennessee, Knoxville

References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text rev.). Washington, DC: Author.
- Barkley, R. A. (1990). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment*. New York: Guilford.

- DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHA Rating Scale-IV: Checklists, norms, and clinical interpretation*. New York: Guilford.
- Jacobsen, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.