Psychometric Properties of the Vanderbilt ADHD Diagnostic Parent Rating Scale in a Referred Population

Mark L. Wolraich, MD, Warren Lambert, PhD, Melissa A. Doffing, MA, Leonard Bickman, PhD, Tonya Simmons, BS, and Kim Worley, MD

1Child Study Center, University of Oklahoma Health Sciences Center, and 2Center for Mental Health Policy, Vanderbilt University

Objectives  To determine the psychometric properties of the Vanderbilt Attention Deficit/Hyperactivity Disorder Parent Rating Scale (VADPRS), which utilizes information based on the Diagnostic and Statistical Manual of Mental Disorders, 4th Ed. (DSM-IV). The VADPRS was created to collect uniform patient data and minimize the time burden of lengthy interviews. Methods  Participant data (N = 243) was used from the first 2 years of a longitudinal study on communication among physicians, teachers, and parents in diagnosing, treating, and managing children with attention deficit/hyperactivity disorder (ADHD). The reliability, factor structure, and concurrent validity of the VADPRS were evaluated and compared with ratings of children in clinical and nonclinical samples on the Vanderbilt ADHD Teacher Rating Scale and the Computerized Diagnostic Interview Schedule for Children–IV, Parent version. Results  The internal consistency and factor structure of the VADPRS are acceptable and consistent with DSM-IV and other accepted measures of ADHD. Conclusion  The VADPRS is a reliable, cost-effective assessment for ADHD in clinical and research settings.

Key words  ADHD; parent; behavior; rating scale.

Attention deficit/hyperactivity disorder (ADHD) is the most common neurobehavioral diagnosis affecting children today (Olfson, 1992; Shaywitz & Shaywitz, 1988). Given the widespread attention that ADHD is receiving in health care and the media (Angier, 1994; Diller, 1996; McGinnis, 1997), a uniform process to evaluate children who present with the core symptoms of inattention, hyperactivity, impulsivity, or poor academic achievement is warranted. Guidelines of both child psychiatry (Dulcan, 1997) and pediatrics (American Academy of Pediatrics, 2000, 2001) encourage clinicians to employ criteria of the Diagnostic and Statistical Manual of Mental Disorders, 4th Ed. (DSM-IV) (American Psychiatric Association [APA], 1994) in making the diagnosis.

Behavior rating scales have been one method for obtaining information from parents and teachers efficiently. Most earlier scales, such as the Conners Rating Scales (Goyette, Conners, & Ulrich, 1978) and the Child Behavior Checklist (Achenbach & Edelbrock, 1983), differ from DSM-IV in several ways: (a) They were more broad based, (b) they did not include all the specific DSM criteria required to make a diagnosis, and (c) they derived their categories based on deviations from the norm. Scales specific for ADHD utilizing the 18 core symptoms have been developed (Conners, Sitarenios, Parker, & Epstein, 1998; DuPaul et al., 1997; Molina, Smith, & Pelham, 2001; Swanson, Nolan, & Pelham, 1982; Wolraich, Feurer, Hannah, Pinnock, & Baumgaertel, 1998) for parents and/or teachers. In addition to the ADHD core symptoms, some of the scales include symptoms for at least the other disruptive behaviors. The Vanderbilt ADHD Teacher Rating Scale (VADTRS) (Wolraich et al., 1998) is a relatively simple instrument that follows the DSM-IV criteria for ADHD. In addition, the VADTRS has included a screen for some of the mood and anxiety symptoms and a rating of the child’s performance (Jellinek, Patel, Froehle, 2002). While the psychometric properties of the VADTRS have been reported (Wolraich et al., 1998), those of a par-
ent version have not. The purpose of the current study was to examine the psychometric properties of a Vanderbilt ADHD Parent Rating Scale (VADPRS).

Methods
Overview
Data were gathered as part of a longitudinal study of communication among physicians, teachers, and parents in diagnosing, treating, and managing children with ADHD. Children were recruited from an urban elementary school system; data were then collected from their parents, teachers, and physicians. Data used in the present study come from the first 2 years of a 4-year longitudinal study.

Longitudinal Data Collection Procedures
The data collection for the current study included an initial time-zero screen and three follow-up data collection waves. Information during the teacher screening process was collected anonymously, and parental consent was obtained from all the participants through procedures approved by the university institutional review board. The time-zero screen was conducted during the 1998–1999 school year. The communication study was presented to 61 out of 67 eligible elementary schools. The school system had an overall total K-12 enrollment of 69,400 students (46.6% black and 46.0% white). Average K-3 class size was 20 students. Of the 975 teachers, 317 (at 57 schools) chose to participate in the study by completing behavior rating scales (the VADTRS) on all the students in their classrooms. According to chi-squared tests, participating teachers were more likely than nonparticipating to have an education at the master's level and above, but did not differ by sex \( (p = .12) \), race \( (p = .07) \) and years experience \( (p = .18) \), or type of teacher \( (e.g., \text{Title 1, certified, grades K–4, } p = .16) \). According to \( t \) tests, average age \( (p = .07) \) and years experience \( (p = .12) \) were not significantly different for participating teachers.

There were 6,171 rating scales completed, of which 1,536 students were eligible for the study due to a teacher-reported clinical diagnosis of ADHD or by meeting rating scale criteria for ADHD. The teacher-reported ADHD screen showed a very high rate of 23% possible ADHD cases. This rate is much higher than the usual 3–5% estimates (APA, 1994) for two reasons. First, rates of ADHD were high in this district. The rate of “any ADHD” in a suburban county measured by the same method was 14%. Second, the teacher screen doesn't include all the criteria required to make a diagnosis. It does not include duration of symptoms and whether they started before the age of 7. In addition, the teacher screening does not include the requirement of dysfunction in more than one setting for the diagnosis of ADHD (criterion C in DSM-IV). The additional requirements are likely to produce lower rates of ADHD. When ADHD, as measured by the teacher screen, is used to predict the clinical diagnosis of ADHD (as reported by the teacher), the sensitivity was 63% and the specificity was 78%.

The first wave of the follow-up included multiple attempts made by the school staff to contact the parents of the 1,536 eligible students by letters and/or phone calls. (The project had to remain blind to confidential identifying information until after parents consented to participate.) The parents who participated \( (N = 288) \) completed a fully structured Computerized Diagnostic Interview Schedule for Children (C-DISC-IV; National Institute of Mental Health, 1997) interview in person by researchers. In April-November 1999, the VADTRS was sent to the teachers of the 288 participants; 89.9% of the surveys were returned \( (N = 259) \).

The second wave, approximately 6 months later, included a second interview (by phone) utilizing the parent rating scale (VADPRS) with the parents of 261 of the children (90.6%). Reasons for attrition of the parents varied: 4 were no longer interested, 15 had moved out of the area, and 8 could not be located.

The third wave, 6 months after the second interview, included a phone interview using the ADHD section of the C-DISC-IV and the VADPRS with 256 (95.2%) of the remaining 269 parents. Reasons for attrition varied: 2 moved in with a new caregiver, 4 had moved out of the area, and 7 could not be located. Additionally, the VADTRS was mailed to the teachers of the 269 children, with 89.6% \( (N = 241) \) of the surveys returned.

Measures
The VADPRS is the parents’ version of the teacher rating scale, the VADTRS (Wolraich et al., 1998). It includes all 18 of the DSM-IV criteria for ADHD. In addition, 8 criteria for oppositional defiant disorder (ODD) and 12 criteria for conduct disorder (CD) are included, along with 7 criteria from the Pediatric Behavior Scale (Lindgren & Koepp, 1987) that screen for anxiety and depression. The wording has been simplified so that the reading level is slightly below third grade. As with the teachers’ form, the parents are asked to rate the severity of each behavior on a 4-point scale (“never” to “very often”). The diagnosis is considered present if scores of 2 or 3 on a 0–3 scale (indicating that a behavior is “often” or “very often” present) are checked for the requisite number of criteria based on the DSM-IV definition of ADHD diagnosis.

The performance section of the VADPRS is an eight-
item scale with four items relating to academic performance: (a) overall academic performance, (b) reading, (c) mathematics, and (d) written expression. Another four items evaluate relationships: (e) peers, (f) siblings, (g) parents, and (h) participation in organized activities. The parent rates each of these on a 5-point scale from “problematic” to “above average.”

Results
The eligible sample \( (N = 1,536) \) shown in Table I includes children who met the behavior rating scale criteria for ADHD on the teacher screen or who had a teacher-reported clinician diagnosis of ADHD. All children were in elementary school (grades K–4), and a little more than half (52%) were African American. The modal age for participants was 7 (mean 7.41). As expected, the sample included mostly boys (68–69%). Two hundred forty-three participants were still in the project after completing the screening and the three data collection points with parents, teachers, and physicians. For simplicity, we call this a clinical sample of children with ADHD. Strictly speaking, though, children positive on the teacher screen did not necessarily meet the two requirements of DSM-IV: age of onset and impairment in two or more settings.

In most ways the study sample of 243 children with
ADHD, or ADHD characteristics, resembled the sample of eligible nonparticipants, e.g., by VADTRS scores, internalizing and externalizing comorbidity, race, grade in school. However, significant differences appear in recognition of ADHD, with higher percents of the participant sample having a clinical diagnosis, being treated with stimulant medication, or being referred to a special/support team (S team) for problems in school.

**Attrition**

For confidentiality, the research team was blinded to identifying data until after parents volunteered to participate. Invitations to participate had to be sent out by schools, leaving two reasons for nonparticipation: (a) parents chose not to participate, or (b) schools failed to get invitations to the parents. The higher rates of clinical diagnosis, medication, and S-team referrals in the study sample (N = 243) suggest that parents who chose to participate were somewhat more likely than nonparticipants (N = 1,293) to have children with recognized ADHD problems.

**Analysis**

Analysis of the VADPRS appears under five headings: (a) internal consistency reliability, (b) item analysis, (c) factor structure, (d) concurrent validity, and (e) comorbid scales for factor structure, reliability, and validity.

**Internal Consistency Reliability.** This analysis compares the VADPRS with the VADTRS and C-DISC-IV using available samples, including teacher and parent ratings of clinical samples and a large screening sample of school children. The main question concerns the reliability of the VADPRS, but the various samples also let us determine whether the DSM-IV criteria are reliable in low-end samples (school children without ADHD) and high-end samples (children identified as having ADHD).

Whole scale internal consistency reliabilities for the VADPRS, VADTRS, and C-DISC-IV appear in Table II. Results show good internal consistency in all methods and samples, with the overall Cronbach’s alpha ≥.90 in every case.

**Item Analysis.** The next analysis examines item correlations to determine whether some of the symptoms prove unreliable when used for parent reports in a clinical sample. Item reliabilities were evaluated with part-whole correlations, the correlation of a given item with the sum of all other items. Low part-whole correlations indicate items that are unreliable in a given sample, perhaps because of the restricted range.

Figure 1 compares item correlations for the different instruments (VADPRS, VADTRS, and C-DISC-IV). Ranking the part-whole correlations for each instrument sample made it possible to compare the overall quality of item correlations from best (highest) to worst (lowest). These item correlations show no items with inadequate part-whole correlations, which would be anticipated for the well-researched symptoms of ADHD. In this respect the parent rating scale is no worse, and possibly a bit better, than the VADTRS and C-DISC-IV.

The best and worst items are identified on the chart, showing for example that speech symptoms (excessive talking, blurring out of answers) have the lowest correlations with other symptoms of ADHD.

**Factor Structure.** Factor analysis of the VADPRS ADHD scales, as opposed to the comorbidity scales, is done to assess consistency with the DSM-IV measurement model for ADHD, as shown in Figure 2. This model views ADHD as comprising two separate but correlated components: inattention and hyperactivity/impulsivity.

The measurement model in Figure 2 is based on DSM-IV Symptoms 1–9 measure inattention, symptoms 10–18 measure hyperactivity/impulsivity, and the two groups of symptoms are correlated. In addition to showing the

<table>
<thead>
<tr>
<th>Alpha*</th>
<th>N</th>
<th>Rater</th>
<th>Sample</th>
<th>Instrument*</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.94</td>
<td>255</td>
<td>Parent</td>
<td>Clinical study sample</td>
<td>V Wave 2</td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td>253</td>
<td>Parent</td>
<td>Clinical study sample</td>
<td>V Wave 3</td>
<td></td>
</tr>
<tr>
<td>0.92</td>
<td>4,582</td>
<td>Teacher</td>
<td>School screen, ADHD–</td>
<td>V Time zero screen</td>
<td></td>
</tr>
<tr>
<td>0.91</td>
<td>1,268</td>
<td>Teacher</td>
<td>School screen, ADHD+ no follow-up</td>
<td>V Time zero screen</td>
<td></td>
</tr>
<tr>
<td>0.90</td>
<td>240</td>
<td>Teacher</td>
<td>School screen, ADHD+ with follow-up</td>
<td>V Time zero screen</td>
<td></td>
</tr>
<tr>
<td>0.93</td>
<td>251</td>
<td>Teacher</td>
<td>Clinical study sample</td>
<td>V Wave 1</td>
<td></td>
</tr>
<tr>
<td>0.94</td>
<td>234</td>
<td>Teacher</td>
<td>Clinical study sample</td>
<td>V Wave 3</td>
<td></td>
</tr>
<tr>
<td>0.93</td>
<td>286</td>
<td>Parent</td>
<td>Clinical study sample</td>
<td>D Wave 1</td>
<td></td>
</tr>
<tr>
<td>0.94</td>
<td>245</td>
<td>Parent</td>
<td>Clinical study sample</td>
<td>D Wave 3</td>
<td></td>
</tr>
</tbody>
</table>

*Raw Cronbach alphas. Standardized alphas were identical to two significant figures because all scales used the same 0–3 ratings.

*V = Vanderbilt ADHD Teacher or Parent Rating Scale; and D = Computerized-Diagnostic Interview Schedule for Children, 4th Ed. “in the last month.”
hypothesized model, Figure 2 shows the empirical loadings, which will be explained later.

The first step in the factor analysis was the examination of the eigenvalues of the correlations matrix in a scree plot (Figure 3), which provides an initial estimate of how many factors might reasonably be extracted.

For comparison, eigenvalues for the teacher screen were included, and the analysis was limited to the 222 children who had all three measures, the VADPRS at Wave 2, the VADPRS at Wave 3, and the VADTRS at Wave 3. In all 3 samples, the third eigenvalue was less than 1.0, which suggests that the correlation matrix can be explained by two factors and that adding a third would be of little use.

According to confirmatory factor analysis with EQS software (Multivariate Software, Inc., Encino, California; Bentler, 1985; Bentler & Bonett, 1980), the VADPRS is consistent with the two-factor model of ADHD. In both parent samples, the fit between parent-reported data and the two-factor model was satisfactory (> .90), as shown in Table III. The correlation between inattention and hyperactivity/impulsivity was high (r = .75 to .79).

The confirmatory model assumes that each symptom is due to its latent source (inattentive or hyperactive/impulsive ADHD) and error; the standardized coefficients in Figure 2 are quite uniform and consistent with a two-factor model, again indicating the maturity of the DSM-IV symptom list used in the VADPRS.

**Concurrent Validity.** The next analysis examines the correlation between VADPRS and the C-DISC-IV in order to evaluate the concurrent validity of the VADPRS, i.e., its correlation with an instrument of established reliability and validity. The VADPRS and ADHD section of the C-DISC-IV ratings were completed on the same occasion. Concurrent validity of the item total of the VADPRS is high (r = .79), suggesting that it measures much the same thing that the C-DISC-IV does. Internal consistency reliabilities of the C-DISC-IV and VADPRS were .93 or higher; a reliability of .93 can be interpreted as an expectation of r = .93 for two parallel forms of a given test. Since the observed correlation is lower than .93, we conclude that the VADPRS and the C-DISC-IV are very similar, but are not parallel forms of the same test.

**Comorbid Scales: Factor Structure, Reliability, and Validity.** In addition to the two ADHD scales, the VADPRS has two comorbidity scales to assess internalizing problems (anxiety and depression) and externalizing problems (ODD and CD) that often complicate ADHD. In this sample, 8% of the children were positive for internalizing comorbidity, and 23% were positive for externalizing comorbidity. Histograms for criteria counts appear in Figure 4.
Comorbidity counts are highly skewed, having modes of zero and a few very high scores. Since the teacher form has a four-factor structure (Wolraich, Hannah, Baumgaertel, & Feurer, 1998), we tested this structure in confirmatory factor analyses using the same methods as noted earlier in this paper, except for the addition of 20 items for ODD or CD and 7 for anxiety or depression. Factors 1 and 2 were again the nine symptoms of inattention and nine symptoms of hyperactivity/impulsivity. When the 28 comorbid items were forced into a single Factor 3, the model’s fit was inadequate (comparative fit index [CFI] = .88). Fit was satisfactory for a model with four factors (inattention, hyperactivity/impulsivity, ODD-CD, and anxiety-depression). In this case CFI = .93. This result suggests that the VADPRS, like the related teacher scale, can be considered a four-factor test instrument.

Reliability analysis suggested that the comorbid scales’ internal consistency was adequate. Cronbach’s alpha for ODD-CD (20 items) was .91; and for anxiety-depression (7 items), .79. The comorbidity scales showed some validity in correlations with hypothesized events and measures, but not enough to claim that convergent and discriminant validity (Campbell & Fiske, 1959; Fiske & Campbell, 1992) were well established. For example, the internalizing problem count had a significant correlation with scores on the Columbia Impairment Scale (r = .43, p < .0001) and children above the internalizing cut score were significantly more likely to have an S-team (p = .01) or school suspension (p = .002). However, they were no more likely to have contact with a mental health provider (p = .15) or detention in school (p = .18) or to miss more days of school (r = .06, p = .40).

The externalizing problem count had significant correlations with both Columbia Impairment Scale scores (r = .52, p < .0001) and days of school missed during the year (r = .24, p < .0001), and children above the cut score for externalizing comorbidity were significantly more likely to have been suspended (p = .003) and have contact with a mental health provider (p = .003). However, they were no more likely for detention in school (p = .07) or having an S-team (p = .13).

Discussion

While the VADPRS uses the well-tested ADHD symptom list from the DSM-IV, it still had to be evaluated empirically before field use. This was done in the present study. The reliability, factor structure, and concurrent validity of the VADPRS were evaluated and compared with VADTRS and C-DISC-IV ratings of children in clinical and nonclinical samples. Results suggest that the internal consistency and factor structure of the VADPRS are acceptable and consistent with DSM-IV and other accepted measures of ADHD. Table II (internal consistency in nine samples) suggests that the criteria are robust, having good reliability under a variety of conditions, including different respondents (parent and teacher), different methods (Vanderbilt checklists vs. C-DISC-IV interview), and different severities (ranging from the 4,582 ADHD-negative children to the various clinical samples). This adequacy is not unexpected. Since the VADPRS, VADTRS, and C-DISC-IV all use the same 18-symptom list from the DSM-IV, large differences in reliability or validity would not be anticipated. Confirmatory factor analysis and reliability assessment also confirm the structure of the two comorbid dimensions of internalizing and externalizing conditions.

The present study suggests that the VADPRS has good
psychometric properties in a high-risk sample; however, a study of a referred sample cannot establish the sensitivity and specificity of diagnostic method. It is possible that the VADPRS would perform differently on a low-risk sample with different characteristics. However, when the same items are used with teachers, confirmatory factor analysis showed that urban and suburban samples shared the same factor structure, and that the factor structure showed only small differences by gender, age, school grade, or severity of ADHD symptom counts (Wolraich et al., 2002). Like the VADTRS, the VADPRS makes no attempt to enforce the DSM-IV's ≥2-setting requirement, so prevalence rates may be higher than the true rates. Wolraich et al. (in press) found that parent-teacher agreement was low, with symptom count correlations ranging from \( r = .27 \text{ to } .34 \) and ADHD diagnostic agreement ranging from Kappa = .11 to .15. Strict enforcement of the ≥2-setting requirement led to rates of diagnosis much lower than parent or teacher alone. The VADPRS is not a substitute for clinical diagnosis. Together, the VADPRS and VADTRS can be useful adjuncts to an evaluation.

The VADPRS is easy to complete (having a reading level slightly below third grade), easy to score following the criteria required to make a DSM-IV diagnosis, and includes both the core symptoms and a rating of performance as well as providing a screen for common comorbid conditions. Its ease of administration and scoring as well as its provision of specific DSM-IV information make it a useful clinical tool in evaluating patients and a useful research tool for evaluating ADHD in a cost-efficient manner.

Rating scales such as the VADPRS are systematic and require less professional time than do interviews to determine the presence of specific behaviors. The high correlations between rating scales and direct interview in an urban inner-city sample suggest that rating scales may be used as accurate measures of a child’s behavior for the purposes of diagnosing ADHD, even with a sample of parents with low reading levels. A rating scale can reduce the time a clinician requires to gather information and thereby allow the clinician to focus the clinical interview on other issues. However, it is important to note that the scales are

Table III. Results of Confirmatory Analysis of Measurement Model

<table>
<thead>
<tr>
<th>Method or Result</th>
<th>Wave 2 VADPRS</th>
<th>Wave 3 VADPRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>N parental ratings</td>
<td>242</td>
<td>242</td>
</tr>
<tr>
<td>Item scores missing</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Bentler Comparative Fit Index (CFI)</td>
<td>0.923</td>
<td>0.914</td>
</tr>
<tr>
<td>Bentler robust CFI</td>
<td>0.941</td>
<td>0.940</td>
</tr>
<tr>
<td>Correlation ( r ) (inattention hyperactivity)</td>
<td>.792</td>
<td>.753</td>
</tr>
</tbody>
</table>

VADPRS = Vanderbilt Attention Deficit/Hyperactivity Disorder Parents Rating Scale.

CFI’s from EQS software (Bentler & Wu, 1993; Hu & Bentler, 1995) indicate how well model fits observed data. CFI > .90 considered adequate (Byrne, 1994, p. 35). Robust estimate corrects for possible non-normal distribution of item scores. Sample includes same children, each rated twice with no missing items.
not substitutes for clinical interviews to obtain information from parents and children in evaluating a child for possible ADHD.

Rating scales such as the VADPRS are important for research. They can obtain information on large groups of children at low costs. Its reasonable correlation with the C-DISC-IV makes it possible to attain DSM-IV diagnostic information for ADHD without having to employ more costly structured interviews. Such a process could reduce the reliance on convenience-based samples, which is currently the norm in many studies, and allow for samples that better reflect the overall population. This change will provide more accurate prevalence estimates.

The present study had limitations. The rate of participation among eligible families was low. However, there were few differences between participants and nonparticipants. Another limitation was the absence of test-retest reliability. While we were able to use a high-risk sample to compare the VADPRS to a gold standard for ADHD, the C-DISC-IV, and we could demonstrate a relationship between the comorbid conditions and measures of impairment, we could not specifically assess the concurrent validity of the comorbid conditions and performance. A study to obtain this assessment is currently in progress.

Acknowledgement

This study was supported by a grant from the National Institute of Mental Health (HS/MH 0905).

Received August 19, 2002; first revisions received November 11, 2002; second revisions received January 27, 2003; accepted March 21, 2003

References


---

**Psychometric Properties of the VADPRS** 567