The relationship of malingering test failure to self-reported symptoms and neuropsychological findings in adults referred for ADHD evaluation

Julie Suhr*, Dustin Hammers, Katy Dobbins-Buckland, Eric Zimak, Carrie Hughes

Ohio University, United States
Accepted 7 May 2008

Abstract

Diagnosis of adult attention-deficit/hyperactivity disorder (ADHD) adults is difficult, as neither symptom report nor neuropsychological findings are specific to ADHD. Few studies address the possibility that noncredible performance influences both symptom report and neuropsychological findings. The present study utilized archival data from young adults referred for concerns about ADHD, divided into three groups: (1) those who failed a measure of noncredible performance (the Word Memory Test; WMT), (2) those who met diagnostic criteria for ADHD, and (3) controls with psychological symptoms but no ADHD. Results showed a 31% failure rate on the WMT. Those who failed the WMT showed clinical levels of self-reported ADHD symptoms and impaired neuropsychological performance. Neither self-report measures nor neuropsychological tests could distinguish ADHD from psychological controls, with the exception of self-reported current hyperactive/impulsive symptoms and Stroop interference. Results underscore the effect of noncredible performance on both self-report and cognitive measures in ADHD.

Keywords: Malingering; Attention-deficit disorder with hyperactivity; Neuropsychological assessment; Self-report

In increasing numbers, adults are presenting for evaluation with concerns that they have attention-deficit/hyperactivity disorder (ADHD). The diagnosis of ADHD in adults is complicated by several factors, most notably the difficulty establishing evidence for either childhood or current presence of impairing symptoms of the disorder, outside of self-report (Murphy & Schachar, 2000). Further complicating this issue is that neuropsychological assessment is still not broadly accepted as part of a diagnostic evaluation for either children or adults presenting with concerns about ADHD (Stefanatos & Baron, 2007); self-reported symptoms and functioning are still relied upon as major components of diagnosis.

One problem with reliance on self-report in diagnosis of ADHD is the high base rate with which “ADHD” symptoms are reported in the general population. Although only 1–5% of adults are diagnosed with ADHD (Du Paul et al., 2001; Harrison, 2004), a large number of adults report frequently experiencing ADHD-like symptoms, either currently or in childhood. For example, Murphy and Barkley (1996) found that 22% of current ADHD symptoms and 56% of childhood ADHD symptoms were endorsed as occurring “often or very often” by 20% or more of a sample of adults applying...
ADHD symptoms were recorded as clinically significant in at least 20% of the then used their ratings to form both probable and definite diagnosis of childhood ADHD. Notably, 32% of childhood symptoms of inattention, hyperactivity, and impulsivity based on the adult interview data. Interviewers unaware of the diagnostic status of participants and the study’s purpose, rated the presence of clinically significant (based on strict research criteria) and adults who showed no evidence of ADHD in childhood. Interviewers, who were unaware of the diagnostic status of participants and the study’s purpose, rated the presence of clinically significant childhood symptoms of inattention, hyperactivity, and impulsivity based on the adult interview data. Interviewers then used their ratings to form both probable and definite diagnosis of childhood ADHD. Notably, 32% of childhood ADHD symptoms were recorded as clinically significant in at least 20% of the control participants, and 11% of controls were diagnosed with ADHD in adulthood based on their retrospective symptom report, despite having been carefully screened for the absence of ADHD when they were children. Such results support the argument that reliance on self-reported symptom data is problematic in the diagnosis of ADHD in adult samples (Murphy, Gordon, & Barkley, 2000).

Not surprisingly, the base rate of reported ADHD symptoms is also high in samples of individuals seeking treatment, even for non-ADHD related concerns. For example, Harrison (2004) asked 224 university undergraduates presenting to university clinics for treatment (180 for medical concerns, 12 for psychiatric concerns, and 21 for counseling) to complete the Brown Attention-Deficit Disorder Scale (BADD; Brown, 1996), a measure of current adult ADHD symptoms commonly used in clinical practice. Although none had a history of ADHD diagnosis, 35% met criteria for probable ADHD using the cutoff score given in the test’s manual. Specifically, when broken down by referral source, 27% of the individuals self-referred for medical concerns and over 65% of the psychiatric/counseling referrals scored above the cutoff on the scale. With regard to retrospective childhood symptom report, studies also suggest a high number of false positive scores using the commonly used cutoff on a popular retrospective childhood ADHD scale (the Wender Utah Rating Scale (WURS); Ward, Wender, & Reimherr, 1993). For example, almost half of a sample of individuals who did not meet diagnostic criteria for ADHD but who were seeking treatment for psychiatric concerns fell above this cutoff (McCann & Roy-Byrne, 2004). A similar number of false positives (40–60%) were found in a sample of adults specifically presenting with concerns that they may have ADHD but who did not meet diagnostic criteria for the disorder (Roy-Byrne et al., 1997). The nonspecificity of “ADHD” symptoms to ADHD raises significant concerns about reliance on self-reported current or childhood symptoms for differential diagnosis.

Another potential problem with reliance on self-report measures in ADHD diagnosis is their vulnerability to non-credible symptom reporting, which is not addressed by existing studies. There are ample secondary gain issues in ADHD, including but not limited to receipt of psychostimulant medications and academic/work accommodations (Alfano & Boone, 2007; Sullivan, May, & Galbally, 2007), that may motivate individuals to misrepresent themselves with regard to current or past symptoms (i.e., malingering). The face validity of both current and retrospective symptom report measures makes these instruments easy to exaggerate (Quinn, 2003). Furthermore, the symptoms of ADHD are well known in the general population, due to television advertisements for ADHD medications that list the symptoms, popular press articles and books describing the syndrome, and easily accessed Internet information on ADHD symptoms and the measures used to assess them (Conti, 2004; Murphy, 1994). Of course, there are many potential reasons for high symptom report other than malingering that may be equally noncredible (i.e., not diagnostic of ADHD); for example, high report of any psychological symptoms as a “cry for help” in someone experiencing psychological distress. Although many psychological instruments have subscales to assess the validity of self-report, measures of self-reported ADHD symptoms typically do not (Quinn, 2003).

Given that ADHD is a neurodevelopmental disorder with neuropsychological impairments as core symptoms, another component of many ADHD assessments is a neuropsychological evaluation. However, as mentioned above, neuropsychological test results are not yet a universally accepted part of an ADHD evaluation in either child or adult cases. One reason for this may be that, although reviews show robust differences between ADHD and control groups in many neuropsychological domains, neuropsychological tests are not always sensitive to ADHD and are not specific to the disorder (Nigg, 2005). It is interesting that, despite clear problems with sensitivity and specificity of self-reported symptoms in ADHD diagnosis, clinicians and researchers continue to rely on self-report for diagnosis, but then consider the poor diagnostic “accuracy” in neuropsychological tests to be a fatal flaw. In fact, a factor that may contribute to the poor diagnostic sensitivity of neuropsychological tests in ADHD is the inclusion of false positives.
in an ADHD diagnosis group, particularly in studies that rely on self-reported childhood and current symptoms for diagnosis. In addition, the role of noncredible performance is not considered in neuropsychological studies of ADHD, and noncredible performance may influence both participant diagnosis (i.e., noncredible self-report leading to invalid diagnosis) and neuropsychological test performance. As with self-reported symptoms, neuropsychological performance can be vulnerable to conscious attempts to deceive an evaluator for the purpose of external gain (i.e., malingering). Both Quinn (2003) and Harrison, Edwards, and Parker (2007) found that individuals asked to simulate ADHD can readily identify behaviors to emulate that would assist them in diminishing their performance on neuropsychological tasks. Further, malingering is a major contributor to neuropsychological test scores in other populations (Green, Rohling, Lees-Haley, & Allen, 2001). Because it is not possible to determine whether someone has conscious intent to deliberately deceive an assessor for the purposes of obtaining an external gain, we will continue to use the term “noncredible performance” (Boone, 2008) when referring to invalid behavioral presentation of a disorder (such as failing a test designed to measure poor effort), and reserve the word “malingering” for simulation studies in which individuals were asked to deliberately simulate ADHD, where conscious intent to deceive is known to be the motivation behind the invalid behavioral presentation.

As noted above, few studies have examined the effect of noncredible performance on ADHD assessment, either on self-report instruments or neuropsychological tests, and most existing studies of noncredible performance in ADHD have utilized a simulated malingering design with undergraduate samples. Undergraduates asked to malinger ADHD either report similar or higher levels of childhood and current ADHD symptoms than adults diagnosed with ADHD (Booksh, 2005; Harrison et al., 2007; Quinn, 2003), and malingerers generate scores on clinical ADHD instruments that lead to high levels of false positive diagnosis (for example, 65–95% in Jachimowicz & Geiselman, 2004). With regard to neuropsychological test performance, malingerers perform worse than individuals diagnosed with ADHD on continuous performance tasks (Booksh, 2005; Quinn, 2003) and achievement test-based measures of processing speed (Harrison et al., 2007), but not on the Trail Making Test (TMT) or on the Weschler Adult Intelligence Scale-III measures of working memory and psychomotor processing speed (Booksh, 2005). Of note, Booksh’s (2005) study included the Word Memory Test (WMT; Green, 2005) as a measure of noncredible performance, which correctly identified 58% of individuals asked to malinger ADHD, with no false positives in a control group or in the group of individuals diagnosed with ADHD.

To date, the only published work that included assessment of noncredible performance in a clinical sample referred for ADHD evaluation is Sullivan et al. (2007). They analyzed the archives of 66 consecutive adult assessment cases referred for either ADHD or learning disability (LD) from a campus-based psychology clinic. In all cases the client had completed the WMT to assess for noncredible performance. There were 21 individuals referred for concerns about ADHD, 13 for concerns about an LD, and 32 for concerns about both. In the total sample, 22.4% of individuals failed the WMT using standard cutoffs listed in its manual. Of those referred specifically for concerns about ADHD, 47.6% failed, while 15.4% of referrals for only LD and 9.4% of those referring themselves for both disorders failed. Failure on the WMT was strongly related to poorer performance on the WAIS-III and the California Verbal Learning Test, as well as higher self-reported symptoms of ADHD on the CAARS. As this is the first published study to examine the use of measures of noncredible performance in ADHD, it is important to replicate and extend their findings in an independent sample. The present study sought to replicate and extend their findings by including multiple measures of noncredible performance, including a better-defined control group of individuals with other psychological conditions and by examining both current and childhood ADHD symptom report in addition to neuropsychological test performance.

The purpose of the present study was to examine rate of failure on a measure of noncredible performance (specifically the WMT) in a sample of young adults self-referred for evaluation for ADHD concerns, and to compare those who failed the WMT to those who passed on other measures of noncredible performance, self-reported current and childhood ADHD symptoms, and performance on neuropsychological tests commonly used in ADHD assessment. We hypothesized that those who failed the WMT would be more likely to fail other measures of noncredible performance, would self-report more current and more childhood symptoms of ADHD and would do worse on neuropsychological tests of attention, memory, executive functioning and processing speed, compared to individuals with ADHD and to controls. As an important differential diagnosis in ADHD evaluation is the presence of psychological conditions that can cause attention and concentration problems, and given the evidence above for high ADHD symptom report in individuals with other psychological disorders, we also examined whether there were differences in symptom report and neuropsychological test performance in adults with ADHD who passed the WMT relative to a control group of individuals with psychological diagnoses who also passed the WMT.
1. Method

1.1. Participants

Participants were drawn from a sample of 85 individuals who had undergone neuropsychological evaluation in a university psychology clinic from 2004 to 2007. The 85 individuals were selected among consecutive referrals for concerns about ADHD who met the following criteria: (1) must have signed for permission to use their de-identified interview and test results in archival research at the time of evaluation, (2) reported no history of significant neurological injury or illness, as judged by unstructured interview information, and (3) had completed the WMT during their evaluation. There were 48 males in the sample and 5 individuals who reported race/ethnicity status other than Caucasian (4 African American, 1 biracial). The average age was 22.7 (range 18–56 years) and the average years of completed education was 13.9, with a range of 12–20 years. Participants were divided into three groups based on records review.

Individuals who failed any one of the first four subtests of the WMT using norms available for the test (performance 82.5% or below on first three subtests, 70% or lower on the 4th) were assigned to the noncredible group \((n = 26)\). Of note, this represented a 31% failure rate on the WMT in the total sample. Of the 26, 9 failed all 4 subtests, 4 failed 3 subtests, 8 failed 2 subtests, and the rest failed 1 subtest.

To be included in the ADHD group \((n = 15)\), individuals (1) showed evidence of previous impairment in childhood related to ADHD symptoms, based on at least two pieces of evidence (self-report, parent report, school records, prior medical/psychological records), (2) evidenced clinically high ADHD symptoms at the present time through both self-report (including but not limited to the CAARS) and either collateral report or behavioral observation during the evaluation, and (3) passed the WMT. Based on their report of childhood and adult symptoms, 47\% \((n = 7)\) met criteria for predominantly inattentive subtype, while the rest met criteria for combined subtype. Of note, 27\% \((n = 4)\) were also diagnosed with and/or receiving treatment for depression at the time of the evaluation. Also of note, 33\% \((n = 5)\) were prescribed medication for ADHD at the time of the evaluation. Individuals who were taking stimulants were, with physician consent, not taking the stimulants while they completed their neuropsychological testing. In addition, individuals who were taking stimulants were instructed to complete self-report ADHD items based on their typical behavior, rather than their behavior while on medication.

To be included in the psychological symptom group \((n = 24)\), individuals (1) had no evidence of impairment in childhood related to ADHD complaints—note virtually all individuals in this group dated the onset of any current ADHD symptoms to high school or college, (2) evidenced current diagnosis of/treatment for a non-ADHD psychological condition independent of our clinical evaluation OR in some cases, had not yet been diagnosed but met diagnostic criteria for a psychological disorder, usually major depressive disorder, at the time of their evaluation, and (3) passed the WMT.

There were 20 additional participants who, for sample size reasons, were not included in the present analyses. These included 10 individuals who did not meet diagnostic criteria for ADHD but were diagnosed with learning disability during their evaluation and 10 individuals who were not given any diagnosis after their evaluation but did not meet the criteria described above.

2. Measures

Individuals presenting for neuropsychological ADHD evaluation in the clinic from which participants were drawn typically received a comprehensive neuropsychological evaluation, although not all individuals received the identical set of measures. Most individuals were tested over at least two different days of testing, usually more. A typical evaluation in a referral for adult ADHD included the Wechsler Adult Intelligence Scale-III (WAIS-III; Wechsler, 1997), either the Wechsler Individualized Achievement Test-II or the WIAT-II Abbreviated (The Psychological Corporation, 2002), the Auditory Verbal Learning Test (AVLT; Lezak, Howieson, & Loring, 2004), the Complex Figure Test (Lezak et al., 2004), either the long form or the short form of the Wisconsin Card Sorting Test (Heaton, Chelune, Talley, Kay, & Curtiss, 1993; Kongs, Thompson, Iverson, & Heaton, 2000), the Stroop Color and Word Test (SCWT; Golden, 1978), verbal fluency (Lezak et al., 2004), the Trail Making Test (TMT; Lezak et al., 2004), and the WMT (Green, 2005), not in any particular order. Typical self-report measures included the WURS, the Conners Adult ADHD Rating Scale Self-Report Long Form (CAARS; Conners, Erhardt, & Sparrow, 1998), and the Beck Depression Inventory-II (Beck,
other self-report and psychological measures were also administered as clinically indicated but less frequently. Less than half the sample received the Conners’ Continuous Performance Test-II (CPT-II; Conners & MHS Staff, 2000), in part due to computer problems. In addition, the clinic from which the study data were drawn discontinued routine use of continuous performance tests for both child and adult ADHD evaluations, due to growing data demonstrating poor specificity of continuous performance tasks in ADHD diagnosis (see Cohen & Shapiro, 2007, for a review). We have previously reported on the data available from the subsample of individuals in the present study who completed the CPT-II, which showed no difference in the number of individuals showing CPT impairment among individuals failing the WMT, individuals diagnosed with ADHD, and controls (some with psychological diagnoses) (Suhr, Dobbins-Buckland, & Hammers, 2007).

Although the neuropsychological battery included many measures, for the present study we focused attention on tests and cognitive domains shown to have at least medium effect sizes in studies of adult ADHD (Boonstra, Oosterlaan, Sergeant, & Buitelaar, 2005; Hervey, Epstein, & Curry, 2004; Schoechlin & Engel, 2005), including verbal learning (AVLT learning over trials, immediate and delayed recall, and delayed recognition raw scores), processing speed (TMT part A time to completion, verbal fluency raw score, SCWT Color t score, and Processing Speed Index from the WAIS-III), and working memory/executive functioning (Working Memory Index from the WAIS-III, TMT part B time to completion, and SCWT Interference t score). In addition, we compared performance in the groups on other indices that have been validated as measures of noncredible performance, including the Exaggeration Index for an expanded AVLT (EIAVLTX; Barrash, Suhr, & Manzel, 2004), WAIS-III digit span less than 5 (Iverson & Franzen, 1994), Working Memory Index less than 70 (Etherton, Bianchini, Ciota, Heinly, & Greve, 2006), AVLT recognition less than 10 (Boone, Lu, & Wren, 2005; Meyers, Morrison, & Miller, 2001), and WAIS-III Vocabulary—Digit Span score of 2 or greater (Greve, Bianchini, Mathias, Houston, & Crouch, 2003).

3. Results

The three groups were not different in age, $F(2, 62) = 1.52, p = .23$, educational level, $F(2, 62) = 0.79, p = .46$, verbal intellect (VCI), $F(2, 62) = 2.54, p = .09$, nonverbal intellect (POI), $F(2, 62) = 2.83, p = .07$, or gender distribution, $\chi^2(2) = 0.37, p = .83$, see Table 1.

With regard to other measures of noncredible performance, in all but one case the indices were 100% specific to the noncredible group, though sensitivity was low. The EIAVLTX identified 2 of the noncredible group members, the Digit Span identified 5, the Working Memory Index identified 1, and AVLT recognition identified 3. Although the Vocabulary—Digit Span cutoff identified 8 members of the noncredible group, 6 members of the psychological symptom group were also identified by the cutoff. The CAARS inconsistency score was also not entirely specific to the noncredible group. Scores above the clinical cutoff on the inconsistency scale were obtained by 3 individuals in the noncredible group and 1 each in the ADHD and psychological symptom groups. This score, however, reflects inconsistency in responding to items that measure similar content, and thus is not a measure of an over-reporting response bias. The CAARS manual suggests that t scores above 80 on any of the subtests should be considered possible evidence for symptom exaggeration, although the manual also cautions that scores this high could represent severity of symptoms in individuals with ADHD. Of note, on CAARS e scale (DSM-IV Inattention symptoms), 65% of the noncredible group, 60% of the ADHD group, and 54% of the psychological symptom group scored at 80 or above. On CAARS f scale (DSM-IV hyperactive/impulsive symptoms), 33% of the noncredible group, 7% of the ADHD group, and 4% of the psychological symptom group scored at 80 or above.

Table 1

<table>
<thead>
<tr>
<th>Demographic information for the three groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invalid group</td>
</tr>
<tr>
<td>Age (mean, S.D.)</td>
</tr>
<tr>
<td>Education (mean, S.D.)</td>
</tr>
<tr>
<td>WAIS-III VCI (mean, S.D.)</td>
</tr>
<tr>
<td>WAIS-III POI (mean, S.D.)</td>
</tr>
<tr>
<td># Male</td>
</tr>
</tbody>
</table>

Note: ADHD, attention-deficit/hyperactivity disorder. WAIS-III, Wechsler Adult Intelligence Scale-III. VCI, Verbal Comprehension Index. POI, Perceptual Organizational Index.
Table 2
Performance on self-report attention-deficit/hyperactivity disorder (ADHD) measures in the three groups

<table>
<thead>
<tr>
<th></th>
<th>Invalid group</th>
<th>ADHD group</th>
<th>Psychological controls</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td>Mean (S.D.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAARS a</td>
<td>65.8 (15.4)</td>
<td>71.7 (10.5)</td>
<td>66.3 (9.8)</td>
<td>0.98</td>
<td>.38</td>
</tr>
<tr>
<td>CAARS b</td>
<td>61.3 (12.7)</td>
<td>63.6 (7.8)</td>
<td>56.8 (9.6)</td>
<td>1.90</td>
<td>.15</td>
</tr>
<tr>
<td>CAARS c</td>
<td>59.2 (14.4)</td>
<td>61.5 (9.6)</td>
<td>55.0 (11.8)</td>
<td>1.26</td>
<td>.29</td>
</tr>
<tr>
<td>CAARS d</td>
<td>59.1 (13.6)</td>
<td>55.7 (7.7)</td>
<td>59.7 (10.7)</td>
<td>0.5</td>
<td>.56</td>
</tr>
<tr>
<td>CAARS e</td>
<td>78.1 (17.3)</td>
<td>82.8 (6.9)</td>
<td>76.5 (15.5)</td>
<td>0.69</td>
<td>.50</td>
</tr>
<tr>
<td>CAARS f</td>
<td>68.0 (17.6)</td>
<td>68.8 (11.6)</td>
<td>58.6 (11.1)</td>
<td>3.37</td>
<td>.04</td>
</tr>
<tr>
<td>CAARS g</td>
<td>78.2 (12.9)</td>
<td>81.0 (7.3)</td>
<td>73.5 (11.4)</td>
<td>2.10</td>
<td>.13</td>
</tr>
<tr>
<td>CAARS h</td>
<td>64.8 (12.5)</td>
<td>66.3 (7.4)</td>
<td>61.8 (11.1)</td>
<td>0.79</td>
<td>.45</td>
</tr>
<tr>
<td>WURS</td>
<td>48.5 (12.5)</td>
<td>48.5 (16.2)</td>
<td>34.5 (18.9)</td>
<td>5.31</td>
<td>.008</td>
</tr>
</tbody>
</table>

Note: CAARS = Conners' Adult Attention Rating Scale. WURS = Wender Utah Rating Scale.
* Invalid group and ADHD group higher than psychological controls, both p < .01. 1 Although this score was significant in univariate analyses, the MANOVA including all CAARS subscales was not significant.

3.1. Self-report measures

Differences among groups on CAARS subtests were examined using MANOVA. CAARS subtests could not distinguish the groups, Wilk’s lambda $F(16, 94) = .81$, $p = .65$, see Table 2. Differences among groups on age- and gender-based clinical cutoffs (t score 70 or higher) of the CAARS were examined using chi square, with a Bonferroni-corrected significance level of .006. No CAARS scales were specific to ADHD; all subscales were endorsed at clinical levels by at least 20% of the noncredible group, and scales a (inattention/memory problems), e (DSM-IV inattentive symptoms), g (DSM-IV total ADHD symptoms), and h (ADHD index) were endorsed by at least 20% of the psychological symptom group. The only scale that could distinguish the three groups was subscale f (DSM-IV hyperactive-impulsive symptoms), $\chi^2(2) = 12.05$, $p = .002$. This was due to low endorsement at a clinical level in the psychological symptom group (8%, lower than the noncredible group at $p = .003$ and lower than the ADHD group at $p = .001$); the invalid and ADHD groups were not distinguishable from one another (46% of the invalid group and 60% of the ADHD group above clinical cutoffs, $p = .55$) see Table 3.

The groups scored differently on the WURS, $F = 5.31$, $p < .008$, with the psychological symptom group scoring significantly lower than the other two groups ($p < .01$), who were not distinguishable from one another (see Table 2).

3.2. Performance on neuropsychological measures

Using MANOVA, the groups were not different in processing speed (WAIS-III PSI, TMT part A, COWA, SCWT color t score), Wilk’s lambda $F(8, 68) = 0.94$, $p = .49$. The groups were significantly different in memory performance

Table 3
Percentage of participants in the three groups who fell above t score = 70 on CAARS subscales

<table>
<thead>
<tr>
<th></th>
<th>Invalid group</th>
<th>ADHD group</th>
<th>Psychological controls</th>
<th>$\chi^2(2)$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>42</td>
<td>67</td>
<td>33</td>
<td>3.62</td>
<td>.16</td>
</tr>
<tr>
<td>B</td>
<td>27</td>
<td>13</td>
<td>8</td>
<td>3.18</td>
<td>.20</td>
</tr>
<tr>
<td>C</td>
<td>27</td>
<td>13</td>
<td>13</td>
<td>1.93</td>
<td>.38</td>
</tr>
<tr>
<td>D</td>
<td>23</td>
<td>0</td>
<td>17</td>
<td>3.18</td>
<td>.19</td>
</tr>
<tr>
<td>E</td>
<td>81</td>
<td>100</td>
<td>79</td>
<td>2.88</td>
<td>.23</td>
</tr>
<tr>
<td>F</td>
<td>46</td>
<td>60</td>
<td>8</td>
<td>12.05</td>
<td>.002</td>
</tr>
<tr>
<td>G</td>
<td>92</td>
<td>80</td>
<td>67</td>
<td>3.99</td>
<td>.13</td>
</tr>
<tr>
<td>H</td>
<td>27</td>
<td>33</td>
<td>21</td>
<td>.73</td>
<td>.69</td>
</tr>
</tbody>
</table>

Note: CAARS = Conners’ Adult Attention Rating Scale. ADHD = attention-deficit/hyperactivity disorder.
Table 4
Performance on neuropsychological measures in the three groups

<table>
<thead>
<tr>
<th></th>
<th>Invalid group</th>
<th>ADHD group</th>
<th>Psychological controls</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Memory</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVLT LOT</td>
<td>42.2 (9.3)</td>
<td>51.0 (8.2)</td>
<td>50.2 (14.0)</td>
<td>4.32</td>
<td>.021</td>
</tr>
<tr>
<td>AVLT IR</td>
<td>7.4 (2.8)</td>
<td>10.4 (2.4)</td>
<td>11.2 (2.8)</td>
<td>13.23</td>
<td>&lt;.0011</td>
</tr>
<tr>
<td>AVLT DR</td>
<td>7.5 (2.8)</td>
<td>9.4 (2.5)</td>
<td>10.7 (3.0)</td>
<td>7.75</td>
<td>.0011</td>
</tr>
<tr>
<td>AVLT recognition</td>
<td>12.7 (2.1)</td>
<td>14.1 (1.1)</td>
<td>14.0 (1.3)</td>
<td>5.80</td>
<td>.0051</td>
</tr>
<tr>
<td><strong>Processing speed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS-III PSI</td>
<td>94.9 (14.5)</td>
<td>100.4 (14.6)</td>
<td>97.3 (10.1)</td>
<td>0.85</td>
<td>.43</td>
</tr>
<tr>
<td>COWA</td>
<td>35.3 (8.7)</td>
<td>41.0 (13.9)</td>
<td>37.4 (11.7)</td>
<td>0.98</td>
<td>.38</td>
</tr>
<tr>
<td>TMT part A</td>
<td>32.7 (3.8)</td>
<td>23.1 (4.9)</td>
<td>28.3 (3.7)</td>
<td>1.19</td>
<td>.31</td>
</tr>
<tr>
<td>SCWT color</td>
<td>39.7 (2.4)</td>
<td>48.1 (3.1)</td>
<td>45.8 (2.4)</td>
<td>2.73</td>
<td>.08</td>
</tr>
<tr>
<td><strong>Executive functioning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS-III WMI</td>
<td>96.3 (11.9)</td>
<td>108.4 (15.5)</td>
<td>104.6 (9.9)</td>
<td>5.51</td>
<td>.0061</td>
</tr>
<tr>
<td>TMT part B</td>
<td>73.8 (31.7)</td>
<td>57.2 (15.9)</td>
<td>59.7 (14.4)</td>
<td>3.24</td>
<td>.051</td>
</tr>
<tr>
<td>SCWT interference</td>
<td>53.2 (10.5)</td>
<td>49.1 (9.7)</td>
<td>60.6 (12.1)</td>
<td>4.18</td>
<td>.022</td>
</tr>
</tbody>
</table>

Note: ADHD = attention-deficit/hyperactivity disorder. WAIS-III = Wechsler Adult Intelligence Scale-3rd edition. VCI = Verbal Comprehension Index. POI = Perceptual Organizational Index. WMI = Working Memory Index. PSI = Processing Speed Index. AVLT = Auditory Verbal Learning Test. LOT = learning over trials. IR = immediate recall. DR = delayed recall. TMT = Trail Making Test. COWA = Controlled Oral Word Association. SCWT = Stroop Color and Word Test. 1Invalid group worse than other groups. 2Psychological controls better than other groups.

On every measure, the noncredible group performed significantly worse than the other two groups, which were not significantly different from one another. Groups were also different in executive functioning (WAIS-III WMI, TMT part B, SCWT Interference t score), Wilk’s lambda $F(6, 84) = 3.26, p = .006$. In every case, the noncredible group was significantly worse than the other two groups, which were not significantly different from one another. In addition, the ADHD group also performed significantly worse than the psychological controls on SCWT interference, $p < .01$, see Table 4.

4. Discussion

Results underscore the importance of screening for noncredible performance in referrals for adult ADHD. Individuals failing the WMT were also identified by other indices of noncredible performance, at low sensitivity but 100% specificity in all but one case. Rates of noncredible performance as identified by the WMT were similar to those reported by Sullivan et al. (2007) in another university clinic sample with similar referral concerns. Thus, the current data provide further evidence for the importance of using measures of noncredible performance in adults referring themselves for ADHD evaluation. It is important to note that failure on the WMT does not automatically indicate that the individuals in the noncredible group were malingering (i.e., consciously exaggerating or simulating symptoms for secondary gain). Although information about secondary gain was not collected and/or recorded comprehensively in this clinical sample to warrant its use in our analyses, it was common to see in the clinical charts that individuals who presented for evaluation specifically requested medication for ADHD, academic accommodations, or accommodations for specialized testing such as the GRE or Praxis exams. Of course, requesting treatment or accommodation does not automatically indicate that an individual is malingering. However, future studies should carefully document information about potential secondary gain, as well as include multiple measures of noncredible performance, in order to further explore factors that contribute to noncredible performance in adults referred for ADHD evaluation.

Future studies should consider the development of measures of noncredible performance within self-report ADHD instruments. Our findings suggest that the CAARS inconsistency index is not useful for the assessment of overendorsement of symptoms. Further, $t$ scores of above 80 on CAARS indices, suggested by the test’s manual as raising suspicion of symptom exaggeration, were common in all three-study groups, at least in terms of inattentive symptoms. Further work examining self-reported “ADHD” items that are generally infrequently endorsed may lead to the development of validity indices that are more helpful in the detection of noncredible symptom report in ADHD.
Although we hypothesized that participants who failed the WMT would score higher on self-reported ADHD symptoms than other groups, results did not support this hypothesis. However, the noncredible group was not distinguishable from the ADHD diagnosis group in current ADHD symptom report, both when using mean difference scores and when using clinical cutoffs for age and gender-matched peers. Similarly, the noncredible group was not distinguishable from the ADHD group in retrospective childhood symptom report, using either mean difference scores or clinical cutoffs. Overall these results demonstrate the vulnerability of self-reported ADHD symptoms to noncredible presentation.

Even when controlling for noncredible performance, current ADHD symptoms were frequently reported by the psychological controls, with the exception of CAARS scale f. Retrospective symptom report suggested that ADHD individuals could be distinguished from psychological controls; however, use of a commonly used clinical cutoff for the WURS still resulted in 29% false positives in the psychological symptom group. These results are consistent with other literature suggesting that current and retrospective ADHD symptoms are not specific to ADHD and further emphasize the problems with relying on self-reported symptoms in the differential diagnosis of ADHD in adults.

As hypothesized, individuals who failed the WMT performed worse on neuropsychological tests (specifically those assessing memory and executive functioning) than the other groups. These findings are consistent with Sullivan et al. (2007) in that they show the importance of accounting for noncredible performance when interpreting neuropsychological test performance in individuals referred for ADHD assessment. Of note, the ADHD group was not distinguishable from the psychological symptom group on neuropsychological measures, with the exception of the SCWT interference task. The SCWT is commonly used in ADHD, although results of several meta-analyses have shown mixed results in terms of its effect size in ADHD (Frazier, Demaree, & Youngstrom, 2004; Hervey et al., 2004; Homack & Riccio, 2007) in that they show the importance of accounting for noncredible performance when interpreting neuropsychological test performance in individuals referred for ADHD assessment. Of note, the ADHD group was not distinguishable from the psychological symptom group on neuropsychological measures, with the exception of the SCWT interference task. The SCWT is commonly used in ADHD, although results of several meta-analyses have shown mixed results in terms of its effect size in ADHD (Frazier, Demaree, & Youngstrom, 2004; Hervey et al., 2004; Homack & Riccio, 2004). The present results suggest the interference subscale may be useful in distinguishing adults with ADHD from a control group with psychological symptoms, once noncredible performance has been controlled for.

In addition to limitations already noted, the present study was limited in that some of the measures of interest in analyses played a role in the diagnosis of the participants. It is noteworthy that, for the present analysis, data was taken systematically from clinic charts and each participant was reviewed for inclusion based on the criteria given above. As a result, 20 potential participants who did not meet clear criteria for inclusion for the present study were excluded from analyses. Because this was an archival data set, there was no control for which tests were administered or in what order (or on what day of testing), which is also a study limitation. Furthermore, although no individuals who had been taking ADHD medications were taking them at the time of the neuropsychological evaluation, there was no control for other variables that might have affected neuropsychological test performance (i.e., nicotine or caffeine use, substance use). Finally, as the sample was predominantly young adults in the early years of undergraduate education, the results may not generalize to broader samples of adults referring themselves for concerns about ADHD.

Future studies should include a control group of individuals with learning disability, who also share many of the same cognitive complaints and some of the same neuropsychological weaknesses. For example, Sullivan et al. (2007)’s study suggested that individuals referred for learning disability assessment fail noncredible performance tests such as the WMT at lower rates than individuals referred for ADHD. In addition, future studies should create psychological control groups based on specific diagnoses (e.g., mood disorders, anxiety disorders), in order to see whether there is some specificity to both symptom report and neuropsychological test performance in specific psychological disorders that might be helpful in differential diagnosis of ADHD in adults. Similarly, in future work, neuropsychological profiles associated with specific ADHD subtypes should be examined (once noncredible performance has been controlled for).

Overall, present results are consistent with existing literature on the importance of assessing for noncredible performance in order to understand and interpret both self-report of psychological symptoms and neuropsychological test data. This is particularly important in disorders such as ADHD, where overdiagnosis is a “legitimate concern” (Murphy, 1994). There are major social implications for ADHD diagnosis that affect not only those who are performing in a noncredible manner, but also for those who actually have the disorder and are in need of appropriate treatment and/or accommodation for their condition.

References


