Further Validation of the Conner's Adult Attention Deficit/Hyperactivity Rating Scale Infrequency Index (CII) for Detection of Non-Credible Report of Attention Deficit/Hyperactivity Disorder Symptoms

Carolyn M. Cook, Elizabeth Bolinger, Julie Suhr*

Department of Psychology, Ohio University, Athens, OH, USA

*Corresponding author at: Department of Psychology, 200 Porter Hall, Ohio University, Athens, OH 45701, USA. Tel.: +1-740-593-1091; fax: +1-740-593-0579.
E-mail address: suhr@ohio.edu (J. Suhr).

Abstract

Objective. Attention deficit/hyperactivity disorder (ADHD) can be easily presented in a non-credible manner, through non-credible report of ADHD symptoms and/or by non-credible performance on neuropsychological tests. While most studies have focused on detection of non-credible performance using performance validity tests, there are few studies examining the ability to detect non-credible report of ADHD symptoms. We provide further validation data for a recently developed measure of non-credible ADHD symptom report, the Conner's Adult ADHD Rating Scales (CAARS) Infrequency Index (CII).
Method. Using archival data from 86 adults referred for concerns about ADHD, we examined the accuracy of the CII in detecting extreme scores on the CAARS and invalid reporting on validity indices of the Minnesota Multiphasic Personality Inventory-2 Restructured Format (MMPI-2-RF). We also examined the accuracy of the CII in detecting non-credible performance on standalone and embedded performance validity tests.
Results. The CII was 52% sensitive to extreme scores on CAARS DSM symptom subscales (with 97% specificity) and 20%–36% sensitive to invalid responding on MMPI-2-RF validity scales (with near 90% specificity), providing further evidence for the interpretation of the CII as an indicator of non-credible ADHD symptom report. However, the CII detected only 18% of individuals who failed a standalone performance validity test (Word Memory Test), with 87.8% specificity, and was not accurate in detecting non-credible performance using embedded digit span cutoffs.
Conclusions. Future studies should continue to examine how best to assess for non-credible symptom report in ADHD referrals.

Keywords: Malingering/symptom validity testing; Attention deficit/hyperactivity disorder; Assessment; Self-report; Test construction

Introduction

There is a growing body of evidence that non-credible presentation is a problem in attention deficit/hyperactivity disorder (ADHD) evaluation. Non-credible presentation occurs when individuals being evaluated present in an inaccurate way during the assessment, either by reporting their history or symptoms in an inaccurate fashion (non-credible reporting) and/or by behaving in a manner inconsistent with their actual abilities (non-credible performance). When non-credible presentation is done consciously and with the purpose of receiving external gain, it is often referred to as malingering; however, even in the absence of clear evidence for intent to appear more impaired or external gain, non-credible presentation can occur. Non-credible presentation in ADHD evaluation is a significant problem and has many potential consequences, including substantial costs to society for unnecessary assessments and treatment, unjustified use of limited medical resources, damage to the public's confidence in the effectiveness of treatment, disadvantage to students who do not feign ADHD because others are receiving inappropriate academic

advantage through accommodations, and the passive support of potential drug abuse (Tucha, Fuermaier, Koerts, Groen, & Thome, 2015).

Rates of non-credible presentation in adults presenting for ADHD evaluation range from ∼8% (using very strict criteria for malingering diagnosis) to almost 48% (based on failure rates on performance validity tests or PVTs) (see review by Musso & Gouvier, 2014). It is important to note this is well above the prevalence rate of ADHD (5% of children; 2.4% of adults) (American Psychiatric Association, 2013). Further, these rates do not account for non-credible reporting of symptoms in the absence of failure on PVTs, which is likely of much higher base rate. Thus, when evaluating adults for concerns about ADHD, it is important to assess for the validity of the purported deficits.

Most studies to date have focused on the detection of non-credible presentation in ADHD through the use of PVTs. A review by Tucha and colleagues (2015) concluded that the most promising tests were the Auditory Verbal Learning Test—Exaggeration Index (Barrash, Suhr, & Manzel, 2004), Reliable Digit Span (Babikian, Boone, Lu, & Arnold, 2006), Word Memory Test (Green, 2003), Test of Variables of Attention (Greenberg, Kindschi, Dupuy, & Corman, 1996), and Conners Continuous Performance Test-II (Conners, 2000). However, as Musso and Gouvier (2014) concluded in their qualitative review of both simulated and clinical studies of ADHD malingering, few PVTs have adequate sensitivity to malingering of ADHD when used in isolation (although they generally show specificity around or above 90%). They suggested that measures of non-credible presentation specific to ADHD should be developed.

While many studies have examined the detection of non-credible ADHD presentation through PVTs, few have investigated the detection of non-credible report of ADHD symptoms. This is a significant limitation, given that the most common assessment approach in adult ADHD is self-report, via interviews and/or questionnaires (Bruchmuller, Margraf, & Schneider, 2012; Musso & Gouvier, 2014; Sibley et al., 2012). Because symptoms of ADHD are face valid and well known to the public, it is not surprising that individuals asked to feign ADHD report high levels of both childhood and current ADHD symptoms at levels sometimes, but not always, higher than individuals actually diagnosed with ADHD (Booksh, Pella, Singh, & Gouvier, 2010; Harp, Jasinski, Shandera-Ochsner, Mason, & Berry, 2011; Harrison, Edwards, & Parker, 2007; Jachimowicz & Geiselman, 2004; Jasinski et al., 2011; Quinn, 2003; Sollman, Ranseen, & Berry, 2010; Tucha, Sontag, Walitza, & Lange, 2009). Even outside the context of deliberate exaggeration, however, symptoms of ADHD are reported by individuals with other psychological conditions; thus, it is possible that some non-credible symptom report is related to general psychological distress and/or a communication of a need for treatment/intervention (Lewandowski, Lovett, Codding, & Gordon, 2008; Schoechlin & Engel, 2005; Suhr, Zimak, Buelow, & Fox, 2009; Van Vorhees, Hardy, & Kollins, 2011).

Existing studies on assessment for non-credible symptom reporting in adult ADHD have used simulation designs. For example, Rios and Morey (2013) found that validity scales on the Personality Assessment Inventory-Adolescent accurately detected 38%–82% of individuals asked to simulate ADHD, but with only 73%–81% specificity. Young and Gross (2011) showed that simulators scored higher on many validity scales on the MMPI-2, with the most sensitive scales being Infrequency psychopathology (Fp) (59.4% sensitivity, 94.4% specificity) and the Henry Heilbronner Index (Henry, Heilbronner, Mittenberg, & Enders, 2006) (sensitivity 46.9%, specificity 88.7%). Harp and colleagues (2011) found that ADHD simulators' MMPI-2-RF profiles looked similar to those of honest responders, with the exception of Infrequent Responses revised (F-r). Using a different cutoff for Infrequent Psychopathology Responses revised (Fp-r) detected 63.6% of simulators, holding specificity at 90% or greater.

While some clinical studies have shown that failure on PVTs is associated with higher endorsement of ADHD symptoms (Harrison & Edwards, 2010; Marshall et al., 2010; Suhr, Hammers, Dobbins-Buckland, Zimak, & Hughes, 2008; Sullivan, May, & Galbally, 2007), no studies have attempted to identify cutoffs on ADHD self-report scales that would accurately detect individuals who fail PVTs. However, because non-credible symptom reporting does not always relate to failure on PVTs (Heilbronner et al., 2009; Nelson, Sweet, Berry, Bryant, & Granacher, 2007), there is a need to develop and validate measures of non-credible report of ADHD symptoms using methodology other than simulated malingering designs and using criteria beyond performance on PVTs.

Recently, a measure to detect non-credible report of ADHD symptoms on the Conner's Adult Attention Deficit/Hyperactivity Rating Scale (CAARS; Conners, Erhardt, & Sparrow, 1998) was developed, known as the CAARS Infrequency Index (CII; Suhr, Buelow, & Riddle, 2011). Items for the CII were obtained by identifying items on the CAARS that were infrequently endorsed in a large sample of non-treatment seeking university students, including those with self-reported ADHD diagnosis. Initial validation data were then gathered from a sample of individuals who presented to a university psychology clinic for ADHD evaluation and gave permission for their deidentified clinical data to be used in research. The CII identified 30%–80% of individuals showing extreme scores (≥80) on the CAARS subscales E (DSM Inattentive Symptoms) and/or F (DSM Hyperactive/Impulsive Symptoms), which, according to the CAARS manual (Conners et al., 1998), are suspect for invalidity; specificity was >90%. The CII also accurately detected 24% of individuals who failed a PVT (WMT), with 95% specificity. However, this study did not examine how the CII compared with other self-report validity scales.

In the current study, we further examined the utility of the CII as a measure of non-credible ADHD symptom report by assessing its relationship to other self-report validity scales. Specifically, we expected to replicate results from Suhr and colleagues (2011) demonstrating that the CII could accurately detect extreme scores indicative of over-reporting on the CAARS DSM subscales. We also hypothesized that the CII would accurately detect individuals who showed invalid symptom report on the MMPI-2-RF. Finally, we examined the relationship of the CII to performance on both standalone and embedded PVTs.

Materials and Methods

Participants

Participants included individuals ages 18 and older referred to a university psychology training clinic for neuropsychological evaluation for concerns about ADHD and/or a learning disability and who signed consent at the time of their initial appointment in the clinic for their deidentified clinical data to be used in research being conducted in the training clinic (study approved by the institution's IRB). While some participants were self-referred, others were referred by counselors/psychologists, professors, or physicians. Of note, this archival sample does not overlap with that reported in Suhr and colleagues (2011). All evaluations were conducted by advanced graduate students in clinical psychology under the supervision of a licensed psychologist. Exclusionary criteria included any self-reported history of neurological illness or injury (seizures, moderate to severe traumatic brain injury, brain tumors, etc.). Severity of reported brain injury was determined based on reported loss of consciousness, post-traumatic amnesia, presence/absence of any neurological/neuroimaging findings, and diagnosis given at the time of the injury. Of the 171 individuals in the archival sample, 96 participants completed both the CAARS and the MMPI-II-RF. Of those 96, 10 individuals did not provide consistent data on either the CAARS or the MMPI-2-RF and thus were not included in the present analysis.

The final sample consisted of 86 adults with an average age of 22 (range 18–42, *SD* 5.0) and an average of 14.5 years of education (*SD* 1.9, range 12–18 years). The sample was 53% female, and 79.1% self-identified as Caucasian non-Hispanic (∼3.5% self-identified as Asian American, 3.5% identified as African American, 2.3% identified as Hispanic, and the rest identified as multiple race/ethnic background or other). The final sample of 86 did not differ in any significant way on any demographic variables from the total archival sample of 171 individuals. The sample included 4 individuals diagnosed with ADHD (who did not perform non-credibly on any of the measures in the study); an additional 11 individuals were diagnosed with a mood disorder (usually major depression), 14 with an anxiety disorder, 8 with an adjustment disorder, 4 with a substance abuse disorder, 1 with a math learning disability, and 1 with schizotypal personality disorder.

Procedure

Data were collected as part of each participant's clinical evaluation, and thus measures of interest for the present analyses were not presented in a fixed order and were administered as part of a much larger set of measures that varied based on clinical presentation, presenting problems, and history. All participants provided demographic information and medical and psychological histories as part of a clinical interview. They also completed psychological measures and neuropsychological tests as part of a comprehensive assessment for ADHD and any additional clinical concerns.

Measures

Non-credible report. The CAARS Self-Report Long Form (Conners et al., 1998) is a 66-item self-report measure of ADHD symptoms. Responses are scored on a 4-point scale, where 0 = not at all, 1 = just a little, 2 = pretty much, and 3 = very much. Test–retest reliabilities have been shown to be strong (Erhardt, Epstein, Conners, Parker, & Sitarenios, 1999). The CAARS is highly correlated with other self-report ADHD measures, and initial studies on diagnostic accuracy for adult ADHD found the CAARS to have diagnostic sensitivity of 82% and specificity of 87% relative to healthy controls (Erhardt et al., 1999). The Inconsistency Index was used in the present study to exclude individuals who appeared to be responding inconsistently to this measure (as noted above), using standard cutoffs provided in the clinical manual (Conners et al., 1998). For the present analyses, we calculated the CII from item scores on the CAARS; CII scoring is described in detail in Suhr and colleagues (2011). In addition to the Inconsistency Index and the CII, we also used extreme scores ($T \geq 80$) on the CAARS DSM-based subscales (E, F, and G) as indicators of non-credible symptom reporting, based on recommendations in the clinical manual (Conners et al., 1998).

The MMPI-2-RF is a true/false self-report measure of psychopathology with well-established validity data (Ben-Porath & Tellegen, 2008; Tellegen & Ben-Porath, 2011). The Variable Response Inconsistency revised and the True Response Inconsistency revised subscales were used to exclude participants who responded inconsistently on the MMPI-2-RF, as noted above. The MMPI-2-RF validity scales used in the present analyses were Infrequent Responses revised (F-r), Infrequent

Psychopathology Responses revised (Fp-r), Infrequent Somatic Responses (Fs), Symptom Validity revised (FBS-r), and the Response Bias Scale (RBS). F-r, Fp-r, Fs, FBS-r, and RBS have all been demonstrated to successfully identify symptom over-reporting (Ben-Porath & Tellegen, 2008; Tellegen & Ben-Porath, 2011). We used MMPI-2-RF manual guidelines to identify cutoffs for over-reporting for each validity scale (≥ 79 for F-r, ≥ 70 for Fp-r, ≥ 80 for all others; Ben-Porath & Tellegen, 2008).

Non-credible performance. WMT (Green, 2003) is a well-validated PVT, with high diagnostic accuracy in a variety of clinical populations (Bauer, O'Bryant, Lynch, McCaffrey, & Fisher, 2007; Green, 2003). For the present analyses, passing or failing the WMT was based on performance on the first three subtests, using standard cutoffs recommended in the manual. Failing any one of the first three subtests was coded as a failure on the WMT.

Several cutoff scores calculated from the Digit Span subtest of the Wechsler Adult Intelligence Scale—IV (Wechsler, 2008) were used as another indicator of non-credible performance, and included longest digits forward < 5, longest digits backward < 4, age-corrected scaled score < 7, and Reliable Digit Span < 8 (sum of the longest string of digits an individual can repeat forward without error over two trials and the longest string of digits an individual can repeat backward without error over two trials; Greiffenstein, Baker, & Gola, 1994). Failure on any one of these indices was coded as a failure on Digit Span as a PVT.

Data Analysis

For tests of the main study hypotheses regarding the CII's ability to detect non-credible self-report, receiver operating characteristic (ROC) analyses were utilized, with extreme scores on the CAARS and invalid scores on the MMPI-2-RF validity scales as criterion measures (using cutoffs defined above). In addition to ROC analyses, we also attempted to identify CII cutoffs that would maximize specificity (defined as near or above 90%). For tests of the exploratory analyses regarding the ability of the CII to detect non-credible performance, ROC analyses were conducted with non-credible and credible criterion groups established using the WMT and Digit Span, as defined above. We also attempted to identify CII cutoffs that would maximize specificity to PVT failure.

Results

CII and Other Conner's Adult Attention Deficit/Hyperactivity Rating Scale Scores

Of the whole sample, 51 participants scored ≥ 80 on one of the CAARS scales E, F, or G, while 35 did not. Using CII as a predictor, AUC was 0.87, $SE = 0.04$, $p < .001$, 95% CI (0.80, 0.95). To use a more extreme definition of non-credible CAARS report, we also divided the groups by individuals who scored ≥ 80 on at least two of the CAARS E, F, or G subscales. Of the whole sample, 31 participants did not score ≥ 80 on any of CAARS scales E, F, or G, while 33 scored ≥ 80 on at least two of CAARS scales E, F, or G; the remaining participants were excluded because they did not meet criteria for either of these groups. Using CII as a predictor, AUC was 0.95, $SE = 0.03$, $p < .001$, 95% CI (0.89, 1.0). Analysis of the ROC curve suggested a cutoff of 21 or higher on the CII was 52% sensitive to non-credible self-report (extreme scores on at least two of CAARS scales E, F, and G), with 97% specificity.

CII and Minnesota Multiphasic Personality Inventory-2 Restructured Format

Table 1 presents the ROC analyses for the CII as a predictor of invalidity on each of the MMPI-2-RF validity scales. AUCs were at the low end of moderate in accuracy (as defined by Strainer & Cairney, 2007) for all validity scales, ranging from 70% to 72% except Fp-r, which was 68%. The ROC curves were examined to find cutoffs on the CII that would create specificity values at or

Table 1. Area under the curve (AUC) findings for the Conner's Adult Attention Deficit/Hyperactivity Rating Scale Infrequency Index for Minnesota Multiphasic Personality Inventory-2-Restructured Format Validity Scales

	AUC	Standard error	*p*-value	95% confidence interval
F-r	0.70	0.06	.009	0.58–0.83
Fp-r	0.68	0.07	.02	0.55–0.81
Fs	0.72	0.07	.01	0.59–0.85
FBS-r	0.70	0.07	.04	0.56–0.84
RBS	0.70	0.07	.01	0.57–0.83

Note: F-r = Infrequent Responses revised; Fp-r = Infrequent Psychopathology Responses revised; FBS-r = Symptom Validity revised; RBS = Response Bias Scale.

close to 90%; these analyses suggested a cutoff of >22 on the CII. Table 2 provides the accuracy values at this cutoff relative to each MMPI-2-RF validity scale. Sensitivities ranged from 20% (FBS-r) to 36% (Fs).

CII and Non-Credible Performance on Standalone and Embedded Performance Validity Tests

Of the whole sample, 28 individuals failed at least one of the first three subtests of the WMT. The CII showed some accuracy in the detection of WMT failure, $AUC = 0.67$, $SE = 0.076$, $p = .01$, 95% CI (0.55, 0.80). To maximize specificity, a score of >21 was selected; this score was 88.9% specific and 17.8% sensitive to failing the WMT.

Of the whole sample, 16 failed at least one of the Digit Span PVT indicators. The CII was not accurate in the detection of Digit Span failure, $AUC = 0.56$, $SE = 0.08$, $p = .47$, 95% CI (0.40, 0.72). The score of >21 on CII only detected 13% of those who failed the Digit Span PVT indicators, with 87% specificity.

Of note, we also explored whether RBS, an MMPI-2-RF validity scale developed by identifying items that distinguished individuals who failed PVTs, could accurately identify individuals who failed either the WMT or the Digit Span indicators. The ROC analysis for WMT failure was not significant, $AUC = 0.55$, $SE = 0.07$, $p = .48$, 95% CI (0.41, 0.69). The ROC analysis for Digit Span was also not significant, $AUC = 0.57$, $SE = 0.09$, $p = .37$, 95% CI (0.40, 0.75).

Discussion

The purpose of the present study was to further examine the utility of the CII as a measure of non-credible ADHD symptom report. A cutoff of 21 or higher on the CII was 52% sensitive to non-credible self-report as assessed by extreme CAARS scale scores, with 97% specificity. This replicates and extends previous findings from Suhr and colleagues (2011) and shows the CII can detect likely non-credible symptom report on the CAARS DSM-based subscales.

At near 90% specificity to invalid symptom report on the MMPI-2-RF, the CII (with a cutoff of >22) showed sensitivity of 20%−36% for detecting invalidity on various MMPI-2-RF validity scales. This suggests that the CII is a less sensitive predictor of invalidity on the MMPI-2-RF validity scales than on the CAARS DSM-based subscales. One potential reason for the lower accuracy is test-related method variance. That is, while CAARS responses are scored on a 4-point scale, MMPI-2-RF responses are scored as true or false. Individuals may respond differently to these different formats, which may impact the ability of the CII to predict invalidity on MMPI-2-RF scales.

Another possible reason for the difference in sensitivity is the nature of the constructs being measured by the two instruments. While the CAARS DSM-based subscales assess for the presence of symptoms of ADHD, MMPI-2-RF validity scale items are focused on infrequently endorsed psychiatric and medical symptoms. Furthermore, although the FBS-r includes infrequently endorsed cognitive complaints, the cognitive complaints assessed by the FBS-r scale are more closely tied to diffuse and/or vague medical and/or neurological complaints, rather than the type of complaints frequently endorsed in ADHD evaluations.

Results from PVT analyses showed that a CII cutoff of >21 was 88.9% specific to failure on the WMT, but only detected ∼18% of those who failed the WMT. The CII was not able to accurately detect failure on embedded Digit Span. Digit Span findings might be explained by the fact that the sample was high functioning and the rates of failure on the embedded Digit Span indicators were low. Notably, the MMPI-2-RF RBS scale was not able to accurately detect failure on either the WMT or the Digit Span PVT cutoffs. The RBS findings are somewhat surprising, given the RBS was developed using regression analyses to identify MMPI-2 items that predicted invalid performance on several well-validated memory tests in a large sample of non-head-injured disability claimants (Gervais, Ben-Porath, Wygant, & Green, 2008). Differences in our findings may be related to differences in the samples utilized in these studies (e.g., disability status, referral question, demographics, general level of functioning) and speak to the need to develop ADHD-specific indicators of non-credible presentation.

Table 2. Sensitivity and specificity of Conner's Adult Attention Deficit/Hyperactivity Rating Scale Infrequency Index >22 relative to invalidity on Minnesota Multiphasic Personality Inventory-2-Restructured Format Validity Scales

Criterion scale	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
F-r	24.5	89.5	36.3	82.1
Fp-r	28.5	89.3	36.3	88.8
Fs	36.3	90	36.3	85.1
FBS-r	20	91.8	25	89.4
RBS	33.3	89.7	36.3	83.5

Note: F-r = Infrequent Responses revised; Fp-r = Infrequent Psychopathology Responses revised; FBS-r = Symptom Validity revised; RBS = Response Bias Scale.

In addition, prior research shows there is often a discrepancy between self-reported symptoms and observed or reported behavior, particularly in the context of non-credible presentation (Heilbronner et al., 2009; Nelson et al., 2007). It is possible that participants in our sample reported symptoms in a non-credible manner but performed in a more credible manner on behavioral tests not perceived to be directly linked to ADHD symptomatology. Thus, it is important to routinely assess for accuracy of symptom report in addition to assessing for non-credible performance through the use of PVTs.

Given that, in many settings, adults are diagnosed with ADHD on the basis of self-report alone, there remains a need to develop measures to identify non-credible ADHD symptom report. Previous work (Suhr et al., 2011) and the present study results suggest that the CII is useful for detecting extreme scores on the CAARS DSM-based scales that are potentially indicative of non-credible symptom report. The present study results also suggest that the CII is useful, albeit with minimal sensitivity, for detecting invalid symptom report as measured by the MMPI-2-RF. Thus, elevated scores on the CII should raise suspicion about the self-reported ADHD symptoms of an individual being evaluated for this condition. Although the limited sensitivity of the CII means that not all individuals who report symptoms in a non-credible fashion will be identified, the high specificity of the cutoff means that there is a low risk of false positives. In addition, a lower cutoff score may improve the sensitivity of the CII while still obtaining adequate specificity, particularly in the context of multiple indicators of non-credible presentation. Further replication and research in different clinical contexts and analogue studies would provide additional data on the appropriateness of the cut scores identified in this and previous studies (Suhr et al., 2011). Although the present results were promising, there is still value in including other measures of non-credible symptom report in ADHD evaluation, given that a large number of individuals who showed invalid self-report on the MMPI-2-RF were not detected by the CII. Further, the results point to the need to include evaluation for both non-credible report and non-credible performance (via PVTs) in the assessment of ADHD.

Conflict of Interest

None declared.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Babikian, T., Boone, K. B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various digit span scores in the detection of suspect effort. *The Clinical Neuropsychologist*, *20*, 145–159.

Barrash, J., Suhr, J., & Manzel, K. (2004). Detecting poor effort and malingering with an expanded version of the Auditory Verbal Learning Test (AVLTX): Validation with clinical samples. *Journal of Clinical and Experimental Neuropsychology*, *26*, 125–140.

Bauer, L., O'Bryant, S. E., Lynch, J. K., McCaffrey, R. J., & Fisher, J. M. (2007). Examining the test of memory malingering trial 1 and word memory test immediate recognition as screening tools for insufficient effort. *Assessment*, *14*, 215–222.

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI-2-RF: Manual for administration, scoring, and interpretation*. Minneapolis: University of Minnesota Press.

Booksh, R. L., Pella, R. D., Singh, A. N., & Gouvier, W. D. (2010). Ability of college students to simulate ADHD on objective measures of attention. *Journal of Attention Disorders*, *13*, 325–338.

Bruchmuller, K., Margraf, J., & Schneider, S. (2012). Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. *Journal of Consulting and Clinical Psychology*, *80*, 128–138.

Conners, C. (2000). *Conners' continuous performance test II*. Tonawanda, NY: Multi Health Systems.

Conners, C. K., Erhardt, D., & Sparrow, E. P. (1998). *Conners adult attention rating scale—Self-Report: Long version*. North Tonawanda, NY: Multi-Health Systems.

Erhardt, D., Epstein, J. N., Conners, C. K., Parker, J. D. A., & Sitarenios, G. (1999). Self-ratings of ADHD symptoms in adults II: Reliability, validity, and diagnostic sensitivity. *Journal of Attention Disorders*, *3*, 153–158.

Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2008). Differential sensitivity of the Response Bias Scale (RBS) and MMPI-2 validity scales to memory complaints. *The Clinical Neuropsychologist*, *22*, 1061–1079.

Green, P. (2003). *The word memory test*. Alberta, CA: Green's Publishing.

Greenberg, L., Kindschi, C., Dupuy, T., & Corman, C. (1996). *Test of variables of attention*. Los Alamitos, CA: Universal Attention Disorders.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*, 218–224.

Harp, J. P., Jasinski, L. J., Shandera-Ochsner, A. L., Mason, L. H., & Berry, D. T. (2011). Detection of malingered ADHD using the MMPI-2-RF. *Psychological Injury and Law*, *4*, 32–43.

Harrison, A. G., & Edwards, M. J. (2010). Symptom exaggeration in post-secondary students: Preliminary base rates in a Canadian sample. *Applied Neuropsychology*, *17*, 135–143.

Harrison, A. G., Edwards, M. J., & Parker, K. C. (2007). Identifying students faking ADHD: Preliminary findings and strategies for detection. *Archives of Clinical Neuropsychology*, *22*, 577–588.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., & Millis, S. R., & Conference Participants (2009). American Academy of Clinical Neuropsychology Consensus Conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, *23*, 1093–1129.

Henry, G. K., Heilbronner, R. L., Mittenberg, W., & Enders, C. (2006). The Henry–Heilbronner Index: A 15-item empirically derived MMPI-2 subscale for iden-
    tifying probable malingering in personal injury litigants and disability claimants. *The Clinical Neuropsychologist*, *20*, 786–797.
Jachimowicz, G., & Geiselman, R. E. (2004). Comparison of ease of falsification of attention deficit hyperactivity disorder diagnosis using standard behavioral
    rating scales. *Cognitive Science Online*, *2*, 6–20.
Jasinski, L. J., Harp, J. P., Berry, D. T., Shandera-Ochsner, A. L., Mason, L. H., & Ranseen, J. D. (2011). Using symptom validity tests to detect malingered ADHD in
    college students. *The Clinical Neuropsychologist*, *25*, 1415–1428.
Lewandowski, L. J., Lovett, B. J., Codding, R. S., & Gordon, M. (2008). Symptoms of ADHD and academic concerns in college students with and without ADHD
    diagnoses. *Journal of Attention Disorders*, *12*, 156–161.
Marshall, P., Schroeder, R., O'Brien, J., Fischer, R., Ries, A., Blesi, B., et al. (2010). Effectiveness of symptom validity measures in identifying cognitive and
    behavioral symptom exaggeration in adult attention deficit hyperactivity disorder. *The Clinical Neuropsychologist*, *24*, 1204–1237.
Musso, M. W., & Gouvier, W. D. (2014). "Why is this so hard?": A review of detection of malingered ADHD in college students. *Journal of Attention Disorders*, *18*,
    186–201.
Nelson, N. W., Sweet, J. J., Berry, D. T. R., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic
    evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society*, *13*, 440–449.
Quinn, C. A. (2003). Detection of malingering in assessment of adult ADHD. *Archives of Clinical Neuropsychology*, *18*, 379–395.
Rios, J., & Morey, L. C. (2013). Detecting feigned ADHD in later adolescence: An examination of three PAI–A negative distortion indicators. *Journal of
    Personality Assessment*, *95*, 594–599.
Schoechlin, C., & Engel, R. R. (2005). Neuropsychological performance in adult attention-deficit hyperactivity disorder: Meta-analysis of empirical data. *Archives
    of Clinical Neuropsychology*, *20*, 727–744.
Sibley, M. H., Pelham, W. E., Molina, B. S. G., Gnagy, E. M., Waxmonsky, J. G., Waschbusch, D. A., et al. (2012). When diagnosing ADHD in young adults em-
    phasize informant reports, DSM items, and impairment. *Journal of Consulting and Clinical Psychology*, *80*, 1052–1061.
Sollman, M. J., Ranseen, J. D., & Berry, D. T. (2010). Detection of feigned ADHD in college students. *Psychological Assessment*, *22*, 325–335.
Strainer, D. L., & Cairney, J. (2007). What's under the ROC? An introduction to receiver operating characteristics curves. *The Canadian Journal of Psychiatry*, *52*,
    121–128.
Suhr, J., Hammers, D., Dobbins-Buckland, K., Zimak, E., & Hughes, C. (2008). The relationship of malingering test failure to self-reported symptoms and neuro-
    psychological findings in adults referred for ADHD evaluation. *Archives of Clinical Neuropsychology*, *23*, 521–530.
Suhr, J., Zimak, E., Buelow, M., & Fox, L. (2009). Self-reported childhood attention-deficit/hyperactivity disorders symptoms are not specific to the disorder.
    Comprehensive Psychiatry, *50*, 269–275.
Suhr, J. A., Buelow, M., & Riddle, T. (2011). Development of an infrequency index for the CAARS. *Journal of Psychoeducational Assessment*, *29*, 160–170.
Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assess-
    ments. *Applied Neuropsychology*, *14*, 189–207.
Tellegen, A., & Ben-Porath, Y. S. (2011). *Minnesota multiphasic personality inventory-2 restructured form technical manual*. Minneapolis: University of
    Minnesota Press.
Tucha, L., Fuermaier, A. B., Koerts, J., Groen, Y., & Thome, J. (2015). Detection of feigned attention deficit hyperactivity disorder. *Journal of Neural Transmission*,
    122(*Suppl. 1*), S123–S134.
Tucha, L., Sontag, T. A., Walitza, S., & Lange, K. W. (2009). Detection of malingered attention deficit hyperactivity disorder. *Attention Deficit and Hyperactivity
    Disorders*, *1*, 47–53.
Van Vorhees, E. E., Hardy, K. K., & Kollins, S. H. (2011). Reliability and validity of self- and other-ratings of symptoms of ADHD in adults. *Journal of Attention
    Disorders*, *15*, 224–234.
Wechsler, D. (2008). *WAIS-IV: Administration and scoring manual*. San Antonio, TX: The Psychological Corporation.
Young, J. C., & Gross, A. M. (2011). Detection of response bias and noncredible performance in adult attention-deficit/hyperactivity disorder. *Archives of Clinical
    Neuropsychology*, *26*, 165–175.