Development of an Infrequency Index for the CAARS

Julie A. Suhr[1], Melissa Buelow[1],
and Tara Riddle[1]

Abstract

There is a clinical need for measurement of noncredible self-reporting of symptoms of attention deficit hyperactivity disorder (ADHD) in adults presenting for ADHD evaluation. The present study describes the development of initial validity data for an Infrequency Index for the Conner's Adult Attention Deficit/Hyperactivity Rating Scale (CII). Items for the CII were obtained from a large sample of nontreatment seeking university students, including individuals with a self-reported history of ADHD diagnosis. Items endorsed infrequently in the sample, including those with ADHD diagnoses, were identified and summed to create the CII. Initial validation data were gathered from a sample of individuals seeking clinical evaluation for ADHD. The CII was strongly related to noncredibly high symptom report and was also related to noncredible performance on cognitive measures. Results provide initial support for the CII's use in assessing noncredible overreporting on the Conner's Adult Attention-Deficit/Hyperactivity Rating Scale.

Diagnosis of attention deficit hyperactivity disorder (ADHD) in adulthood is often based on clinical interview and completion of self-report scales that assess for current and childhood symptoms consistent with the disorder. However, there are several problems with reliance on self-report for ADHD diagnosis, including the high base rate of endorsement of "ADHD" symptoms in both general and clinical populations and the face validity of ADHD symptoms, which makes them vulnerable to noncredible responding.

Both current and childhood ADHD symptoms are endorsed with at high rates in the general population, making it difficult to rely on self-reported symptoms for ADHD diagnosis (DuPaul et al., 2001; Heiligenstein, Conyers, Berns, Miller, & Smith, 1998; Murphy & Barkley, 1996; Murphy, Gordon, & Barkley, 2000; Suhr, Zimak, Buelow, & Fox, 2009; Weyandt, Linterman, & Rice, 1995). This problem is seen not only when adult ADHD symptoms are assessed via self-report questionnaires but also with structured clinical interviews (Mannuzza, Klein, Klein, Bessler, & Shrout, 2002). Data from clinical and/or treatment-seeking samples provide even more powerful

[1]Ohio University, Athens, OH, USA

Corresponding Author:
Julie A. Suhr, Department of Psychology, 200 Porter Hall, Ohio University, Athens, OH 45701, USA
Email: suhr@ohio.edu

evidence against reliance on self-reported ADHD symptoms in assessment and diagnosis, as both current and childhood symptoms are endorsed at extremely high rates by adults self-referred for a variety of medical and psychological reasons (Harrison, 2004; McCann & Roy-Byrne, 2004; Roy-Byrne et al., 1997). Such findings should lead to healthy skepticism regarding interpretation of reported ADHD symptoms as specifically indicative of an ADHD diagnosis.

One factor that may contribute to the high base rate of symptom report in clinical samples is that some individuals could be motivated to present in a noncredible fashion, by overreporting number of symptoms, severity of symptoms, or degree to which they impair everyday functioning. There are many potential reasons for noncredible symptom reporting, including report of any and all psychological symptoms as a "cry for help" in someone experiencing a high rate of general psychological distress or demoralization, or consciously exaggerating or reporting symptoms in an attempt to receive secondary gain (i.e., malingering). Regardless of the reason for the behavior, however, noncredible symptom report can and should be assessed for in clinical assessment. Nevertheless, it remains clear that although many psychological instruments include subscales for the assessment of noncredible responding, existing adult ADHD symptom report measures and interviews do not (Quinn, 2003).

With regard to malingering as a reason for noncredible symptom report in ADHD, there is ample evidence for potential secondary gain, including receipt of psychostimulant medications, academic or work accommodations, or accommodations for high-stakes testing situations such as the GRE, LSAT, or MCAT exams (Alfano & Boone, 2007; Sullivan, May, & Galbally, 2007). Because ADHD symptoms are easily recognized as symptoms of ADHD because of widely available public information about the disorder (Alfano & Boone, 2007; Conti, 2004; Murphy, 1994; Quinn, 2003), it would not be difficult for a motivated individual to identify the symptoms that should be endorsed, allowing for malingering on self-report scales or even in a structured interview. There is evidence that self-reported ADHD symptoms are easy to simulate. Studies have shown that individuals asked to simulate ADHD symptoms either report similar or higher levels of both childhood and current ADHD symptoms than individuals diagnosed with ADHD (Lee Booksh, Pella, Singh, & Gouvier, 2010; Harrison, Edwards, & Parker, 2007; Jachimowicz & Geiselman, 2004; Quinn, 2003). There is also evidence for high rates of noncredible cognitive performance in those who refer themselves for ADHD evaluation (Suhr, Hammers, Dobbins-Buckland, Zimak, & Hughes, 2008; Sullivan et al., 2007). It may be difficult to distinguish noncredible symptom report as a means to secondary gain from noncredible symptom report related to significant distress; however, noncredible symptom report can be measured and its impact on assessment results can be considered regardless of the motivations for the noncredible behavior.

A common method for assessment of noncredible responding of self-reported psychological symptoms is based on the premise that individuals who are exaggerating or even malingering their symptoms are more likely to endorse symptoms that are infrequently endorsed by others (i.e., an overreporting response bias). Examples of such methods include the F family of the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher et al., 2001) and Minnesota Multiphasic Personality Inventory-2–Restructured Form (MMPI-2 RF; Ben-Porath & Tellegen, 2008), and the Critical Items of the Personality Assessment Inventory (PAI; Morey, 2007). Typically, infrequency scales focus on the overreport of broadly defined psychopathology rather than specific sets of symptoms or symptoms related to a specific disorder. One exception is the Structured Inventory of Malingered Symptoms (Widows & Smith, 2005), a standalone self-report measure of noncredible psychological and neuropsychological symptoms with subscales for feigned cognitive impairment, including neurologic symptoms, amnestic symptoms, and low intelligence. Ideally, however, an infrequency scale would consist of items embedded within a valid measure of the construct being assessed, to make it less readily apparent to someone consciously trying to simulate a disorder.

The Conner's Adult Attention Deficit/Hyperactivity Rating Scale (CAARS; Conners, Erhardt, & Sparrow, 1998) is a popular self-report instrument for adults who present with concerns about ADHD. Although the CAARS has a validity index, the Inconsistency Index, it is designed to assess inconsistency in responding to items that measure similar content, and thus is not a measure of an overreporting response bias. The CAARS manual also suggests that T scores above 80 on any of the subtests should be considered possible evidence for invalidity, in the form of symptom exaggeration, although the manual cautions that scores this high could validly represent severity of symptoms in individuals with ADHD. The purpose of the present study was to develop an infrequency index for the CAARS and present preliminary data on its validity as a measure of noncredible ADHD symptom report.

Study 1: Development of the CAARS Infrequency Index (CII)

Method

Participants. Participants were 1,173 individuals from a large Midwestern university who were part of a much larger institutional review board–approved study of personality, affect, and neuropsychological correlates of ADHD diagnosis and ADHD symptom report in nontreatment-seeking undergraduate students. The data for the present analyses were taken from a deidentified database of participants who had completed the screening and diagnosis phase of the larger project, involving completion of several self-report measures, between November 2005 and July 2007, and for whom complete information from key study measures for the present study (CAARS, demographic, and diagnostic history variables) was available. Briefly, the 1,173 individuals participated in small groups of about 10 to 20 participants at a time, under the observation of one or two trained research assistants, and completed about 2 hours of questionnaires to see if they would qualify for a second study in which neuropsychological correlates of adult ADHD diagnosis were to be explored.

Participants ranged in age from 18 to 25 years (average age = 19 years), with 401 male participants. Participants self-reported their racial/ethnic status as follows: 1,080 White, 5 Native American, 14 Asian/Asian American, 53 Black/African American, 22 Hispanic, 17 other racial/ethnic identity, and 2 unreported. For some analyses, participants were further divided into the following groups.

The *ADHD group* included 71 individuals (36 male) who reported having received a diagnosis of ADHD in the past, regardless of current diagnosis or treatment for any other psychological condition or score on a self-report depression scale.

The *Psychological Control group* included 147 individuals (30 male) who denied having a prior diagnosis of ADHD, but who reported that they were currently diagnosed with/receiving treatment for a psychological condition and/or had scores on a self-report depression scale that suggested current depressive symptoms in at least the moderately severe range.

The *Control group* included 955 individuals (335 male) who reported no prior diagnosis of ADHD, no current diagnosis and/or treatment for a psychological condition, and who had scores on a self-report measure of depressive symptoms indicative of minimal to no depression.

Groups were not different in age; however, there were differences in the distribution of gender among the groups. Proportionally, more females were in the psychological control (80% female) and normal control (65% female) groups, compared with 49% in the ADHD group, $\chi^2(2) = 21.35$, $p < .001$. In the ADHD group, 40 reported current ADHD medication use (5 Ritalin, 14 Adderall, 2 Strattera, 9 Concerta, 1 Focalin, 9 other/multiple medications). Of note, all individuals were instructed to answer CAARS items as though they were not on medication.

Measures. Participants completed questionnaires assessing basic demographic characteristics (e.g., age, ethnicity, high school and college grade point average) as well as relevant psychological and physical history, including ADHD diagnostic status, ADHD medication status, functional impairment related to the presence of ADHD symptoms if applicable, presence of alcohol and substance use difficulties, current and past psychiatric status (including history of treatment), and history of head injury.

The CAARS Self-Report–Long Form (Conners et al., 1998) is a 66-item inventory of self-reported ADHD symptoms. Responses are scored on a 4-point scale (0 = *not at all*; 1 = *just a little*; 2 = *pretty much*; 3 = *very much*). Test–retest reliabilities are strong (Erhardt, Epstein, Conners, Parker, & Sitarenios, 1999), and the measure is highly correlated with other self-report ADHD measures, as well as showing good sensitivity and specificity data in initial studies of diagnostic accuracy for adult ADHD (Erhardt et al., 1999).

Scores from the Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996) were examined to assess for the presence of current depressive symptomatology and its relation to noncredible symptom report as well as for group placement as described above. The BDI-II is a 21-item self-report measure of the presence and severity of depressive symptoms in adolescents and adults ages 13 and older. Items on the BDI-II correspond with criteria for *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (*DSM-IV*; American Psychiatric Association, 1994) depressive disorders. Participants are asked to select the one out of four statements that best characterize their mood in the past 2 weeks. Responses are summed to yield a total symptom score, with higher scores representing greater symptom severity. The BDI-II shows strong internal consistency and is highly correlated with other measures of depressive symptomatology (Beck et al., 1996).

Statistical analyses. For initial selection of infrequently endorsed items, responses to CAARS items were recoded as follows: Items endorsed with a 2 or 3 were coded 1, and items endorsed with a 0 or 1 were coded 0. We chose this method because all CAARS items are coded such that higher scores mean more pathology, and we assumed that noncredible presentation of ADHD symptoms would lead to overendorsement of pathological symptoms. Frequency analyses were then conducted for all CAARS items, collapsed across group assignment.

As the goal was to find infrequently endorsed items in order to minimize false positives, items in which at least 90% of total respondents scored a 0 on the binary scale were selected for further analysis (12 items total; item numbers 21, 22, 23, 26, 30, 34, 43, 45, 49, 51, 52, 62). Two items (49, 51) come from Scale A (Inattention/Memory Problems), three items (30, 43, 52) come from Scale C (Impulsivity/Emotional Lability), three items (21, 22, 62) come from Scale F (*DSM-IV* Hyperactive/Impulsive Symptoms), two items come from Scale H (ADHD Index), and two items come from more than one scale (23, 26). These 12 items were each endorsed as occurring "pretty much, often" to "very much, very frequently" by 10% or less of the total sample of participants.

A total infrequency score (CII) was then created by summing the original scores for the 12 items, thus creating a scale with a 0-36 score range. We decided to use the original scoring method for ease of calculation for future examiners, should the index prove to be valid.

After the initial item selection and creation of the CII using the entire data set, CII scores were calculated for the three groups outlined above. We then compared frequency of CII scores across groups, with the goal of finding a cutoff score to maximize specificity in all groups (at least 90% of participants in all groups falling below that score). Given our desire to minimize diagnosis of false positives, we chose to identify a value that would maximize specificity in all groups. To explore the relationship of CII to other evidence for noncredible self-report, CII scores were also compared with other indices of CAARS validity (Inconsistency Index, *T* scores on CAARS DSM Inattentive [E] and *DSM* Hyperactive/Impulsive [F] scales).

Results

Cronbach's alpha for the CII was .86, which is quite acceptable considering the low base rate with which any one of these items occurred in the overall sample.

A score of 20 or less on the CII occurred in 90.1% of the ADHD group, 94.6% of the psychological control group, and 99.5% of the normal control group. Thus, scores of 21 or higher on the CII occurred infrequently in all the study groups. However, there were gender differences on the CII, with males ($\chi^2 = 6.03$, $SD = 5.59$) scoring significantly higher than females ($\chi^2 = 4.91$, $SD = 4.66$), $t(1171) = 3.65$, $p < .001$. A cutoff score of 21 was needed for 90% or greater specificity in all groups for males (i.e., scores of 22 or higher occur infrequently in all groups), whereas a cutoff score of 20 was adequate for at least 90% specificity across all female groups.

Although we did not expect that an infrequency scale would be strongly related to inconsistency in responding, given they are two different types of noncredible responding, we examined the relationship of the CII to the Inconsistency Index of the CAARS. In the whole sample, 85 individuals were noncredibly inconsistent using stated cutoffs for the Inconsistency Index in the CAARS manual. Using a cutoff of 20, the CII identified only 4% of those who were noncredibly inconsistent.

To determine whether the CII was related to report of clinically significant ADHD symptoms, we examined the relationship of CII scores to clinical elevations (*T* score of 65 or higher) on the *DSM* Inattentive (E) and Hyperactive/Impulsive (F) scales of the CAARS. In the whole sample, 16% obtained a clinical elevation on E (34% of the ADHD group, 31% of psychological controls, 9% of controls), whereas 7% obtained a clinical elevation on F (17% of the ADHD group, 12% of psychological controls, 4% of controls). A score of 21 or higher on the CII identified 12% of those scoring clinically high on the E subscale, with >99% specificity among those not clinically high on the E subscale. A score of 21 or higher on the CII identified 23% of those scoring clinically high on the F subscale, with >99% specificity.

Finally, we examined the relationship of CII scores to elevations on the CAARS that are suggestive of noncredible self-report. As mentioned above, the CAARS manual recommends a cutoff of $T = 80$ or higher as indicative of noncredible responding on the CAARS. In the whole sample, very few people obtained CAARS E and F subscale *T* scores of 80 or higher (4% on E, 1% on F). However, a CII of 21 or higher identified 28% of those scoring at *T* score of 80 or higher on E (with 92% specificity) and 47% of those scoring at *T* score of 80 or higher on F (with 99% specificity).

Study 1 Discussion

The CII showed good internal consistency. A cutoff on the CII was identified that occurred infrequently in this nontreatment-seeking sample, even in those previously diagnosed with ADHD or experiencing current psychological distress. Findings raised some concerns about gender differences on the CII, with the possibility that a slightly higher cutoff would be necessary for males. Of note, *T* scores of 65 and above on CAARS subscales were relatively common, providing further evidence of the nonspecificity of ADHD symptoms, particularly given that this elevated level of symptoms was seen outside of a clinical context (i.e., non-treatment-seeking healthy young adults participating in research). However, the CII identified only a small number of those reporting clinically high symptoms, suggesting the CII is tapping something other than report of clinically elevated ADHD-like symptoms.

Consistent with expectations and its intended use, the CII score was more strongly related to overreporting of symptoms on the CAARS than to inconsistency in responding. Furthermore, the

CII was specific (but with low sensitivity) to extreme CAARS subscale elevations suggestive of invalid performance, also consistent with its intended use.

An important note is that this sample was exclusively university students, and thus the infrequency of symptom report is tied to this particular population, but the large sample size does lend support to the relative infrequency of CII scores above 20 in this population. As Study 1 findings were taken from a non-treatment-seeking sample, it is important to examine the utility of the CII in a clinic sample. In addition, Study 1 was limited by the fact that the criterion of noncredible symptom report was defined by extreme elevations of scales E and F on the CAARS, the instrument from which the CII was developed. Of note, only 3 of the 12 CII items come from those two scales, and exclusively from F, which attenuates this concern to some degree, but inclusion of an independent measure of noncredible performance would strengthen conclusions about the relation of the CII to noncredible responding.

Thus, in Study 2, we gathered initial CII validity data from a clinic sample. In addition, we examined the relationship of the CII to noncredible self-report, as defined as in Study 1 (extreme elevations on CAARS subscales E and F), and additionally to an index of noncredible cognitive responding, which provides an external criterion for validity beyond the CAARS scale. Although noncredible cognitive responding, which is performance based, does not always correlate with noncredible symptom reporting, evidence that individuals are performing noncredibly on both self-report scales and in cognitive performance might be potentially useful when considering the cause of the noncredible behavior (i.e., a "cry for help" vs. malingering).

Study 2: Initial Validation of the CII
Method

Participants. Participants were 124 individuals who presented for psychological evaluation at a large Midwestern university's psychology clinic and who provided consent for data to be used for archival research purposes. They included a subset of participants from a previous study (Suhr et al., 2008) for whom complete data with endorsement of all CAARS items could be located in the clinic archives, plus 58 additional new archived clinical records for participants who were coded for group placement using the same criteria described in the prior study. All archived data were taken from among consecutive referrals of adults 18 years and older with concerns about ADHD who had given permission for use of their deidentified evaluation results in archival research and who reported no history of neurological injury or illness during their clinical interview.

Participants completed demographic and self-report symptom questionnaires as well as psychological and neuropsychological tests as part of a comprehensive assessment battery for academic concerns. Participants ranged in age from 18 to 59 years (average age = 22 years), and on average were in their second year of undergraduate education (range = 12-20 years of education). There were 66 male participants. Participants' self-reported racial/ethnic status was predominantly Caucasian (93%; 4% African American, 2% Hispanic/Latino, 1% Indian). Participants were divided into three groups based on records review. Details about group placement appear in the previous study (Suhr et al., 2008) but are briefly summarized here.

There were 29 individuals (12 male) who failed at least one of the first four subtests of the Word Memory Test (WMT) using standard clinical norms; they were assigned to the *Noncredible Performance group*.

There were 19 individuals (11 male) who were diagnosed with ADHD based on their evaluation results and using criteria described in Suhr et al. (2008). In brief, all individuals assigned to the *ADHD Diagnosis group* had to have at least two pieces of evidence documenting childhood impairment related to ADHD symptoms, evidence clinically significant ADHD symptoms currently

based on at least two of the following: behavioral observation, collateral report, or self-report, and pass the WMT. Individuals who were taking medication for ADHD at the time of their evaluation ($n = 6$) were instructed to complete self-report ADHD items based on their typical behavior rather than on their behavior while on medication.

There were 43 individuals (22 male) in the *Psychological Symptom group*. This included individuals with no evidence of impairment in childhood related to ADHD complaints, as described above; in addition, they had diagnosis of/treatment for a non-ADHD psychological condition independent of the current clinical evaluation (i.e., from another health care provider or another clinic) *or* in some cases, had not yet been diagnosed but met diagnostic criteria for a psychological disorder, based on clinical interview and psychological testing, usually major depressive disorder, at the time of their evaluation.

Finally, there were 33 individuals (21 male) who did not meet criteria for placement in the other three groups, but were assigned to the *Other group* as yet another specificity check for the CII. This group included 12 individuals who were diagnosed with learning disability but not ADHD during their evaluation and 21 individuals who did not meet diagnostic criteria for any psychological or neurodevelopmental disorder.

Measures. Although each clinical participant completed a long battery of psychological and neuropsychological tests as part of their clinical evaluation and which were related to their placement in diagnostic groups as described above, only measures of interest for the present analyses are described below.

The CAARS Self-Report–Long Form (Conners et al., 1998) is a 66-item inventory of self-reported ADHD symptoms; its psychometric properties were described above.

The WMT (Green, 2005) is a behavioral measure of noncredible responding. The WMT immediate recognition has high diagnostic accuracy as a measure of noncredible cognitive performance in a variety of clinical populations, including psychiatric patients, disability claimants, and traumatic brain injury, with high concordance rates compared with other measures of noncredible cognitive responding (Bauer, O'Bryant, Lynch, McCaffrey, & Fisher, 2007; Green, 2005). In addition, the WMT has shown excellent ability to detect those simulating brain damage (Tan et al., 2002). Split-half reliability of the WMT is high ($r = .86$-$.90$; Green, 2005).

Statistical analyses. Given the gender differences suggested in Study 1, we first assessed for gender differences in the clinic data (collapsing across all groups).

To examine the relationship of CII to other CAARS indices of noncredible self-report, as in Study 1, we used the entire sample collapsed across all groups. To understand CII's relationship to overreporting of symptoms, we examined its relation to inattentive (Scale E) and hyperactive/impulsive (Scale F) symptom report on the CAARS, using T scores of 80 or higher on E and F as criteria for noncredible symptom report. We thus divided the entire sample into two groups (noncredible on E, credible on E) and conducted a receiver operating characteristic (ROC) analysis, a graphical plot of sensitivity versus ($1 -$ specificity) at each level of the CII for predicting the criterion (noncredible symptom report). ROC analysis allows for determination of an overall accuracy of classification, as well as classification statistics to address specific goals (high sensitivity, high specificity). We repeated this analysis using F as the criterion for credible report (i.e., dividing the full sample into noncredible on F, credible on F).

The final set of statistical comparisons was conducted within and between study groups, rather than collapsing across the full sample. The first set examined the accuracy rates of detection using E, F, and CII within each group. Finally, we used independent criterion for validity (noncredible cognitive performance) by assessing the accuracy of the CII in detecting members of the group who failed the WMT (Noncredible Performance group) but not members of the ADHD Diagnosis group.

Table 1. Percentage of Individuals in Each Study Group Who Performed Above Cutoffs for Noncredible Performance

Group	Above CII Cutoff	Above E Cutoff	Above F Cutoff
Noncredible	24	72	24
ADHD	5 (*n* = 1)	42	5 (*n* = 1)
Psychological Group	23	60	12
Other	6	36	6

Note. ADHD = attention deficit hyperactivity disorder, CII = Conner's Infrequency Index, E = subscale E on the Conner's Adult ADHD Rating Scale, F = subscale F on the Conner's Adult ADHD Rating Scale.

Results

There was not a significant difference between males (mean = 14.8, *SD* = 6.2,) and females (mean = 13.3, *SD* = 6.3) on the CII, *t*(122) = 1.39, *p* = .17. Therefore, all subsequent analyses collapsed across gender.

Total sample analyses. In the entire sample, the CII cutoff of 20 identified 1 of 6 people scoring in the suspect range on the Inconsistency Index (with 87% specificity).

Using *T* scores of 80 or higher on E as a criterion for noncredible symptom report, we divided the entire sample into two groups; credible symptom reporters (*n* = 57) and noncredible symptom reporters (*n* = 67). We then conducted an ROC analysis of the CII. Area under the curve (AUC) graphs showed that the CII demonstrated moderate accuracy (78%) in identifying group membership, *SE* = .04, *p* < .001. Using the curve to select a cutoff with 100% specificity (i.e., no false identification of anyone in the credible symptom group) suggested use of a cutoff of 20 on the CII, which also showed 30% sensitivity to noncredible symptom report.

This analysis was repeated using *T* scores of 80 or higher on F as a criterion for noncredible symptom report, resulting in 109 people in the credible symptom group and 15 people in the noncredible symptom group. The ROC analysis and AUC graphs showed that the CII demonstrated good accuracy (92%) in identifying group membership, *SE* = .03, *p* < .001. The suggested cutoff of 20 on the CII identified 80% of noncredible *F* scores, with 93% specificity.

Study group analyses. Table 1 shows the percentage of individuals in each study group who scored in the noncredible range on each of the E, F, and CII scales. Of note, 100% of those who were noncredible on the CII were also noncredible on either E or F. Among those who were noncredible on either E or F, 32% of the Noncredible group, 13% of the ADHD group (*n* = 1), 38% of the Psychological Control group, and 15% of the Control group were also noncredible on the CII.

To further examine CII's relation to noncredible responding, in this case on cognitive tasks, we examined the accuracy of the CII in detecting the Noncredible Performance group (*n* = 29) relative to the ADHD group (*n* = 19). The AUC graph showed that the CII was at least moderately accurate, with 67% overall accuracy, *SE* = .08, *p* < .05. Using a cutoff of 20, the CII identified 24% of the Noncredible Performance group, with 95% specificity (*n* = 1 of the ADHD group scored above 20; this individual had a comorbid diagnosis of major depressive disorder and was also identified by noncredible scores on both CAARS subscales E and F). Neither CAARS Scale E nor CAARS Scale F showed better accuracy for distinguishing the Noncredible Performance group from the ADHD group. The ROC analyses for CAARS Scale E showed an overall 58% accuracy, *SE* = .09, *p* = .33, with a cutoff of 80 or higher showing 72% sensitivity but only 68% specificity. The ROC analyses for CAARS Scale F showed an overall 58% accuracy, *SE* = .09, *p* = .37; however, a cutoff of 80 on Scale F demonstrated the same sensitivity and specificity as the CII, in fact identifying the same individuals in both groups.

Study 2 Discussion

Study 2 provided initial support for the CII as a measure of noncredible self-report in ADHD assessment. As in Study 1, the CII was not strongly related to inconsistency in self-report, and thus the CII may serve as an independent way in which to check the validity of self-report on the CAARS. In addition, as in Study 1, the CII showed moderate sensitivity to extreme scores on subscales E and F of the CAARS, with good specificity. Within-group analyses showed that the CII identifies only a subset of those extremely high on E and F. As the CAARS manual points out, extreme scores on CAARS subscales are suspect for invalidity but could also indicate severe ADHD symptomatology; in identifying only a subset of those who are extreme on CAARS subscales, the CII may be showing a stronger relationship to noncredible self-report.

There were a relatively high number of individuals in the Psychological group who scored high on the CII, which would be expected, given that an overreporting response bias can also occur in those who are acutely distressed at the time of evaluation (i.e., the "cry for help;" Butcher et al., 2001). Our data are consistent with this, in that 60% of the Psychological Controls in Study 2 scored noncredibly high on CAARS inattentive symptoms (Scale E), and 23% of them scored noncredibly high on CII. These findings further illustrate that face valid "ADHD" symptoms are commonly reported in treatment-seeking populations and do not necessarily distinguish ADHD from other psychological conditions.

Finally, Study 2 showed that CII is relatively sensitive and quite specific to an external criterion of validity, noncredible cognitive performance (i.e., failure of the WMT). The CII more accurately distinguished the Noncredible Performance group from the ADHD group than either the E or F scales. Although neither noncredible symptom report nor noncredible cognitive performance in and of themselves is diagnostic of malingering, the combination of noncredible symptom report and noncredible performance on cognitive measures, in concert with evidence for the potential for secondary gain, and lack of other clinical evidence for ADHD, should make a clinician more suspicious of malingering as a possible explanation for such a clinical presentation.

The present results emphasize the need for existing self-report measures such as the CAARS to include indices of noncredible responding that are based on infrequently endorsed items, such as those in the CII, and will hopefully inform the revision of existing or development of new self-report ADHD scales. The CII items were developed using the grand tradition of existing measures of psychopathology that include infrequency scales, such as the MMPI (Ben-Porath & Tellegen, 2008; Butcher et al., 2001) and the PAI (Morey, 2007). Ideally, however, infrequency measures would include items specifically developed to be infrequently endorsed rather than statistically based as in the present study. Currently, work to develop new infrequently endorsed "ADHD" items to embed within self-report instruments such as the CAARS is underway (A. Harrison, personal communication, December 18, 2009), and it is also hoped that this will inform the next version of self-reported ADHD symptom scales.

In summary, these preliminary results on the CII suggest that it is a potentially useful indicator of an overreporting bias on the CAARS. However, the score of 21 or higher on the CII needs independent replication in other clinical samples, particularly given the limited number of well-documented adult ADHD participants in Study 2 and the limited age and education range of the participants in both studies. For example, it is possible that adults in midlife or not in college would endorse the infrequent items with more frequency. Furthermore, given that the students with ADHD in the present sample were attending university (although often not successfully), the present study did not likely include individuals with severely impairing ADHD symptomatology. Thus, the CII at present is most appropriate for evaluations for ADHD in a university setting and should not be interpreted in isolation from other clinical data. In addition, given the

potential gender differences in the scale identified in Study 1, further research should continue to examine gender differences on the CII that would influence the accuracy of our initial cutoffs for the scale. Additional validity studies should be conducted with the CII, including use of simulated ADHD samples and comparison with other self-report measures that include assessment noncredible self-report (e.g., the MMPI).

Should continued examination of the CII confirm these preliminary findings, this scale could provide valuable information for clinicians who wish to ensure that they are not invalidly diagnosing individuals with a neurodevelopmental condition such as ADHD, which might lead to inappropriate treatment and/or lack of treatment for another more appropriate diagnosis.

Declaration of Conflicting Interests

Funding

References

Alfano, K., & Boone, B. K. (2007). The use of effort tests in the context of actual versus feigned attention-deficit/hyperactivity disorder and learning disability. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 366-383). New York, NY: Guilford Press.

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

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. doi:10.1177/1073191106297617

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory* (2nd ed.). San Antonio, TX: Psychological Corporation.

Ben-Porath, Y. S., & Tellegen, A. (2008). *Minnesota Multiphasic Personality Inventory-2 Restructured Form: Manual for administration, scoring, and interpretation*. Minneapolis: University of Minnesota Press.

Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). *Minnesota Multiphasic Personality Inventory–2: Manual for administration and scoring* (rev. ed.). Minneapolis: University of Minnesota Press.

Conners, C. K., Erhardt, D., & Sparrow, E. P. (1998). *Conners Adult Attention Rating Scale–Self-Report: Long Version*. North Tonawanda, NY: Multi-Health Systems.

Conti, R. (2004). Malingered ADHD in adolescents diagnosed with conduct disorder: A brief note. *Psychological Reports, 94,* 987-988. doi:10.2466/PR0.94.3.987-988

DuPaul, G. J., Schaughency, E. A., Weyandt, L. L., Tripp, G., Kiesner, J., Ota, K., & Stanish, H. (2001). Self-report of ADHD symptoms in university students: Cross-gender and cross-national prevalence. *Journal of Learning Disabilities, 34,* 370-379. doi:10.1177/002221940103400412

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. doi:10.1177/108705479900300304

Green, P. (2005). *Green's Word Memory Test for Windows: User's manual*. Edmonton, Alberta, Canada: Author.

Harrison, A. G. (2004). An investigation of reported symptoms of ADHD in a university population. *The ADHD Report, 12*(6), 8-11. doi:10.1521/adhd.12.6.8.55256

Harrison, A. G., Edwards, M. J., & Parker, K. C. H. (2007). Identifying students faking ADHD: Preliminary findings and strategies for detection. *Archives of Clinical Neuropsychology, 22,* 577-588. doi:10.1016/j.acn.2007.03.008

Heiligenstein, E., Conyers, L. M., Berns, A. R., Miller, M. A., & Smith, M. A. (1998). Preliminary normative data on DSM-IV attention deficit hyperactivity disorder in college students. *Journal of American College Health, 46,* 185-188. doi:10.1080/07448489809595609

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.

Lee Booksh, R., 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. doi:10.1177/1087054708329927

Mannuzza, S., Klein, R. G., Klein, D. F., Bessler, A., & Shrout, P. (2002). Accuracy of adult recall of childhood attention deficit hyperactivity disorder. *American Journal of Psychiatry, 159,* 1882-1888. doi:10.1176/appi.ajp.159.11.1882

McCann, B. S., & Roy-Byrne, P. (2004). Screening and diagnostic utility of self-report attention deficit hyperactivity disorder scales in adults. *Comprehensive Psychiatry, 45,* 175-183. doi:10.1016/j.comppsych.2004.02.006

Morey, L. C. (2007). *The Personality Assessment Inventory (PAI): Professional manual* (2nd ed.). Lutz, FL: Psychological Assessment Resources.

Murphy, K. (1994). Guarding against overdiagnosis of ADHD in adults. *The ADHD Report, 2*(6), 3-4.

Murphy, K., & Barkley, R. A. (1996). Attention deficit hyperactivity disorder adults: Comorbidities and adaptive impairments. *Comprehensive Psychiatry, 37,* 393-401. doi:10.1016/S0010-440X(96)90022-X

Murphy, K., Gordon, M., & Barkley, R. (2000). To what extent are ADHD symptoms common? A reanalysis of standardization data from a DSM-IV checklist. *The ADHD Report, 8*(3), 1-5.

Quinn, C. (2003). Detection of malingering in assessment of adult ADHD. *Archives of Clinical Neuropsychology, 18,* 379-395.

Roy-Byrne, P., Sheele, L., Brinkley, J., Ward, N., Wiatrak, C., Russo, J., et al. (1997). Adult attention-deficit hyperactivity disorder: Assessment guidelines based on clinical presentation to a specialty clinic. *Comprehensive Psychiatry, 38,* 133-140. doi:10.1016/S0010-440X(97)90065-1

Suhr, J., Hammers, D., Dobbins-Buckland, K., Zimak, E., & Hughes, C. (2008). The relationship of malingering test failure to self-reported symptoms and neuropsychological findings in adults referred for ADHD evaluation. *Archives of Clinical Neuropsychology, 23,* 521-530. doi:10.1016/j.acn.2008.05.003

Suhr, J., Zimak, E., Buelow, M., & Fox, L. (2009). Self reported childhood Attention Deficit/Hyperactivity symptoms are not specific to the disorder. *Comprehensive Psychiatry, 50,* 269-275. doi:10.1016/j.comppsych.2008.08.008

Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assessments, *Applied Neuropsychology, 14,* 189-207.

Tan, J. E., Slick, D. J., Strauss, E., & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist, 16,* 495-505.

Weyandt, L., Linterman, I., & Rice, J. (1995). Reported prevalence of attentional difficulties in a general sample of college-students. *Journal of Psychopathology and Behavioral Assessment, 17,* 293-304. doi:10.1007/BF02229304

Widows, M. R., & Smith, G. P. (2005). *Structured Inventory of Malingered Symptomatology*. Odessa, FL: Psychological Assessment Resources.