OXFORD
UNIVERSITY PRESS

Archives of Clinical Neuropsychology 33 (2018) 832–844

Archives of CLINICAL NEUROPSYCHOLOGY

Utility of the Personality Assessment Inventory in detecting feigned Attention-Deficit/Hyperactivity Disorder (ADHD): The Feigned Adult ADHD index

Stephen L. Aita[1,*], Channing A. Sofko[2], Benjamin D. Hill[1], Mandi W. Musso[3], Anneliese C. Boettcher[4]

[1]*Psychology Department, University of South Alabama, Mobile, AL, USA*
[2]*Department of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University, Rhode Island Hospital, Providence, RI, USA*
[3]*Department of Graduate Medical Education, Our Lady of the Lake Medical Center, Baton Rouge, LA, USA*
[4]*Department of Psychiatry and Behavioral Sciences, University of Kansas School of Medicine, Wichita, KS, USA*

*Corresponding author at: Psychology Department, University of South Alabama, Mobile, AL, USA.
Tel.: +1-251-460-6371; fax: +1-251-460-6320. E-mail address: sla1521@jagmail.southalabama.edu (S.L. Aita)

Editorial Decision 6 October 2017; Accepted 31 October 2017

Abstract

Objective: The high potential for secondary gain among college students presenting for Attention-Deficit/Hyperactivity Disorder (ADHD) evaluations highlights the need for psychometrically sound embedded validity indicators. The purpose of this study was to develop new validity indicators specific to feigned ADHD for the Personality Assessment Inventory (PAI) and compare them to preexisting imbedded PAI validity measures.

Method: PAI scales that were theoretically related to feigned ADHD were evaluated. A binomial (ADHD simulators, $n = 138$, and genuine ADHD, $n = 142$) logistic regression was conducted with selected PAI scales and subscales. Classification rates were compared between the new and existing validity scales. A similar approach was used for item-level data in a second study in a subgroup of the original sample.

Results: The derived PAI scale-based logistic regression had a sensitivity of 54% and specificity of 92%. This algorithm accurately identified 97% of healthy controls as not feigning ADHD and correctly identified 98% of a no diagnosis group and 75% of a mood/anxiety disorders group. Classification accuracy of the new index was superior to the majority of existing PAI validity scales across groups. An item-level PAI algorithm had a sensitivity of 85% and specificity of 97% for identifying feigned ADHD.

Conclusions: New validity measures were compared to existing PAI validity indicators and performed better than many of them in this study. The algorithms developed in this study of ADHD simulators and genuine ADHD cases have adequate sensitivity and good specificity and appear to function differently than other PAI symptom validity scales.

Keywords: ADHD; Malingering/symptom validity testing; Assessment

Introduction

ADHD is a neurodevelopmental disorder characterized by persistent deficits in attention and/or hyperactivity/impulsivity that occurs in approximately 5% of children and 2.5% of adults across cultures (American Psychological Association; APA, 2013). Symptoms are numerous and include difficulty with organization, distractibility and unrelated thoughts, and intrusive behaviors. Clinicians faced with accurately diagnosing this condition encounter immense challenges including a lack of consistent clinical profiles for ADHD (see Boonstra, Oosterlaan, Sergeant, & Buitelaar, 2005; Hervey, Epstein & Curry, 2004; Schoechlin & Engel, 2005 for meta-analytic reviews) and neuropsychological measures have demonstrated poor discriminate validity of ADHD (i.e., Pettersson, Söderström, & Nilsson, 2015). In addition, high rates of ADHD symptoms are reported among college students (ranging from 2% to 8%; Weyandt & DuPaul, 2008), and the availability of ADHD symptomatology information on the internet (Conti, 2004), and the ease with which individuals without ADHD can easily be coached to simulate ADHD on symptom rating scales (e.g., Quinn, 2003) compounds the problem.

doi:10.1093/arclin/acx113 Advance Access publication on 23 November 2017

The academic and social incentives available to college students feigning Attention-Deficit/Hyperactivity Disorder (ADHD) can be both abundant and appealing. Specifically, along with an ADHD diagnosis comes the opportunity to receive testing accommodations (i.e., extra time on exams, recordings of lectures). While these accommodations may allow more free time for social activities, an ADHD diagnosis may also be appealing to college students due to the resulting ability to access stimulant medications (Alfano & Boone, 2007; Frazier, Frazier, Busch, Kerwood, & Demaree, 2008; Sollman, Ranseen, & Berry, 2010). Despite the many risks associated with stimulant misuse (e.g., psychosis, myocardial infarction, sudden death; see Lakhan & Kirchgesser, 2012 for a review), these substances are often recreationally misused by students with and without a diagnosis (Babcock & Byrne, 2000; Lakhan & Kirchgesser, 2012; McCabe, Knight, Teter, & Wechsler, 2005; Sharp & Rosén, 2007; Upadhyaya et al., 2005; White, Becker-Belase & Grace-Bishop, 2006). In fact, research has demonstrated that for students with ADHD, some misuse their medication to attain a "high" (Upadhyaya et al., 2005), while some students without a diagnosis endorse using stimulants to aid with focus and concentration (Herman et al., 2011). The prevailing notion that stimulant misuse is not objectively harmful may be contributing to its use and acceptance on college campuses (see Arria & DuPont, 2011 for a review and call for action).

Potentially due to the abundance of information about ADHD and the knowledge of available secondary gains, the high base rate of symptom exaggeration and amplification on college campuses is not surprising. In fact, Sullivan, May, and Galbally (2007) found evidence of symptom exaggeration as high as 47.6% when an individual presented solely for an ADHD evaluation. These rates were lower, at 24.5%, when individuals presented for an ADHD and learning disability evaluation. With high base rates of ADHD symptom reporting on college campuses, it is incumbent upon clinicians to accurately detect symptoms, and the exaggeration of symptoms, with a limited number of existing validated approaches. Because of the lack of specificity of neuropsychological tests, some have called for the diagnosis of ADHD to be made simply by clinical interview and self-report when low intellectual functioning and/or learning disabilities are not suspected (e.g., American Academy of Child and Adolescent Psychiatry, 2007; McGough & Barkley, 2004). However, relying on self-report instruments may be problematic due to the potential for secondary gain, comorbidity with mood disorders, poor retrospective recall (Mannuzza & Klein, 2000) and the lack of specificity in detecting feigned ADHD. In particular, while some studies suggest self-report instruments can be useful in differentiating between healthy, genuine respondents, and both feigning and ADHD groups (Booksh, 2005; Booksh, Pella, Singh, & Gouvier, 2010; Harrison, Edwards, & Parker, 2007), other research has demonstrated a lack of specificity in using these instruments to differentiate between genuine ADHD and feigned ADHD (Harrison, 2006; McCann & Roy-Byrne, 2004). Although the use of multiple cognitive stand-alone symptom validity tests (SVTs) during an ADHD evaluation to detect feigned symptoms is promising (e.g., Fuermaier et al., 2016; Jasinski et al., 2011; Sollman et al. 2010), the use of existing embedded validity scales on commonly implemented self-report personality measures has been less promising (e.g., Fuermaier et al., 2016).

Current literature (e.g., Harp, Jasinski, Shandera-Oschner, Mason & Berry, 2011; Musso & Gouvier, 2004; Musso, Hill, Barker, Pella, & Gouvier, 2014) has expanded upon the difficulties in detecting malingered ADHD symptoms with traditionally defined cutoffs on commonly used personality assessment measures such as the Personality Assessment Inventory (PAI; Morey, 1991) and the Minnesota Multiphasic Personality Inventory – 2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008). The PAI is a broad, objective, self-report measure of personality and emotional functioning that is widely implemented in forensic (Archer, Buffington-Vollum, Stredny, & Handel, 2006) and clinical evaluations. While the PAI has multiple broad validity scales, including Positive Impression Management (PIM), Negative Impression Management (NIM), and more specific scales, including the Malingering Index (MAL) and the Rogers Discriminant Function (RDF), standard cutoffs have not been useful in detecting feigned ADHD (e.g., Musso et al., 2014). In fact, Musso and colleagues (2014) found that among college students asked to simulate ADHD, few students scored above the established cutoffs on PAI validity scales. Alternative cutoffs were recommended based on this study; however, sensitivity to feigned ADHD remained in the 0.20–0.33 range. The MMPI-2-RF is a restructured form of the MMPI-2 (Butcher et al., 2001), which is one of the most frequently administered instruments within psychological evaluations (Rabin, Barr & Burton, 2005). Similar to those results found with the PAI (i.e., Musso et al., 2014), Harp and colleagues (2011) encountered similar psychometric issues with established validity indicators on the MMPI-2-RF. Their study found that the MMPI-2-RF scales designed to detect "faking bad" had very low sensitivity in detecting feigned ADHD. It was only when alternative cut scores were implemented that select faking bad scales (e.g., *Fp-r*) demonstrated modest sensitivity with feigned ADHD.

The high base rate of college students exaggerating symptoms of ADHD, in conjunction with the low sensitivity demonstrated by the embedded measures of validity on the PAI (e.g., Musso et al., 2014), highlights a need to develop additional embedded validity measures with improved efficacy. Rios and Morey (2013) examined the adolescent version of the PAI in college students asked to feign symptoms and found that only the un-coached students had higher scores on the MAL compared with the clinical group. Musso and colleagues (2014) found that the MAL cut-off score of ≥3 resulted in excellent specificity, but a sensitivity of 0.22. When Musso and colleagues lowered the cut score to ≥2, the specificity was unacceptable for

most of the groups. Due to psychometric problems with existing validity scales within this population, the purpose of this study was to develop a novel embedded measure with improved psychometric properties for detecting and differentiating between active attempts of ADHD symptom amplification (sensitivity) and genuine psychological conditions or intact functioning (specificity). Drawing on the existing literature, we hypothesized that an embedded validity indicator, developed based on data from ADHD simulators, and those with genuine ADHD, would address previously encountered issues of specificity. Moreover, we hypothesized that the resulting regression equation would allow for good sensitivity and specificity in "no-diagnosis" control groups, and slightly poorer psychometric properties in a group with mood and anxiety disorders. We hypothesized that, when compared to existing embedded measures for the PAI, this new scale would have superior psychometric properties. We addressed these hypotheses in two studies examining the issue of developing an algorithm to detect feigned ADHD in two different ways: a theory-driven approach using PAI scales that focused on content and a more data-driven approach using item-level data. This has the added advantage of giving clinicians and researchers two different ways to apply the results of these studies depending on the data they have at hand (i.e., scales or actual items).

Study 1: Methods

Setting and Participants

The clinical group's data were obtained from archival data review of psychoeducational evaluations, which included the administration of the PAI as part of the test battery. All evaluations were conducted by doctoral graduate students in a clinical psychology program. Evaluations included a thorough clinical interview and all diagnoses were made under the supervision of a licensed psychologist. Records were excluded if the individual's measured full-scale IQ was less than 70, and/or if they had a diagnosis of an intellectual disability, and/or if they were diagnosed with a serious mental illness, such as schizophrenia. This study included individuals from two university-affiliated psychology training clinics within the southeastern United States. In addition, students were prospectively recruited from three southeastern universities to serve as "non-treatment seeking" controls or ADHD simulators. They received course credit for participation. Individuals who received course credit were recruited through classes and did not present to the clinic for evaluations, thus this group was considered the Non-Treatment Seeking Control group. Simulation procedures for the ADHD Simulator group can be found in a previously published article (Musso et al., 2014); however, these procedures will be briefly reviewed. Only those with no reported history of psychopathology, learning disorder, or ADHD were included in the study. Those randomly assigned to the simulation group were provided with a scenario. The scenario was that the students' roommate has been diagnosed with ADHD and is now taking medication and doing well in school, while also having time to socialize. The student decides to take the roommate's medication during midterms and then notices how much easier things were. This leads the student to believe that he/she has undiagnosed ADHD and this leads to "Googling" information about ADHD. The student simulator was then provided time to review ADHD symptoms that result from a simple "Google" search, and they were told that they could take some notes about symptoms from these pseudowebpages that would help them fake on tests. These simulators were then instructed to take the tests as if they were trying to convince someone that they have ADHD. Importantly, these students were reminded that they were trying to fake ADHD as a college student, so they must do at least as well as someone who is enrolled in a university. Moreover, they were instructed that they want to get diagnosed without over-exaggerating the part so that they are not detected as faking. The battery for the student simulators included the Shipley Institute of Living Scale, Second Edition (Shipley-II; Shipley, 2004), the PAI, the State-Trait Anxiety Inventory, Form Y (Spielberger, Gorsuch, & Vagg, 1970), the Wender Utah Rating Scale (Ward, Wender, & Reimherr, 1993), the Wechsler Adult Intelligence Scale, Third Edition, and Reliable Digit Span (RDS; Greiffenstein, Baker, & Gola, 1994), and a Participant Effort-Rating Scale which asked ADHD simulators to report on a Likert-type Scale of 1–5 (1 = not at all; 3 = somewhat; 5 = very much so) how much effort they put forth in performing as someone with ADHD, and how successful they thought they were at performing at that level without being detected. These simulators are analogs for the typical college student that would present feigning ADHD (there is no expectation that the typical student feigning ADHD does so correctly); this sample of simulators has been accepted for use in previous research (Musso et al., 2014).

Participants in the clinical group completed the PAI as part of a larger neuropsychological test battery. Effort was measured using the Wechsler Adult Intelligence Scale – Third Edition (WAIS-III; Wechsler, 1997a) embedded validity indices: Mittenberg Index (Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995), Reliable Digit Span (Greiffenstein et al., 1994), and Vocabulary-Digit Span (Mittenberg et al., 1995). Wechsler Memory Scale – Third Edition (WMS-III; Wechsler, 1997b) embedded indices were also used and included Logical Memory Rarely Missed Items Index (Kilgore & DellaPietra, 2000), Auditory Recognition Delayed raw score (Langeluddecke & Lucas, 2003), and Faces I total score (Glassmire et al., 2003;

Langeluddecke & Lucas, 2003). For the clinical sample, individuals diagnosed with ADHD who failed two or more validity indices (PAI indices and Reliable Digit Span) were classified as putting forth suspect effort (6.8% of the ADHD group in the clinical sample) and were removed from further analysis in order to have a clean "good effort" ADHD group for later comparisons to ADHD simulators. Overall, there were 142 individuals in the "ADHD" group, 137 in the "Mood/Anxiety Disorder" group, 111 in the "No Diagnosis/Clinical" controls group, 138 in the "ADHD Simulators" group, and 133 in the "Non-treatment Seeking" control group (see Table 1 for demographic information). Chi-square analyses revealed that the groups differed significantly by gender and ethnicity. ANOVA analyses revealed the groups were significantly different by age and education. This is not unexpected as these were largely clinical samples with the exception of the simulators but this did require evaluation of the effect of any of these differences on results obtained.

Measures

Personality Assessment Inventory (PAI). The PAI is a self-report personality measure comprised of 344 items on a four-point scale with anchor points of *False* and *Very True*. These items are categorized into four scales that assess validity of responding, 11 clinical syndrome scales with nine having associated subscales, five treatment scales, and two interpersonal scales. Items were developed with an emphasis on content validity with a focus on assessing the breadth of experiences within the disorders. Reported median Cronbach's alpha's (Morey, 1991) were .81 in a census-matched normative sample (N = 1,000). All protocols were entered into a computerized scoring program using information from the census-matched normative sample.

PAI Validity and Clinical Syndrome Scales Relevant to ADHD. A series of exploratory analyses examining correlations and mean differences were conducted using PAI scales that indexed cognitive and behavioral symptoms a layperson may consider to be associated with ADHD as well as to rule-out multicollinearity between scales. Based on theoretical considerations of what symptoms and behaviors would likely be endorsed by individual's feigning ADHD, the following PAI scales were considered strong candidates for evaluation as potential predictors of feigned ADHD for inclusion in a logistic regression equation. These scales and the reason for their inclusion are discussed subsequently.

Positive Impression Management (PIM). This validity index embedded within the PAI was designed to detect attempts to present oneself in a favorable manner. Interestingly, as noted by Morey (1991), McCrae and Costa (1983) found that the tendency toward socially desirable responding is related to other variables that are widely regarded to be signs of appropriate adjustment. These traits include extraversion, adjustment (particularly being considered "well-adjusted"), and openness to experience. The Cronbach's alpha for PIM in a census-matched sample was reported at 0.71 (Morey, 1991), which reflects this notion of the PIM measuring multiple constructs. PIM and the related indicator of Negative Impression Management (NIM) are considered to be failed when individuals score excessively high on these scales. However, low PIM scores are not

Table 1. Demographic information for Study 1 groups

Group	ADHD (n = 142)	ADHD Simulators (n = 138)	Mood/Anxiety Disorder (n = 137)	Clinic No Diagnosis (n = 111)	Healthy Controls (n = 133)
Age *M* (*SD*)	21.77 (3.99)[b,e]	19.83 (1.54)[a,c,d]	22.71 (4.58)[b,e]	22.05 (5.07)[b,e]	19.18 (1.57)[a,c,d]
Education *M* (*SD*)	13.8 (1.61)[b,e]	13.12 (1.08)[a,c,d]	14.25 (1.85)[b,e]	13.79 (2.06)[b,e]	13.17 (0.94)[a,c,d]
Sex					
% Male	54.9[b,c]	26.1[a,d]	35.0[a,d]	57.7[b,c,e]	30.8[a,d]
% Female	45.1	73.9	65.0	42.3	69.2
Racial composition					
% White	82.7[e]	72.5[e]	87.6[e]	84.7[e]	64.7[a,b,c,d]
% African American	10.1	10.1	8.8	10.8	24.1
% Asian	1.4	5.8	2.2	1.8	3.8
% Hispanic	5.8	7.2	1.5	1.8	3.8
% Other	—	4.3	—	0.9	3.8

Note: Means are presented with standard deviations in parentheses.
[a]Significantly different from ADHD, $p < .05$.
[b]Significantly different from ADHD Simulators, $p < .05$.
[c]Significantly different from Mood/Anxiety Disorder, $p < .05$.
[d]Significantly different from Clinic No Diagnosis, $p < .05$.
[e]Significantly different from Healthy Controls, $p < .05$.

traditionally interpreted as indicating possible profile distortion. It was hypothesized that individual's feigning ADHD would not only portray themselves in a negative manner, which would be detected by the NIM scale, but may also endorse a low level of positive attributes or characteristics resulting in a low PIM score. A low PIM score is not interpreted as a failed validity indicator in current PAI scoring methodology making PIM a promising candidate for inclusion in a logistic regression.

Schizophrenia Thought Disorder (SCZ-T). Thought disorder, characterized by improper use of both semantic and relational aspects of language, is considered to be a fundamental symptom in schizophrenia and is associated with loose associations between topics, tangential speech with topic changes, circumstantial speech with excessive detail, and illogical associations (Goldberg et al., 1998). The Thought Disorder (*SCZ-T*) scale focuses on the identification of confusion, distractibility, and concentration issues (Morey, 1996) and this makes this scale an intuitive candidate for a predictor of feigned ADHD symptomology. Morey (1996) noted that moderate scores (*T*-score of 60–69) on the SCZ-T are indicative of decision-making difficulties and problems with concentration often characteristic of individuals with anxiety and depression. He elaborated, stating that scores approaching a *T*-score of 70 are characterized by greater confusion in addition to the subtle difficulties. *T*-scores at or above 70 begin to characterize individuals with loose associations and difficulties with expression and communication. Morey asserts that when this subscale is elevated in the absence of an elevation on the SCZ clinical syndrome scale, the presence of other disorders, including attentional ones, should be further explored. Similar to the internal consistency of scores on the PIM, in a census-matched sample, the Cronbach's alpha for the SCZ-T subscale was reported at 0.73 (Morey, 1991).

Cognitive Depression (DEP-C). The Cognitive subscale within the Depression scale of the PAI is designed to detect aspects of depression related to cognition. Specifically, high scores on this scale suggest that an individual has difficulty concentrating and others may view this person as being indecisive. This focus on cognitive symptoms also makes this scale a strong candidate for a predictor of feigned ADHD symptoms. In a census-matched sample, the Cronbach's alpha for the DEP-C subscale was reported at 0.74 (Morey, 1991).

Stimulus Seeking (ANT-S). The Stimulus Seeking subscale within the Antisocial Features scale of the PAI was developed to detect individuals who have a proclivity towards risk-taking behaviors. Those with high scores tend to act impulsively to draw excitement out of boredom. It was hypothesized that an individual feigning ADHD may misinterpret these items as measuring the hyperactive behaviors associated with ADHD by the general population. The Cronbach's alpha for ANT-S in a census-matched sample was reported at 0.69 (Morey, 1991).

Development and validation. As described earlier, a group of candidate predictor PAI scales was identified. A series of exploratory ANOVAs were conducted to analyze mean differences among the diagnostic groups on the subscales of the PAI that were selected theoretically based on ADHD diagnostic criteria and from research examining effort and ADHD. Overall results indicated statistically significant effects for the following theoretically selected variables: PIM (F (4,656) = 46.65, $p <$ .001), SCZ-T (F (4,656) = 68.94, $p <$ .001), ANT-S (F (4,656) = 44.57, $p <$ .001), and DEP-C (F (4,656) = 62.33, $p <$ .001). PIM was selected because there were no failures on this scale for the simulators (meaning simulators did not score high on PIM above the traditional cut-off and were not identified as likely invalid as the scale is currently utilized) but they did score significantly lower on PIM than the genuine ADHD group. We intentionally did not choose to include NIM as individuals who failed NIM were removed prior to data analysis for the genuine ADHD group and this scale is already commonly utilized to assess symptom validity and would not be a new contribution to PAI validity indexes. The aforementioned variables (i.e., PIM, SCZ-T, ANT-S, and DEP-C) were not highly correlated with one another and were entered into a binary logistic regression function to determine their utility in predicting genuine ADHD from ADHD simulators. An algorithm was derived based on a logistic regression equation to predict a known genuine ADHD group from a feigned ADHD group. This scale-level algorithm became the basis of a new scale to detect feigned ADHD that was then applied to known diagnostic groups to determine the potential false positive rate. Given differences in demographic variables across groups, we also evaluated any possible effect of demographic variables on outcomes. We found that age and education were not significantly correlated with the outcome of the scale-level algorithm. For gender, there was a general trend for males to score higher on the scale-level algorithm than females but this was only a significant difference for the ADHD [M (SD): male = 0.39 (.23) and female = .30 (.18)] and Mood/Anxiety [M (SD): male = 0.53 (.24) and female = .41 (.25)] groups. For ethnicity, the only significant difference on the scale-level algorithm was for the Mood/Anxiety group but as illustrated in Table 1 the number of minority participants was very small and any differences cannot be interpreted reliably.

Finally, this new scale was compared to existing embedded PAI scales including the Malingering Index (MAL; Morey, 1996), Rogers Discriminant Function index (RDF; Rogers, Sewell, Morey, & Ustad, 1996), the Multiscale Feigning Index

(MFI; Gaines, Giles, & Morgan, 2012), and the Negative Distortion Scale (NDS; Mogge, Lepage, Bell, & Ragatz, 2010) in order to compare the applicability of this new scale compared to those that are currently used in practice. The MAL is derived from eight configural features based on differences between scales related to negative impression management, infrequent symptoms, inconsistent responding, paranoia, mania, depression, treatment resistance and antisocial features that are not commonly seen in actual clinical patients. The RDF is a discriminant function that was developed to differentiate genuine patients from individuals simulating depression, pain/somatization, or psychosis. The MFI is based on the average elevation of all PAI clinical scales and essentially indexes extreme global endorsement of psychopathology. The NDS is comprised of items associated with extreme psychopathology but that occur vary rarely in psychiatric populations. While all of these indexes have support as symptom validity measures, they all focus on aspects of psychopathology that would not commonly be endorsed by an individual feigning ADHD and demonstrate the need for the development of a symptom validity measure with content specific to ADHD.

Results

Development of a Regression Equation

All of the following results were obtained using census-normed *T*-scores provided by the PAI computerized scoring program. A logistic regression analysis was performed with the feigned ADHD group as the dependent variable and PIM, SCZ-T, ANT-S, and DEP-C as the predictor variables. The classification cut-value was set to .70 in order to reduce false-positives. This resulted in a significant model (χ^2 (4, 280) = 90.00, $p < .001$) with Nagelkerke $R^2 = .37$. See Table 2 for the results of the regression analysis. This binomial logistic regression resulted in the following equation that can be used in calculating the probability of feigning ADHD: $e^{[-3.270\text{-}.041(\text{PAI\_PIM } T \text{ score}) - .026(\text{SCZ\_T } T \text{ score}) + .037(\text{ANT\_S } T \text{ Score}) + .070(\text{DEP\_C } T \text{ Score})]} / 1 + e^{[-3.270\text{-}.041(\text{PAI\_PIM } T \text{ score}) - .026(\text{SCZ\_T } T \text{ score}) + .037(\text{ANT\_S } T \text{ Score}) + .070(\text{DEP\_C } T \text{ Score})]}$. In this algorithm, e is the base of the natural logarithm. In the 280 individuals (142 diagnosed with ADHD and 138 simulators) this algorithm was derived from, 46.4% of simulators were correctly classified (sensitivity) and 92.3% of individuals diagnosed with ADHD were correctly classified (specificity) using a cut-value of .70.

Application of Regression Equation to Known Diagnostic Groups

When this regression equation was applied to 133 individuals in the non-treatment seeking control group using a cut-value of .70, 97% of individuals were accurately identified as genuine or not feigning ADHD and only 3% of controls were misclassified as feigning ADHD. In applying this equation to the no diagnosis/clinical control group, 98.2% of individuals were correctly identified as not feigning ADHD, while only 1.8% of individuals were misclassified. Finally, when the regression equation was applied to 137 individuals in the mood/anxiety disorder group, 75.2% were correctly identified as not feigning ADHD, and 24.8% were misclassified as feigning ADHD. In order to view this information in table format, see Table 3.

Comparison to Existing Embedded Measures

This newly developed measure, which will be referred to as the scale-level Feigned Adult ADHD (FAA) index, was compared to existing validity indicators in genuine ADHD, mood/anxiety disorders, and feigned ADHD groups. Specifically, the scale-level FAA was compared to the MAL, RDF, MFI, and the NDS that were described earlier. These last two scales are relatively new indexes that have not been extensively evaluated previously but seem promising. Refer to Table 4 for hits and

Table 2. Final variables in the ADHD versus ADHD simulators scale-level logistic regression model

					95% CI for EXP(*B*)	
PAI scales	*B*	*SE*	*p*	Exp(*B*)	Lower	Upper
PIM	−0.041	0.017	.013	.960	0.929	0.991
SCZ-T	−0.026	0.013	.035	.974	0.950	0.998
ANT-S	0.037	0.011	.000	1.038	1.016	1.060
DEP-C	0.070	0.014	.000	1.072	1.042	1.103
Constant	−3.270	1.582	.039	0.038	—	—

Note: Normed *T*-scores were used for all variables in the regression. Coding was 0 for genuine ADHD and 1 for ADHD simulators.

Table 3. Classification rates of the scale-level FAA index in known groups

	Correctly identified	Misidentified
Genuine ADHD Group	92.3%	7.7%
Mood/Anxiety Group	75.2%	24.8%
Healthy Controls	97.0%	3.0%
Clinic No Diagnosis	98.2%	1.8%
Feigned ADHD	46.4%	53.6%

Note: Cut-off value = .70.

Table 4. Classification rates of existing validity measures in the genuine ADHD group

	Hits	Misses
Genuine ADHD Group		
NDS	98.6%	1.4%
RDF	83.0%	17.0%
MAL	100%	0.0%
MFI	100%	0.0%
Scale-Level FAA	92.3%	7.7%
Mood and Anxiety Disorders		
NDS	94.2%	5.8%
RDF	71.9%	28.1%
MAL	100%	0.0%
MFI	97.7%	2.3%
Scale-Level FAA	75.2%	24.8%
Feigned ADHD		
NDS	28.3%	71.7%
RDF	57.7%	42.3%
MAL	0.0%	100%
MFI	12.6%	87.4%
Scale-Level FAA	46.4%	53.6%

Note: Hits = correctly identified individuals. Misses = incorrectly identified individuals. A cut-off of >12 was used for the NDS, a cut-off of >0 was used for RDF, a cut-off of >4 was used for MAL, and a cut-off of $T > 76$ was used for the MFI. For the FAA, a cut-off of ≥.70 was used.

misses of these indices in the genuine ADHD group, the mood/anxiety disorders group, and the feigned ADHD group. The NDS, MAL, and MFI had low rates of false positives but had low sensitivity for feigned ADHD. The RDF was the most sensitive to feigned ADHD but had the highest false positive rate (17.0%) as well. The scale-level FAA demonstrated better specificity to genuine ADHD than RDF and slightly better specificity in the mood and anxiety disorders group.

Study 2: Methods

While Study 1 reported an algorithm derived from *T*-scores from PAI scales that appears useful for detecting likely feigning of ADHD, it is possible that a similar logistic regression approach based on item-level data could provide superior predictive utility for identifying feigned ADHD. This is due to item-level data possibly having more specific content that an individual feigning ADHD would endorse that may be obscured when multiple items are combined into scale composites. Therefore, a secondary study was conducted to utilize a different approach to detect feigned ADHD. In contrast with the theory-driven, composite-level method, a data-driven, item-level approach was used to develop a regression-based logistic equation to identify feigned ADHD. In this additional study, 361 participants from the dataset used for Study 1 but that had item-level PAI data available were used in the selection of items for regression scale development. See Table 5 for demographics of this subsample. The reference groups were initially genuine ADHD versus ADHD simulators. However, this approach yielded good sensitivity, but poor specificity, which was likely due to the limited range of information that was overly specific to ADHD and feigned ADHD. Stated differently, the obtained regression was over-calibrated to only these two groups. Consequently, the reference groups were set to feigned ADHD versus valid performers (i.e., all other groups) to maximize information used for the equation loadings for development of the item-level logistic regression equation, referred to as the item-level FAA.

First, all PAI items were entered into a multiple regression using forward entry to predict ADHD simulator versus non-simulators. Next, of the PAI items retained in the multiple regression equation, tolerance and VIF statistics were examined to

Table 5. Demographic information for Study 2 groups

Group	ADHD (n = 73)	ADHD Simulators (n = 138)	Mood/Anxiety Disorder (n = 52)	Clinic No Diagnosis (n = 45)	Healthy Controls (n = 53)
Age *M* (*SD*)	22.33 (3.93)[b,e]	19.83 (1.54)[a,c,d]	21.98 (4.26)[b,e]	22.80 (5.13)[b,e]	19.45 (1.35)[a,c,d]
Education *M* (*SD*)	14.00 (1.66)[b,e]	13.12 (1.08)[a,c,d]	14.38 (2.23)[b,e]	14.10 (2.27)[b,e]	12.87 (1.01)[a,c,d]
Sex					
% Male	56.2[b,c,e]	26.1[a,d]	34.6[a,d]	62.2[b,c,e]	24.5[a,d]
% Female	43.8	73.9	65.4	37.8	75.5
Racial composition					
% White	75.3	72.5	92.3	84.4	79.2
% African American	9.6	10.1	5.8	8.9	9.4
% Asian	2.7	5.8	1.9	4.4	1.9
% Hispanic	9.6	7.2	—	–	5.7
% Other	—	4.3	—	2.2	3.8

Note: Means are presented with standard deviations in parentheses.
[a]Significantly different from ADHD, $p < .05$.
[b]Significantly different from ADHD Simulators, $p < .05$.
[c]Significantly different from Mood/Anxiety Disorder, $p < .05$.
[d]Significantly different from Clinic No Diagnosis, $p < .05$.
[e]Significantly different from Healthy Controls, $p < .05$.

screen retained items for potential multicollinearity. This procedure resulted in the retention of 44 PAI items for entry in binary logistic regression analysis. The 44 items were then entered into a binary logistic regression analysis to predict ADHD simulator versus valid effort group membership. Non-significant items in the model were removed through an iterative process, which resulted in 24 final PAI items.

Results

Development of an Item-level Regression Equation

All of the following results were obtained using data from the final regression-identified PAI items. The classification cut-value was set to .60 as this optimized sensitivity and specificity for the final 24 PAI items. This is a lower setting for the cut-value than the .70 that was used in Study 1 and represents the probability of identification of the dependent variable with a higher probability essentially meaning it is more difficult to be categorized in the predicted group. This resulted in a significant model (χ^2 (24, 351) = 129.98, $p < .001$) with Nagelkerke $R^2 = .84$. See Table 6 for the results of the regression analysis. The following equation can be used in calculating the probability of symptom exaggeration or amplification using item-level data: Probability of symptom exaggeration = $e^{[-7.632 + 1.445(\text{PAI Item 311}) + 1.621(\text{PAI Item 112}) + 1.157(\text{PAI Item 299}) - 1.473(\text{PAI Item 13}) + 1.384(\text{PAI Item 167}) + 1.628(\text{PAI Item 34}) - 1.287(\text{PAI Item 57}) + 1.418(\text{PAI Item 15}) + .971(\text{PAI Item 233}) + 1.041(\text{PAI Item 201}) - 1.673(\text{PAI Item 169}) + .940(\text{PAI Item 139}) + 1.186(\text{PAI Item 195}) - 1.222(\text{PAI Item 18}) - 2.547(\text{PAI Item 254}) - .811(\text{PAI Item 325}) + 1.058(\text{PAI Item 236}) - 1.312(\text{PAI Item 198}) - .843(\text{PAI Item 21}) - .894(\text{PAI Item 322}) + 1.083(\text{PAI Item 94}) + .720(\text{PAI Item 282}) + 1.089(\text{PAI Item 213}) + 1.408(\text{PAI Item 171})]} / 1 + e^{[-7.632 + 1.445(\text{PAI Item 311}) + 1.621(\text{PAI Item 112}) + 1.157(\text{PAI Item 299}) - 1.473(\text{PAI Item 13}) + 1.384(\text{PAI Item 167}) + 1.628(\text{PAI Item 34}) - 1.287(\text{PAI Item 57}) + 1.418(\text{PAI Item 15}) + .971(\text{PAI Item 233}) + 1.041(\text{PAI Item 201}) - 1.673(\text{PAI Item 169}) + .940(\text{PAI Item 139}) + 1.186(\text{PAI Item 195}) - 1.222(\text{PAI Item 18}) - 2.547(\text{PAI Item 254}) - .811(\text{PAI Item 325}) + 1.058(\text{PAI Item 236}) - 1.312(\text{PAI Item 198}) - .843(\text{PAI Item 21}) - .894(\text{PAI Item 322}) + 1.083(\text{PAI Item 94}) + .720(\text{PAI Item 282}) + 1.089(\text{PAI Item 213}) + 1.408(\text{PAI Item 171})]}$. As simulators versus valid effort was the set reference group, of the 351 individuals with complete data for analysis (223 with valid effort and 128 simulators) that the item-level FAA algorithm was derived from, 85.2% of simulators were correctly classified (sensitivity) and 97.0% of individuals who were non-simulators were correctly classified using a cut-value of .60. To determine specificity for "true" ADHD, the item-level FAA was applied to the participants with an ADHD diagnosis. Using a cut-value of .60, the regression equation correctly identified 94.5% of those with a diagnosis of ADHD (specificity).

Application of item-level regression equation to known diagnostic groups. Similar to the scale-level FAA, when the item-level FAA was applied to 53 individuals in the non-treatment seeking control group using a cut-value of .60, 96.2% of individuals were accurately identified as genuine or not feigning ADHD and only 3.8% of this group were misclassified as feigning ADHD. As applied to the no diagnosis/clinical controls group (n = 45), the equation correctly classified 100% of individuals were correctly identified as not feigning ADHD. Finally, when the item-level FAA was applied to 52 individuals in the mood/

Table 6. Final variables in the valid effort versus ADHD simulators item-level logistic regression model

					95% CI for EXP(*B*)	
PAI scales	*B*	*SE*	*p*	Exp(*B*)	Lower	Upper
PAIItem311	1.445	0.574	.012	4.242	1.376	13.077
PAIItem112	1.621	0.384	<.001	5.057	2.381	10.738
PAIItem299	1.157	0.295	<.001	3.179	1.784	5.667
PAIItem13	−1.473	0.322	<.001	0.229	0.122	0.431
PAIItem167	1.384	0.335	<.001	3.991	2.071	7.688
PAIItem34	1.628	0.377	<.001	5.091	2.432	10.661
PAIItem57	−1.287	0.372	.001	0.276	0.133	0.572
PAIItem15	1.418	0.346	<.001	4.129	2.095	8.136
PAIItem233	0.971	0.299	.001	2.641	1.469	4.749
PAIItem201	1.041	0.359	.004	2.831	1.402	5.718
PAIItem169	−1.673	0.454	<.001	0.188	0.077	0.456
PAIItem139	0.940	0.329	.004	2.561	1.345	4.878
PAIItem195	1.186	0.312	<.001	3.275	1.776	6.038
PAIItem18	−1.222	0.312	<.001	0.295	0.160	0.543
PAIItem254	−2.547	0.627	<.001	0.078	0.023	0.268
PAIItem325	−0.811	0.373	.030	0.445	0.214	0.924
PAIItem236	1.058	0.316	.001	2.881	1.552	5.349
PAIItem198	−1.312	0.309	<.001	0.269	0.147	0.493
PAIItem21	−0.843	0.381	.027	0.430	0.204	0.909
PAIItem322	−0.894	0.342	.009	0.409	0.209	0.800
PAIItem94	1.083	0.342	.002	2.955	1.510	5.780
PAIItem282	0.720	0.304	.018	2.054	1.133	3.725
PAIItem213	1.089	0.338	.001	2.972	1.533	5.763
PAIItem171	1.408	0.344	<.001	4.089	2.083	8.024
Constant	−7.632	1.600	<.001	0.000	—	—

Note: Raw PAI items were used for all variables in the regression. Coding was 0 for valid effort and 1 for ADHD simulators.

Table 7. Classification rates of the item-level FAA in known groups

	Correctly Identified	Misidentified
Genuine ADHD Group	97.3%	2.7%
Mood/Anxiety Group	96.2%	3.8%
Healthy Controls	96.2%	3.8%
Clinic No Diagnosis	100%	0%
Feigned ADHD	85.2%	14.8%

Note: Cut-off value = .60.

anxiety disorder group, 96.2% were correctly identified as not feigning ADHD, and 3.8% were misclassified as feigning ADHD (see Table 7).

Given differences in demographic variables across groups, we again evaluated any possible effect of demographic variables on outcomes. We found that age was not significantly correlated with the outcome of the item-level FAA algorithm but education was significantly associated with item-level FAA outcome for only the ADHD group ($r = -.24$). This was a small effect and means that education accounted for 6% of the variance in the outcome of the item-level FAA algorithm. For gender, there was a general trend for males to score slightly higher on the item-level FAA than females but there were no statistically significant differences between genders. For ethnicity, there were no statistically significant differences in item-level FAA for any groups.

General Discussion

The high potential for secondary gains in obtaining an ADHD diagnosis and the lack of psychometrically sound embedded tools for clinicians has presented a challenge for those conducting ADHD evaluations within a college setting (Pella, Hill, Shelton, Elliott, & Gouvier, 2012). Accurate detection of ADHD remains a public health priority given the prevalence and dangers of stimulant misuse (Babcock & Byrne, 2000; Lakhan & Kirchgesser, 2012; McCabe et al., 2005; Sharp & Rosén, 2007; Upadhyaya et al., 2005; White et al., 2006). As such, the purpose of the present study was to address extant

psychometric issues; namely, many existing self-report instruments lack adequate specificity in differentiating between genuine ADHD and simulated ADHD (Harrison, 2006; McCann & Roy-Byrne, 2004). The aim of this study was to address this problem by developing a regression equation to predict group membership (genuine ADHD versus simulated ADHD) using information theoretically selected from the PAI, and then applying this to known diagnostic groups. This resulted in the development of a new embedded index, the scale-level FAA, which can easily be applied to aid in credible versus non-credible performance decision-making.

The study hypotheses were largely supported and the regression equation demonstrated good sensitivity and specificity when applied to known clinical and non-clinical control groups. When compared to other existing embedded effort indices within the PAI (NDS, RDF, MAL, and MFI) the scale-level FAA performed similarly to the RDF. Notably, only the scale-level FAA and the RDF were effective at identifying individuals feigning ADHD. The other PAI validity indicators that were examined did a poor job at identifying ADHD simulators in this study with the MAL not identifying anyone as feigning ADHD when the standard cut-score of >4 was used. As predicted based on the internal consistency of the PAI subscales used in this study, the psychometric properties (i.e., sensitivity and specificity) were generally poorer for the PAI validity indicators in this study when applied to a known mood/anxiety disorder diagnostic group but this was particularly true for the scale-level FAA and RDF which had a similar false positive rate in the mood/anxiety diagnosis group. A strength of the scale-level FAA compared to the RDF is that the RDF had more than double the rate of false positives for individuals with genuine ADHD. Also, the scale-level FAA algorithm is easier to compute than the RDF as it requires only four scale/subscale scores compared to 20 scales/subscales for RDF. When the scale-level FAA algorithm results in a score of .70 or above there is a high probability that ADHD symptom exaggeration or amplification is present, particularly in the absence of a diagnosable mood or anxiety disorder. The current results support that the scale-level FAA can be used in conjunction with other PAI validity measures to aid in clinical decision making and it seems to provide unique information compared to other validity measures. Additionally, our results demonstrate that the NDS, MAL, and MFI have low rates of false positives in genuine cases of ADHD, mood, or anxiety disorders supporting the utility of these indexes as symptom validity measures.

The data-driven approach (i.e., item-level FAA) to developing an equation to identify feigned ADHD demonstrated superior sensitivity and specificity to genuine ADHD than the theory-driven equation in the first study (i.e., the scale-level FAA). Moreover, this equation showed excellent classification rates in the other study groups. Compared to the scale-level FAA, the item-level FAA functioned well in the mood disorders group, with an excellent classification rate. This is likely due to the item-level equation utilizing a greater amount of information than the theory-driven approach.

The item-level FAA index for the PAI may be particularly useful for the detection of feigned ADHD for many reasons. First, in terms of face validity the PAI does not appear to be a measure of ADHD symptoms meaning that individuals feigning symptoms are not likely to perceive it as a likely measure of symptom validity. However, it has been noted that individuals who endorse a high degree of historical ADHD symptoms tend to have elevated scores on PAI clinical scales (Hill, Pella, Singh, Jones, & Gouvier, 2009). As items related to attention or concentration difficulties are dispersed among the 344 questions, individuals who are feigning ADHD may over-endorse these questions beyond individuals with legitimate ADHD diagnoses. Secondly, because the PAI will also alert clinicians to comorbid mood and anxiety disorders, any failure on the FAA can also be viewed in light of any comorbidities, and this can reduce the occurrence of false-positive attributions. Thirdly, due to the complexity inherent in a regression-based approach to embedded measures, one could speculate that symptom-coaching to perform well on this embedded measure may be especially difficult.

While the present research adds to the existing literature by attempting to address the problem of specificity in differentiating genuine ADHD from ADHD simulators, it is not without limitations. First, future research would do well to apply this equation in a group of known adult ADHD feigners and genuine ADHD individuals. Although the present study utilized simulators and those with a genuine ADHD diagnosis in the development of the regression equation, only control and mood/anxiety disorder groups were available for subsequent validation. Next, because simulators were used in the present study, examining the utility of this regression equation in non-simulators who were deemed to have put forth inadequate effort on multiple SVTs may be beneficial. In particular, it is possible that simulators exaggerated their symptoms to a greater degree than would a sophisticated ADHD feigner on a college campus. In contrast, it is also possible that the ADHD feigners were influenced by their own mental health concerns, and future research would do well to include mood screeners in simulator groups. Next, because the PAI was used in the diagnosis of ADHD for the clinical group, criterion contamination is present, and future studies where ADHD was diagnosed without the PAI would be beneficial. Next, future research should examine this indicator in samples that are more similar in age, gender, and racial/ethnic background. Finally, while some have argued against the theoretical selection of variables in a regression equation (e.g., Wolfe et al., 2010), we believe that our model results in a clinician-friendly equation that also has high sensitivity and specificity in differentiating those feigning ADHD from controls. Limitations also exist with the item-level driven analysis. For this specific approach, the sample may have been underpowered for regression analysis of all PAI items to identify incrementally valid items for inclusion into the equation.

While results from both the scale-level and item-level regression equation approaches are promising, they may be sample specific and lack generalizability to other samples. Consequently, examination of the identified PAI scales and items used in the regression equations should be subject to additional validation in other samples using ADHD simulators and known clinical groups.

To our knowledge, the present study is the first to apply a regression-based approach to the PAI for the purpose of developing an embedded validity measure. Other studies have successfully applied similar methods in developing embedded indices for neuropsychological measures (e.g., Hill, Womble, & Rohling, 2015; Wolfe et al., 2010). In summary, we developed both scale-level and item-level logistic regression equations for the PAI to detect feigned ADHD that demonstrated good psychometric properties and do not appear to have high false positive rates in non-ADHD samples. We hope that these results will provide clinicians with additional tools in detecting feigned ADHD and stimulate further research to investigate new ways to detected ADHD symptom exaggeration.

Funding

The authors do not have any funding sources to report.

Conflict of interest

None declared.

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 Boone K. B. (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 366–383). New York: Guilford.

American Academy of child and Adolescent Psychiatry. (2007). Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 894–921. http://dx.doi.org/10.1097/chi.0b013e318054e724.

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

Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment*, *87*, 84–94. http://dx.doi.org/10.1207/s15327752jpa8701_07.

Arria, A. M., & DuPont, R. L. (2011). Nonmedical prescription stimulant use among college students: Why we need to do something and what we need to do. *Journal of Addictive Disorders*, *29*, 417–426. http://dx.doi.org/10.1080/10550887.2010.509273.

Babcock, Q., & Byrne, T. (2000). Student perceptions of methylphenidate abuse at a public liberal arts college. *Journal of American College Health*, *49*, 143–145.

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

Booksh, R. L. (2005). Ability of college students to simulate ADHD on objective measures of attention. (Doctoral dissertation, Louisiana State University). Retrieved from http://etd.lsu.edu/docs/available/etd-11162005-111135/

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. http://dx.doi.org/10.1177/1087054708329927.

Boonstra, A. M., Oosterlaan, J., Sergeant, J. A., & Buitelaar, J. K. (2005). Executive functioning in adult ADHD: A meta-analytic review. *Psychological Medicine*, *35*, 1097–1108. http://dx.doi.org/10.1017/S003329170500499X.

Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). MMPI-2: Manual for administration and scoring *(Rev Ed.)*. Minneapolis, MN: University of Minnesota Press.

Conti, R. P. (2004). Malingered ADHD in adolescents diagnosed with conduct disorder: A brief note. *Psychological Reports*, *94*, 987–988. http://dx.doi.org/10.2466/pr0.94.3.987-988.

Frazier, T. W., Frazier, A. R., Busch, R. M., Kerwood, M. A., & Demaree, H. A. (2008). Detection of simulated ADHD and reading disorders using symptom validity measures. *Archives of Clinical Neuropsychology*, *23*, 501–509. http://dx.doi.org/10.1016/j.acn.2008.04.001.

Fuermaier, A. B. M., Tucha, L., Koerts, J., Weisbrod, M., Grabemann, M., Zimmermann, M., et al. (2016). Evaluation of the CAARS infrequency index for the detection of noncredible ADHD symptom report in adulthood. *Journal of Psychoeducational Assessment. Advance online publication*. http://dx.doi.org/10.1177/0734282915626005.

Gaines, M. V., Giles, C. L., & Morgan, R. D. (2012). The detection of feigning using multiple PAI scale elevations: A new index. *Assessment*, *20*, 437–447. http://dx.doi.org/10.1177/1073191112458146.

Glassmire, D. M., Bierley, R. A., Wisniewski, A. M., Greene, R. L., Kennedy, J. E., & Date, E. (2003). Using the WMS-III faces subtest to detect malingered memory impairment. *Journal of Clinical and Experimental Neuropsychology*, *25*, 465–481.

Goldberg, T. E., Aloia, M. S., Gourovitch, M. L., Missar, D., Pickar, D., & Weinberger, D. R. (1998). Cognitive substrates of thought disorder, I: The semantic system. *American Journal of Psychiatry*, *155*, 1671–1676.

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. R. (2011). Detection of malingered ADHD using the MMPI-2-RF. *Psychological Injury and Law*, *4*, 32–43. http://dx.doi.org/10.1007/s12207-011-9100-9.

Harrison, A. G. (2006). Adults faking ADHD: You must be kidding! (2006). *The ADHD Report*, *14* (*4*), 1–7. http://dx.doi.org/10.1521/adhd.2006.14.4.1.

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. http://dx.doi.org/10.1016/j.acn.2007.03.008.

Herman, L., Shtayermman, O., Aksnes, B., Anzalone, M., Cormerals, A., & Liodice, C. (2011). The use of prescription stimulants to enhance academic performance among college students in health care programs. *Journal of Physician Assistant Education*, *22*, 15–22. http://dx.doi.org/10.1097/01367895-201122040-00003.

Hervey, A. S., Epstein, J. N., & Curry, J. F. (2004). Neuropsychology of adults with attention-deficit/hyperactivity disorder: A meta-analytic review. *Neuropsychology*, *18*, 485–503. http://dx.doi.org/10.1037/0894-4105.18.3.485.

Hill, B. D., Pella, R. D., Singh, A. N., Jones, G., & Gouvier, W. D. (2009). The Wender Utah Rating Scale: ADHD diagnostic tool or personality index? *Journal of Attention Disorders*, *13*, 87–94. http://dx.doi.org/10.1177/1087054708320384.

Hill, B. D., Womble, M., & Rohling, M. L. (2015). Logistic regression function for detection of suspicious performance during baseline evaluations using CNS Vital Signs. *Applied Neuropsychology: Adult*, *22*, 233–240. http://dx.doi.org/10.1080/23279095.2014.910215.

Jasinski, L. J., Harp, J. P., Berry, D. T. R., 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. http://dx.doi.org/10.1080/13854046.2011.630024.

Killgore, W. D. S., & DellaPietra, L. (2000). Using the WMS- III to detect malingering: Empirical validation of the Rarely Missed Index (RMI). *Journal of Clinical and Experimental Neuropsychology*, *22*, 761–771.

Lakhan, S. E., & Kirchgessner, A. (2012). Prescription stimulants in individuals with and without attention deficit hyperactivity disorder: Misuse, cognitive impact, and adverse effect. *Brain and Behavior*, *2*, 661–677. http://dx.doi.org/10.1002/brb3.78.

Langeluddecke, P. M., & Lucas, S. K. (2003). Quantitative measures of memory malingering on the Wechsler Memory Scale–Third Edition in mild head injury litigants. *Archives of Clinical Neuropsychology*, *18*, 181–197.

Mannuzza, S., & Klein, R. G. (2000). Long-term prognosis in attention-deficit/hyperactivity disorder. *Child and Adolescent Psychiatry*, *9*, 711–726.

McCabe, S. E., Knight, J. R., Teter, C. J., & Wechsler, H. (2005). Non-medical use of prescription stimulants among US college students: Prevalence and correlates from a national survey. *Addiction (Abingdon, England)*, *99*, 96–106. http://dx.doi.org/10.1111/j.1360-0443.2005.00944.x.

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. http://dx.doi.org/10.1016/j.comppsych.2004.02.006.

McCrae, R. R., & Costa, P. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, *51*, 882–888. http://dx.doi.org/10.1037/0022-006x.51.6.882.

McGough, J. J., & Barkley, R. A. (2004). Diagnostic controversies in adult attention deficit hyperactivity disorder. *The American Journal of Psychiatry*, *161*, 1948–1956. http://dx.doi.org/10.1176/appi.ajp.161.11.1948.

Mittenberg, W., Theroux-Fichera, S., Zielinski, R., & Heilbronner, R. L. (1995). Identification of malingered head injury on the Wechsler Adult Intelligence Scale–Revised. *Professional Psychology: Research and Practice*, *26*, 491–498.

Mogge, N. L., Lepage, J. S., Bell, T., & Ragatz, L. (2010). The negative distortion scale: A new PAI validity scale. *The Journal of Forensic Psychiatry & Psychology*, *21*. http://dx.doi.org/10.1080/14789940903174253.

Morey, L. C. (1991). *The Personality Assessment Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Morey, L. C. (1996). *An interpretive guide to the Personality Assessment Inventory: Professional manual* (2nd ed.). Lutz, FL: Psychological Assessment Resources.

Musso, M. W., & Gouvier, W. D. (2004). "Why is this so hard?" A review of detection of malingered ADHD in college students. *Journal of Attention Disorders*, *18*, 186–201. http://dx.doi.org/10.1177/108705471244197.

Musso, M. W., Hill, B. D., Barker, A. A., Pella, R. D., & Gouvier, W. D. (2014). Utility of the Personality Assessment Inventory for detecting malingered ADHD in college students. *Journal of Attention Disorders*. http://dx.doi.org/10.1177/1087054714548031Advance online publication.

Pella, R. D., Hill, B. D., Shelton, J. T., Elliott, E., & Gouvier, W. D. (2012). Evaluation of embedded malingering indexes in a non-litigating clinical sample using control, clinical, and derived groups. *Archives of Clinical Neuropsychology*, *27*, 45–57.

Pettersson, R., Söderström, S., & Nilsson, K. W. (2015). Diagnosing ADHD in adults: An examination of the discrimative validity of neuropsychological tests and diagnostic assessment instruments. *Journal of Attention Disorders*. http://dx.doi.org/10.1177/1087054715618788.

Quinn, C. A. (2003). Detection of malingering in assessment of adult ADHD. *Archives of Clinical Neuropsychology*, *18*, 379–395. http://dx.doi.org/10.1016/S0887-6177(02)00150-6.

Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, *20*, 33–65. http://dx.doi.org/10.1016/j.acn.2004.02.005.

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.

Rogers, R., Sewell, K. W., Morey, L. C., & Ustad, K. L. (1996). Detection of feigned mental disorders on the Personality Assessment Inventory: A discriminant analysis. *Journal of Personality Assessment*, *67*, 629–640. http://dx.doi.org/10.1207/s15327752jpa6703_15.

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. http://dx.doi.org/10.1016/j.acn.2005.04.005.

Sharp, J. T., & Rosén, L. A. (2007). Recreational stimulant use among college students. *Journal of Substance Use*, *12*, 71–82. http://dx.doi.org/10.1080/14659890601022865.

Shipley, W. C. (2004). *Shipley Institute of Living Scale* (2nd ed). Ottowa, Ontario, Candada: MultiHealth Systems.

Sollman, M. J., Ranseen, J. D., & Berry, D. T. (2010). Detection of feigned ADHD in college students. *Psychological Assessment*, *22*, 325–335. http://dx.doi.org/10.1037/a0018857.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. (1970). *The state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologist Press.

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. http://dx.doi.org/10.1080/09084280701509083.

Upadhyaya, H. P., O'Rourke, K., Sullivan, B., Wang, W., Rose, K., Deas, D., et al. (2005). Attention-deficit/hyperactivity disorder, medication treatment, and substance use patterns among adolescents and young adults. *Journal of Child & Adolescent Psychopharmacology*, *15*, 799–809. http://dx.doi.org/10.1089/cap.2005.15.799.

Ward, M., Wender, P., & Reimherr, F. (1993). The Wender Utah Rating Scale: An aid in retrospective diagnosis of children with ADHD. *American Journal of Psychiatry*, *160*, 245–256.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *WMS-III administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Weyandt, L. L., & DuPaul, G. J. (2008). ADHD in college students: Developmental findings. *Developmental Disabilities Research Reviews*, *14*, 311–319. http://dx.doi.org/10.1002/ddrr.38.

Wolfe, P. L., Millis, S. R., Hanks, R., Fichtenberg, N., Larrabee, G. J., & Sweet, J. J. (2010). Effort Indicators within the California Verbal Learning Test-II (CLTI-II). *The Clinical Neuropsychologist*, *24*, 153–168. http://dx.doi.org/10.1080/13854040903107791.

White, B. P., Becker-Belase, K., & Grace-Bishop, K. (2006). Stimulant medication use, misuse, and abuse in an undergraduate and graduate student sample. *Journal of American College Health*, *54*, 261–268. http://dx.doi.org/10.3200/JACH.54.5.261-268.