Psychological Assessment
2017, Vol. 29, No. 12, 1447–1457

1040-3590/17/$12.00 http://dx.doi.org/10.1037/pas0000435

Intentional Inattention: Detecting Feigned Attention-Deficit/Hyperactivity Disorder on the Personality Assessment Inventory

Shannon T. Smith
Francis Marion University

Jennifer Cox
University of Alabama

Elyse N. Mowle and John F. Edens
Texas A&M University

Given the increasing number of college students seeking Attention-Deficit/Hyperactivity Disorder (ADHD) diagnoses as well as the potential secondary gains associated with this disorder (e.g., access to stimulant medication, academic accommodations), the detection of malingered symptom presentations in this population is a major concern. The present study examined the ability of validity indicators on the widely used Personality Assessment Inventory (PAI; Morey, 1991) to distinguish between individuals experiencing genuine ADHD symptoms and individuals instructed to present with ADHD symptomatology for secondary gain. Sixty-six participants who successfully simulated ADHD (based on elevations on the Conners' Adult ADHD Rating Scale; Conners, Erhardt, & Sparrow, 1998) were compared with a sample of undergraduate students meeting diagnostic criteria for ADHD ($N = 22$) and an archival sample of adults who received an ADHD diagnosis at a university psychology clinic following a comprehensive psychological evaluation ($N = 41$). Successful simulators obtained significantly higher scores on all relevant PAI validity indicators compared with the clinical and archival comparison samples, with the Rogers Discriminant Function demonstrating the highest predictive accuracy (AUC = .86). Traditional cut scores on the Negative Impression (NIM) validity scale used to designate probable malingering, however, were not sensitive to simulated ADHD symptoms, although they did demonstrate excellent specificity. The PAI may be informative as an indicator of potentially exaggerated or malingered symptom presentation, but alternative cut scores for symptom validity indicators may be necessary to maximize its utility in these particular types of psychological evaluations.

Public Significance Statement
Results from this study reflected that people instructed to feign Attention-Deficit/Hyperactivity Disorder (ADHD) did not obtain elevations at traditional levels on Personality Assessment Inventory (PAI) scales designed to measure symptom validity that would ordinarily raise concerns about possible symptom exaggeration. Clinicians should consider alternate cut scores on the PAI validity scales in conjunction with other assessment data when assessing profile validity in ADHD evaluations.

Keywords: ADHD, personality assessment inventory, malingering, RDF, NIM

Recent years have seen a dramatic increase in the number of college students seeking psychological evaluations assessing Attention-Deficit/Hyperactivity Disorder (ADHD) symptoms (American Council on Education, 1995; Hagar & Goldstein, 2001; Harrison, Edwards, & Parker, 2007). Estimates of the prevalence of ADHD in college students vary, with most ranging from 4%–8% (e.g., Weyandt & DuPaul, 2006). Confirming an initial diagnosis in this population is particularly difficult given the *Diagnostic and Statistical Manual of Mental Disorders-5* (*DSM–5*; American Psychiatric Association, 2013) criterion regarding symptom presentation prior to the age of 12 (age 7 in the previous *DSM* iteration). Whereas diagnosis in children and adolescents is likely to be informed by more objective information (e.g., school records and parental report), initial diagnosis in adults is often based primarily on self-reported current symptoms and retrospective recall of childhood behaviors (Murphy, Gordon, & Barkley, 2000; Solanto, Wasserstein, Marks, & Mitchell, 2012). Although a number of ADHD symptom inventories have been developed to assist clinicians in making these diagnoses, many of these measures lack specificity for the disorder and are prone to overidentification (Harrison et al., 2007; McCann & Roy-Byrne, 2004).

Shannon T. Smith, Department of Psychology, Francis Marion University; Jennifer Cox, Department of Psychology, University of Alabama; Elyse N. Mowle and John F. Edens, Department of Psychology, Texas A&M University.

John F. Edens is a coauthor of the Personality Assessment Inventory Interpretive Report for Correctional Settings (PAI–CS), which is published by Psychological Assessment Resources, Inc. He receives no royalties from the sale of the PAI itself, however.

Correspondence concerning this article should be addressed to Shannon T. Smith, Department of Psychology, Francis Marion University, P.O. Box 100547, Florence, SC 29502-0547. E-mail: ssmith@fmarion.edu

The dependence on self-report measures is especially problematic when considering the secondary gains that often accompany this diagnosis. For example, students with an ADHD diagnosis are usually entitled to academic accommodations (e.g., extended testing times, isolated test administrations, and enhanced access to lecture notes) due to requirements outlined by the Americans with Disabilities Act (Sollman, Ranseen, & Berry, 2010). Beyond educational allowances, individuals with an ADHD diagnosis may be eligible for government services and programs (e.g., student loan repayments waived; Harrison, 2006). The secondary gain of receiving stimulant medication as treatment for this diagnosis is perhaps of the utmost concern, given that stimulant medications may be relied upon as a study aid to enhance concentration even among those without attention deficit problems (Sollman et al., 2010) or utilized for recreational drug use purposes (e.g., McCabe, Knight, Teter, & Wechsler, 2005).

Given the aforementioned secondary gains associated with a diagnosis of ADHD among college students in particular and heightened concerns about symptom exaggeration, recent research has begun to examine malingered ADHD, the preliminary results of which suggest that the disorder can be successfully feigned (Harrison et al., 2007; Sollman et al., 2010). The accessibility of ADHD literature and research available on the Internet likely contribute to this phenomenon (Conti, 2004). Indeed, at least one study reported that students who presented with motivation to feign symptoms and information about ADHD derived from the Internet produced ADHD-consistent profiles on the ADHD Rating Scale: Current and Childhood Symptoms Checklist (ARS; Barkley & Murphy, 2005; Murphy & Barkley, 1995) and the Conners' Adult ADHD Rating Scale (CAARS; Conners, Erhardt, & Sparrow, 1998; Sollman et al., 2010). In a separate study of undergraduate students, ADHD simulators' elevations on the CAARS were indistinguishable from individuals with clinical diagnoses of ADHD, and the simulators also exhibited diminished performance on tests of reading and processing speed typically administered as objective indicators of concentration difficulties (Harrison et al., 2007). Research (Fisher & Watkins, 2008; Jasinski et al., 2011) has demonstrated similar results among additional self-report inventories designed to assess ADHD symptomatology (e.g., the ADHD Behavior Checklist, Murphy & Barkley, 1995; College ADHD Response Evaluation, Glutting, Sheslow, & Adams, 2002).

Considering the problems inherent to relying solely on self-report or even clinician interview measures to render ADHD diagnoses in adults, researchers have conducted simulation studies utilizing ostensibly objective measures to detect feigned ADHD. For example, one early study reported that on the Integrated Visual and Auditory Continuous Performance Test (Sandford & Turner, 1995), the full scale attention quotient robustly discriminated between simulators and ADHD patients (Cohen's $d = 1.87$, $p < .001$) as well as between non-ADHD controls and the simulators (Cohen's $d = 4.00$, $p < .001$; Quinn, 2003). However, the ADHD control group and feigning group (*SD*'s = 30.1 and 21.3, respectively) both evinced large standard deviations, limiting the utility of this scale to reliably detect simulators from ADHD controls (Frazier, Frazier, Busch, Kerwood, & Demaree, 2008; Quinn, 2003).

Additional promising research has utilized performance or symptom validity measures (e.g., Frazier et al., 2008); however, limitations exist in these studies as well (e.g., by not including a clinically diagnosed control group). Sollma et al. (2010) reported results regarding tests designed to detect feigned cognitive deficits. The Test of Memory Malingering (TOMM; Tombaugh, 1997), a performance validity test, demonstrated high specificity for ADHD (Hedge's $g = 1.60$). Additionally, results from Harrison et al. (2007) suggested that scores on the CAARS, when combined with exaggerated low scores on other standardized tests, may be useful in identifying students feigning ADHD (effect sizes ranging from $r = .30$ to $r = .50$). However, no ADHD measure or symptom or performance validity test has yet been able to successfully screen out feigned ADHD while maintaining high specificity for ADHD.

Although ADHD is most often assessed through symptom checklists and self-reported behavior, it may be possible to identify ADHD in adults using broad measures of psychopathology. For example, the Minnesota Multiphasic Personality Inventory–2 Restructured Form (MMPI-2-RF; Tellegen & Ben-Porath, 2008) validity scales have been examined in simulation studies of feigned deficits associated with ADHD symptoms (Harp, Jasinski, Shandera-Ochsner, Mason, & Berry, 2011). Harp and colleagues (2011) found that participants in their study instructed to feign ADHD obtained mean scores on relevant clinical scales of interest similar to participants diagnosed with ADHD who were instructed to answer honestly. The infrequent psychopathology responses (Fp-r) validity scale demonstrated the greatest sensitivity among the MMPI-2-RF validity scales in identifying feigned ADHD, though only when a revised cut score lower than the Fp-r score recommended in the MMPI-2-RF manual was employed.

One self-report measure that has been the focus of extensive research and clinical practice is the Personality Assessment Inventory (PAI; Morey, 1991). Designed to objectively assess salient clinical syndromes and treatment-related variables, the PAI enjoys widespread popularity as a clinical and research tool (Reidy, Sorensen, & Davidson, 2015). Although the measure does not contain a specific scale designed to directly assess ADHD symptoms, various clinical scales are theoretically and empirically related to this disorder. Specifically, Musso et al. (2011) found that the PAI schizophrenia (SCZ) scale was significantly elevated in college students diagnosed with ADHD compared with control participants. This elevation was driven by the SCZ Thought Disorder (SCZ-T) subscale, which assesses confusion, concentration problems, and disorganization of thought processes (Morey, 1991). Additionally, the PAI may be particularly useful in assisting with differential diagnosis in assessment evaluations due to scales tapping anxiety, depression, and other clinical domains that may better account for symptoms that overlap with ADHD (e.g., difficulty concentrating, problems completing tasks).

The PAI contains various indicators of systematic and nonsystematic response distortion. Indicators tapping systematic response distortion include Negative Impression (NIM) and Positive Impression (PIM) Scales, a Malingering Index (MAL), and the Rogers Discriminant Function (RDF). Multiple studies have examined the ability of these response distortion indicators to predict overreporting of psychopathology (Hawes & Boccaccini, 2009). The NIM scale is composed of PAI items rarely endorsed in the community and clinical samples; however, intentional symptom exaggeration is conflated with genuine distress levels. As such, moderate elevations on NIM cannot be readily attributed to malingering per se (Morey, 2007a). According to the PAI manual, NIM scores greater than or equal to 92*T* should be interpreted as

evidence of potential malingering, whereas optimal cut scores of 110*T* have been identified in simulations studies among participants overtly attempting to present as psychologically distressed (Morey, 2007a). The MAL Index was developed to assess response distortion based on eight features that occur more often among individuals instructed to feign psychological disturbance than in genuine clinical samples (e.g., underemphasizing negative symptoms associated with psychosis). The RDF is composed of a weighted combination of 20 PAI scale scores (Rogers, Sewell, Morey, & Ulstad, 1996), with raw scores greater than or equal to 0 (59*T*) suggesting overt negative response distortion. In contrast to NIM, scores on this indicator are similar in both clinical and community samples, suggesting that RDF elevations are not tapping genuine psychological distress (Morey, 2007a). The Infrequency scale (INF) was designed to detect responses to items that are rarely endorsed in clinical or community samples. Scores between 60*T* and 74*T* indicate an unusual response style that could be attributed to a variety of factors (e.g., carelessness, unique item interpretation, reading difficulties; Morey, 2007a).

A meta-analysis of studies using the PAI to detect overreporting of psychopathology found that each of these three validity indicators evinced strong predictive validity for identifying feigned distress (Hawes & Boccaccini, 2009). These scales produced larger effect sizes in simulation studies (involving presumably psychologically healthy individuals instructed to feign mental distress) in comparison with studies involving individuals judged to be malingering or as experiencing genuine psychopathology based on an external criterion (e.g., scores on the Structured Interview of Reported Symptoms [SIRS], Rogers, Bagby, & Dickens, 1992). Additionally, effect sizes were larger for studies in which disorders characterized by more severe psychopathology (e.g., schizophrenia) were feigned in comparison with less severe diagnoses (e.g., anxiety disorders). RDF was noted to have decreased predictive validity in comparison with NIM in forensic and correctional populations as well as in known groups designs—although most of the "known" groups were identified solely based on the SIRS, which demonstrates generally modest associations with RDF (cf. Edens, Poythress, & Watkins-Clay, 2007; Morey, 2007a). RDF's orthogonal relationship with genuine psychopathology may account for the decreased covariance between the SIRS and RDF to the extent that the SIRS (similar to NIM) may be conflated with genuine psychopathology (Hawes & Boccaccini, 2009).

Finally, the PAI also includes indicators of positive distortion, on which very low scores could suggest attempts at feigned psychopathology. For example, scores on PIM tap overtly defensive responding or lack of insight. Low scores on PIM (i.e., reflecting a tendency to be more open about minor flaws than the average individual) may be expected when people are intentionally exaggerating psychological distress, consistent with results from simulation studies (Morey & Lanier, 1998).

Due to research thus far reflecting that symptom validity measures and other indicators demonstrate at best moderate sensitivity for detecting feigned ADHD, Rios and Morey (2013) investigated the ability of PAI-Adolescent (PAI-A; Morey, 2007b) response distortion indicators to detect simulated ADHD symptoms. College students aged 17 or 18 were instructed to complete the PAI-A and CAARS and simulate ADHD under either coached (i.e., received information regarding ADHD symptoms) or uncoached (i.e., did not receive information regarding ADHD symptoms) conditions. Scores from these participants were then compared with profiles of 17- and 18-year olds with confirmed ADHD diagnoses. Although approximately half of the simulating participants obtained CAARS scores in the range indicative of ADHD symptomatology, MAL, NIM, and RDF successfully distinguished between actual clinical patients and successful simulators. Overall, RDF most accurately differentiated between the simulation and clinical comparison groups, with a Cohen's *d* effect size of 2.26. Of note, the PIM scale and MAL index evinced poorer accuracy differentiating between simulated and clinical groups in the coached condition relative to the uncoached condition. The successful ADHD simulators scored consistently higher on the negative response distortion indices and lower on PIM than did participants diagnosed with ADHD. Importantly, however, mean elevations on the PAI negative dissimulation scales in the aforementioned study were generally not in a range that typically would arouse heightened suspicion of symptom exaggeration or malingering. Indeed, typical NIM scores that would alert clinicians to possible response distortion are considerably higher (e.g., 81*T*+) than that identified as an optimal cut score (60*T*) in the Rios and Morey (2013) study.

In addition to examinations of existing PAI validity scales, one previous study has examined the ability to create an ADHD symptom validity scale (hereafter referred to as the Musso ADHD symptom validity scale or the Musso scale) using the PAI in a college population. Musso (2013) identified 14 PAI items that were significant predictors of simulated ADHD and summed the items into a new scale. The scale is comprised of items largely tapping negative affect (e.g., irritability and anxiety), dysphoric cognitions (e.g., hopelessness), and physiological symptoms of distress (e.g., sleep and appetite disruption) derived from the following PAI scales and subscales: Anxiety-Cognitive subscale (one item), Anxiety-Affective subscale (two items), Anxiety-Physiological subscale (three items), Depression-Physiological subscale (two items), Anxiety Related Disorders-Traumatic Stress subscale (one item), Borderline Features-Affective Instability subscale (two items), Mania-Irritability subscale (one item), Suicidal Ideation scale (one item), and Positive Impression Management scale (one item, reversed). A cut score greater than 16 on this scale demonstrated good specificity and sensitivity for detecting simulated ADHD in the derivation sample.

Given the relatively modest amount of existing research summarized above, there is clearly a need for more investigations of the utility of the PAI to detect malingered ADHD symptoms. In this study we examined the ability of relevant PAI validity indicators to distinguish between individuals experiencing genuine ADHD symptoms and individuals instructed to simulate ADHD symptomatology for secondary gain. We hypothesized that the validity indicators would successfully distinguish between simulators and those instructed to complete the inventory according to standard administration instructions, as evidenced by the simulating group obtaining higher scores on RDF, MAL, NIM, and INF and significantly lower scores on PIM. Regarding clinical scales, we hypothesized that the simulation group would obtain significantly higher scores on each clinical scale of interest, compared to the standard instruction responder groups. Finally, we were interested in exploring the sensitivity and specificity of the Musso ADHD symptom validity scale, as described by Musso (2013). Given that the item content of this scale largely taps dysphoric

affect, which is associated with high NIM scores, we expected that the Musso scale would perform similarly to NIM.

Method

Participants

Three groups, including one experimental (simulation) condition and two control conditions (clinical comparison sample and archival comparison sample), were examined in this study. Participants were recruited through the psychology department research pool at a large southwestern university, with the exception of data analyzed from the archival control group, which was obtained from a psychology training clinic at the same university.

Simulation sample. Participants for the simulation group originally included 107 undergraduate students enrolled in an introductory psychology class. Data for 13 participants were excluded from all analyses due to excessive missing data or failing a manipulation check (described below), resulting in a reduced sample of 94 participants.[1]

So as to not include protocols from individuals who did not present with significant ADHD symptomatology—which would preclude any need to evaluate the veracity of their symptoms—participants were only retained in the final sample if they "successfully feigned" by elevating the CAARS *DSM–IV* ADHD Symptoms Total Scale to a minimal threshold of at least 70*T*. This exclusion criterion resulted in a final sample of 66 participants. The mean age for this group of 66 participants was 18.53 years ($SD = 0.88$). Over half of the participants (54.50%) self-identified as female (data from one participant missing), and the majority of the sample (69.70%) self-identified as European American, followed by Hispanic (21.20%), "other" (7.60%), and African American (1.50%).

T scores on the CAARS greater than or equal to 70*T* correspond to the 98th percentile of the normative sample and are considered clinically elevated (Conners, Erhardt, & Sparrow, 1999). The cutoff of 70*T* is recommended for use in populations of adults without previously identified ADHD symptomatology (Conners et al., 1999). The *DSM–IV* ADHD Symptoms Total Scale was selected as the criterion for identifying successful feigners because this scale, created from the sum of the *DSM–IV* Inattentive Symptoms scale and the *DSM–IV* Hyperactive-Impulsive Symptoms scale, was developed to assess *DSM–IV* ADHD criteria. Those with scores in the clinically significant range on this scale "meet the criteria for ADHD, as described in the *DSM-IV*" (Conners et al., 1999, p. 24). We did not utilize the ADHD index as it was developed from a discriminant function analysis within a relatively small sample. To our knowledge that scale's ability to differentiate genuine ADHD from feigned ADHD has not been subject to cross-validation.

Simulation sample procedures. After giving written informed consent, participants in the experimental group completed the study materials in large groups of approximately 35 people each. Participants first completed a demographic questionnaire which included questions regarding any current/past ADHD diagnosis or treatment for ADHD symptoms as well as current or previous significant difficulties with individual *DSM–IV–TR* defined ADHD symptoms (e.g., sustaining attention when being spoken to directly, interrupting others who are speaking). Participants were then provided a written description of ADHD symptoms that was also read aloud to the group. This description was modeled after information obtained from a popular health information website written for the general public. Information was derived from this source, instead of directly from the *DSM–IV–TR*, to increase the ecological validity of the study design. Presumably, a young adult seeking information about ADHD for the purposes of obtaining a diagnosis for secondary gain is more likely to search the Internet rather than consult a psychiatric or psychological text.

Participants were then given the following instructions for completing the PAI and CAARS:

> Today you will be asked to complete two self-report tests. We would like for you to appear on these tests as if you have Attention-Deficit/Hyperactivity Disorder. These two tests are going to ask about what types of problems or symptoms you are suffering from. We want you to respond to the items on the test in order to create the impression that you are suffering from this disorder. However, **the tests have been designed to identify people** who do not answer honestly, so **we want you to try and beat the test by faking without getting caught doing so. If you are able to beat the test,** that is appear as though you have Attention-Deficit/Hyperactivity Disorder, **you will receive $10**. **If the test detects that you are lying or attempting to manipulate it in any way, you will not receive the $10.** Regardless of your ability to beat the test, you will receive experimental credit for your time.

Despite this instructional set, all participants received the monetary compensation regardless of their performance on the test. Following completion of the PAI and CAARS, participants completed a brief manipulation check consisting of a single face valid, multiple choice question that asked participants to select how they were instructed to respond to the PAI and CAARS from four response options (honestly; as if attempting to look psychotic; as if attempting to create a favorable impression of oneself; as if experiencing ADHD). Finally, participants responded to a series of dichotomous questions designed to ascertain their approach to completing the questionnaires (e.g., "[Did you] try to exaggerate on most or all of the test items?;" "[Did you] try to guess which items were designed to detect faking?") and a single open-ended question ("Please list any other test-taking approaches you used that were not listed above").[2]

Clinical comparison sample. The first control comparison sample (hereafter referred to as the clinical comparison group) originally consisted of 28 undergraduate students enrolled in an introductory psychology class at the same southwestern university. Participants in the sample were recruited based on their responses to questions administered during the departmental subject pool

[1] Data from five additional participants were also considered for exclusion due to obtaining either an ICN score $\geq 73T$ ($N = 3$) or INF $\geq 86T$ ($N = 2$), suggestive of random responding (Morey, 2007a). However, data from these participants were retained for the present analyses given that (a) these elevations might be related to genuine attempts to engage in response distortion in ADHD evaluations and (b) removal of these cases did not substantially impact the overall results obtained. Additional information regarding these analyses can be obtained from the first author.

[2] These strategy-related questions were not used as a manipulation check, and participants' data were not excluded based on their responses to these additional questions. Examination of participants' responses to these questions revealed that they were largely uninformative with respect to results obtained. Further details about these questions and associated analyses can be obtained from the first author.

mass prescreening. Recruitment was undertaken across two semesters during which time the total subject pool consisted of approximately 3,000 students. If students during the prescreening endorsed having a current/past ADHD diagnosis or current/past treatment with prescription medication for ADHD symptoms, they were invited (approximate $N = 150$), via e-mail, to register to participate in a follow-up study.

Clinical comparison procedures. After providing written informed consent, 28 individuals in the clinical comparison group completed individual in-person interviews with a trained research assistant. Evaluations consisted of a semi-structured clinical interview including questions derived by the study authors to measure *DSM–IV–TR* ADHD criteria. Questions focused on past/current ADHD symptoms and also included questions concerning prescription medication used to treat ADHD symptoms. After completing the interview, participants completed the PAI and CAARS measures under standard administration instructions.

The first two authors later independently reviewed participants' interview responses to determine whether their responses met criteria for any *DSM–IV* ADHD diagnosis, with the exception of ADHD Not Otherwise Specified ($\kappa = .73$). *DSM–IV* diagnostic criteria required an individual to meet at least six criteria indicative of inattention (ADHD Predominantly Inattentive Type), hyperactivity/impulsivity (ADHD Predominantly Hyperactive Type), or both (ADHD Combined Type) in at least two domains, and there must have been evidence of impairment prior to age 7. Data from six participants were excluded due to failing to meet these eligibility criteria (i.e., the raters agreed that the participants' responses during the interviews did not meet criteria for a *DSM–IV* ADHD diagnosis), resulting in a final sample of 22.[3] The mean age for this sample of 22 participants was 18.77 years ($SD = 1.19$) with approximately half (54.50%) of participants in this sample identifying as male and the vast majority (90.90%) identifying as European American, followed by Hispanic (4.50%), and "other" ethnicity (4.50%).

Archival comparison sample. The third sample (hereafter referred to as the archival comparison sample; $N = 41$) consisted of individuals who had undergone a psychological evaluation for ADHD at a university psychology training clinic at the same southwestern university. Clients at this clinic were generally self-referred for evaluations or referred by academic counselors or other mental health professionals. The archival data for these participants were included in the current study if (a) files included complete PAI and CAARS profiles and (b) examinees had consented in writing at the time of their assessment that their assessment results could be used for research purposes. If the data from a clinic client's evaluation met these criteria, their PAI and CAARS responses, as well as pertinent demographic and diagnostic information, were coded into the study database by a graduate-level research assistant. Over half of this group identified as male (66%) with a mean age of 23.22 years ($SD = 5.33$). Finally, the majority of this group identified as Caucasian (46.3%), followed by Hispanic (26.8%), African American (22.0%) and "other" (4.9%).

Data for the archival comparison group were obtained through records securely housed in the university clinic. These data, including demographic information, PAI scale and subscale scores (item level data were not available for the analyses), CAARS scores, and diagnostic information, were collected and recorded in a secure database by a graduate student research assistant.

Measures

Personality Assessment Inventory. The Personality Assessment Inventory (PAI; Morey, 1991) is a self-report measure used to assess clinical, interpersonal, treatment, and validity domains. The PAI contains 344 items that are endorsed on a 4-point Likert scale ranging from False to Very True. For the purposes of this research, PAI NIM, PIM, INF, RDF, and MAL (described previously) were investigated with respect to their ability to differentiate simulated and standard instruction profiles. Data from other clinical scales of interest were also examined due to the overlap of item content with symptoms and behaviors consistent with ADHD. These scales included the Schizophrenia Thought Disorder subscale (SCZ-T), the Mania Irritability subscale (MAN-I), and Mania Activity Level (MAN-A) subscale, the Anxiety Physiological symptoms (ANX-P) subscale, the Borderline Features Self-Harm subscale (BOR-S), and the Antisocial Features Stimulus-Seeking subscale (ANT-S). Supplemental analyses were also run investigating the Musso ADHD symptom validity scale (Musso, 2013). For comparison purposes and to aid future investigations, additional exploratory supplemental analyses were run for the PAI scales reported by Rios and Morey (2013) as well as the Treatment Rejection scale (RXR). RXR was investigated given the potential for this scale to provide further relevant information regarding the treatment-seeking factor that may influence the responses of those presenting for psychological evaluations.

Connors Adult Attention Rating Scale (CAARS). The CAARS (Conners, Erhardt, & Sparrow, 1998) is a 66-item self-report inventory designed to identify the presence and severity of ADHD symptoms in adulthood (Conners et al., 1999). Items address symptoms pertaining to attention deficits, hyperactivity, and impulsivity. The measure yields subscales related to core *DSM–IV* ADHD criteria (Inattention/Memory Problems, Hyperactivity/Restlessness, Impulsive/Emotional Lability, and Problems with Self-Concept). In addition, the measure includes three subscales consistent with *DSM–IV* inattentive symptoms, *DSM–IV* hyperactive-impulsive symptoms, and *DSM–IV* total ADHD symptoms. Finally, the measure includes a total ADHD Index, considered to be a summary of ADHD symptomatology, and an inconsistency index, designed to identify inconsistent and random responding.

Results

Noted earlier, participants in the simulation sample were classified as having "successfully feigned" if they obtained elevations $\geq 70T$ on the CAARS *DSM–IV* ADHD Symptoms Total Scale. Using this criteria, 70.21% ($n = 66$) of the participants were designated successful simulators. Successful simulators, not surprisingly, scored significantly higher than unsuccessful simulators on the CAARS *DSM–IV* ADHD Symptoms Total Scale and also

[3] No significant differences in results were obtained when the six omitted cases were included in the analyses. Further information can be obtained from the first author.

on all PAI validity indices and subscales of interest (NIM, PIM, INF, RDF, MAL, ANX-P, MAN-I, SCZ-T, BOR-S, and ANT-S).[4]

Next, one-way analysis of variance (ANOVA) analyses were conducted comparing scores on the scales of interest across the three samples (successful simulation, clinical, and archival comparison). Given the unequal sample sizes, Levene's test of Homogeneity of Variance was employed to determine the extent to which this assumption was violated in our PAI variables of interest. With the exception of the NIM scale, the assumption of homogeneity was not violated. Because violations of this assumption are not expected to significantly impact omnibus F tests (Tabachnick & Fidell, 2007), we proceeded with the ANOVA analyses. Scale score means, standard deviations, and Cohen's d effect size differences associated with Bonferroni post hoc comparisons are presented in Table 1. All scores reported are T scores based on relevant normative groups for the instruments. With three exceptions, mean scores for the clinical and archival comparison samples were significantly lower than for the simulation sample, though score elevations in the clinical and archival comparison samples did not significantly differ. Scores on the BOR-S subscale, which assesses the tendency to act impulsively and destructively in response to negative affect, were significantly lower for the clinical sample than in the archival comparison and successful simulation samples, respectively. Additionally, mean score elevations on the SCZ-T subscale were significantly lower in the clinical sample than in the simulation sample. Finally, PIM scores were significantly higher in the clinical and archival clinical samples than in the successful simulation sample (i.e., successful feigners' responses were less defensive than those from the clinical and archival samples).

Receiver operating characteristic curve and chi square analyses were conducted utilizing the PAI validity scales to predict simulation status. Area under the curve (AUC) and standard errors, confidence intervals, sensitivity, and specificity values generated from these analyses are presented in Table 2. For these analyses, PIM scores were reflected in the MedCalc software program to maintain consistency in the interpretation of AUC values. The NIM cut score of 60T reported by Rios and Morey (2013) evidenced greater sensitivity than the 81T score recommended in the PAI manual. RDF evidenced the highest predictive accuracy with an AUC of .86, and MAL demonstrated the poorest predictive ability overall. A comparison of AUC values was conducted via the DeLong, DeLong, and Clarke-Pearson (1988) recommended procedures with MedCALC statistical software. Results indicated that AUC values did not significantly differ at generally accepted p values less than .01, with two exceptions. The RDF AUC value was significantly greater than the AUC values evinced by both MAL ($z = 4.23$, $p < .001$) and PIM ($z = 3.51$, $p = .001$).

Supplemental analyses next investigated the ability of the Musso ADHD symptom validity scale to identify simulated ADHD. Analyses for this scale were restricted to the successful simulation sample ($n = 66$) and the clinical comparison sample ($n = 22$), given that item-level data necessary to create the new scale were not available for the archival comparison sample. Cronbach's alpha for the Musso scale across both samples was .84. The clinical control group obtained a raw score mean of 8.91 ($SD = 6.60$), and the mean within the successful simulation group was 18.80 ($SD = 8.39$). Investigation of the Musso scale's ability to predict feigning status revealed that the new scale evidenced an AUC value of .82 ($SE = .05$, 95% CI [.72, .93]). Given that the Musso scale demonstrated large positive correlations with NIM ($r = .76$) and MAL ($r = .62$) as well as a moderate positive correlation with RDF ($r = .34$), comparisons of AUC values for the Musso scale with NIM, PIM, RDF, and INF within the restricted samples were performed. Results revealed no significant differences (all p's $> .05$).

Results from exploratory analyses for the PAI clinical scales and Treatment Rejection Scale are reported in Table 3. Generally, participants in the ADHD simulation group obtained mean scores on the clinical scales at approximately two standard deviations above the community normative mean, participants in the clinical comparison sample produced scores on the clinical scales ranging from average to one standard deviation above the mean, and data from the archival comparison sample reflected scores on the clinical scales at approximately one standard deviation above the mean. Exceptions to this overall trend were on the Alcohol Problems (ALC) and Drug Problems (DRG) scales where participants in the simulation group produced an average score at approximately one standard deviation above the mean, and those in the clinical and archival comparison samples produced average scores on these scales.

Discussion

The aim of the present study was to examine the utility of the PAI validity indicators in detecting feigning of ADHD symptomatology in a sample of college students. Given the increase in college students seeking evaluations for ADHD symptoms (American Council on Education, 1995; Hagar & Goldstein, 2001; Harrison et al., 2007) and the secondary gains that may be associated with an ADHD diagnosis (Harrison, 2006; Sollman et al., 2010), it is important to consider the utility of the PAI in differentiating forthright from distorted responding.

Studies regarding investigations of feigned ADHD symptomatology on broad band self-report inventories as well as on instruments designed to assess symptoms associated with particular diagnostic constructs are important because of the frequency of administration of these instruments in assessment evaluations. In particular, broad band personality instruments may assist with the differential diagnosis of other diagnostic constructs (e.g., anxiety, depression) with symptoms that overlap with ADHD (e.g., distractibility). Considering research (e.g., Hawes & Boccaccini, 2009) suggesting that validity scales may be less sensitive to identifying feigned disorders that are less severe (e.g., anxiety), it is crucial to assess the extent to which recommended cut scores on validity indicators of broad band personality inventories such as the PAI detect negative dissimulation of ADHD. Individuals instructed to present themselves as experiencing symptoms consistent with ADHD and who obtained elevations $\geq$70T on the CAARS *DSM–IV* ADHD Symptoms Total Scale obtained significantly higher scores on PAI NIM, INF, MAL, and RDF and significantly lower scores on PIM, compared with clinical and archival comparison groups. These differences appeared robust, with large effect sizes when comparing the clinical and simulation samples and medium to large effect sizes found when comparing the archival with simulation samples. However, consistent with

[4] Results of these analyses can be obtained from the first author.

Table 1
ANOVA Comparisons Between Clinical, Archival, and Simulation Groups

Scale	Simulation (S) *M* (*SD*)	Clinical (C) *M* (*SD*)	Archival clinical (A) *M* (*SD*)	*F* value	Bonferroni comparison	Effect size C vs. S	Effect size A vs. S
NIM	70.86 (20.89)	53.91 (9.94)	56.61 (12.67)	12.91	C, A < S	−1.04	−.82
INF	63.02 (12.95)	51.77 (7.53)	52.83 (8.58)	15.12	C, A < S	−1.06	−.93
MAL	64.57 (15.63)	51.36 (8.44)	55.61 (11.28)	10.58	C, A < S	−1.05	−.66
RDF	65.26 (11.48)	51.58 (8.43)	47.12 (9.71)	41.69	C, A < S	−1.36	−1.71
PIM	34.00 (9.47)	41.86 (8.89)	39.88 (11.67)	7.05	C, A > S	.86	.55
ANX-P	76.80 (12.88)	61.14 (9.85)	61.49 (11.52)	26.70	C, A < S	−1.37	−1.25
BOR-S	76.83 (13.09)	56.14 (11.53)	65.90 (14.11)	23.00	C < A < S*	−1.68	−.80
MAN-I	69.20 (10.15)	62.77 (11.27)	60.78 (10.99)	8.78	C, A < S	−.60	−.80
MAN-A	75.26 (11.37)	61.82 (14.29)	69.34 (12.51)	10.52	C < S	−1.04	−.50
SCZ-T	80.85 (11.63)	70.36 (13.34)	77.51 (12.59)	6.10	C < S	−.84	−.28
ANT-S	81.47 (13.80)	66.50 (15.54)	64.90 (14.78)	19.97	C, A < S	−1.02	−1.16

Note. NIM = Negative Impression; INF = Infrequency; MAL = Malingering Index; RDF = Rogers Discriminant Function; PIM = Positive Impression; ANX-P = Anxiety-Physiological; BOR-S = Borderline Features-Self-Harm; MAN-I = Mania-Irritability; MAN-A = Mania-Activity Level; SCZ-T = Schizophrenia-Thought Disorder; ANT-S = Antisocial Features-Stimulus-Seeking.
* Cohen's *d* effect size for C vs. A on BOR-S = −.76.

data reported by Musso (2013), the simulation group average NIM score did not reach the level to suggest distorted responding based on cut scores developed to detect more severe forms of fabricated psychopathology (e.g., psychosis). This result is also in line with results reported by Harp et al. (2011) where the authors lowered the recommended cut score on the MMPI-2-RF Fp-r validity scale to optimize the scale's sensitivity for detecting feigned ADHD.

In comparison with prior research on this issue (e.g., Rios & Morey, 2013), the present study was unique in that it involved the investigation of individuals undergoing assessment in a clinical setting who were ultimately diagnosed with ADHD. Results obtained from this study speak to the treatment-seeking factor that may purportedly contribute to respondents demonstrating a 'cry for help' in their responses to self-report questionnaires to facilitate diagnosis and treatment for genuine psychological distress. Additionally, results revealed no significant differences between the clinical and archival comparison groups on the validity indicators, although results for the clinical scales were somewhat mixed (see Table 3 and additional discussion below).

When examining the utility of the PAI validity indicators to predict symptom exaggeration status, RDF demonstrated the highest predictive accuracy, with an AUC value of .86 (*SE* = .03). This finding is perhaps not surprising given research suggesting that RDF is independent of genuine psychopathology (Morey, 2007a) and more sensitive than other validity indicators to efforts to malinger minor forms of psychopathology (Morey, 2007a). Further, when examining results obtained with the PAI-A RDF, Rios and Morey (2013) reported a similar AUC (.81, *SE* = .05) providing further support to the utility of this index in detecting ADHD feigning.

These results are particularly striking in conjunction with the relatively low NIM score (60*T*) identified as an optimal cut score in this study as well as the Rios and Morey (2013) study. This NIM score is only a single standard deviation above the mean for the PAI community normative sample and is approximately two standard deviations below optimal cut scores discussed in meta-analytic research on this issue (Hawes & Boccaccini, 2009). Clinicians administering the PAI in ADHD evaluations may be especially likely to render false negative designations of malingering when considering the NIM symptom validity scale in isolation. As discussed previously, NIM taps a pessimistic outlook as well as bizarre symptomatology that resemble psychotic experiences. As such, this scale is unlikely to elevate when specific, relatively mild disorders are exaggerated.

RDF has been the subject of controversy in the research literature, to the extent that some (e.g., Hawes & Boccaccini, 2009) have advocated against using the index in clinical settings. Problems have largely been noted with regard to this index in offender samples and in studies using the SIRS to designate "known" or "suspected" groups of malingerers. However, to the extent that the

Table 2
ROC Analyses Results

Scale	AUC (*SE*)	AUC confidence interval	Source	Threshold	Sensitivity	Specificity
NIM	.75 (.04)	[.67, .83]	PAI manual	81T	27.3	95.2
			Rios and Morey (2013)	60T	65.2	74.6
INF	.75 (.04)	[.66, .83]	PAI manual	60T	76.2	60.9
MAL	.64 (.05)	[.54, .74]	PAI manual	84T	9.1	96.8
RDF	.86 (.03)	[.80, .92]	PAI manual	59T	68.2	82.5
PIM	.67 (.05)	[.58, .75]	This study	35T	56.1	68.3

Note. AUC = Area under the curve; NIM = Negative Impression; INF = Infrequency; MAL = Malingering Index; RDF = Rogers Discriminant Function; PIM = Positive Impression.

Table 3
Exploratory ANOVA Comparisons Between Clinical, Archival, and Simulation Groups on Additional PAI Scales

Scale	Simulation (S) *M* (*SD*)	Clinical (C) *M* (*SD*)	Archival clinical (A) *M* (*SD*)	*F* value	Bonferroni comparison	Effect size C vs. S	Effect size A vs. S
SOM	65.38 (13.20)	54.59 (11.49)	53.37 (9.87)	15.22	C, A < S	−.87	−1.03
ANX	73.27 (12.81)	59.36 (10.09)	62.85 (16.46)	12.10	C, A < S	−1.21	−.71
ARD	67.48 (16.48)	55.45 (12.48)	56.80 (12.48)	9.36	C, A < S	−.83	−.73
DEP	68.67 (11.02)	53.14 (8.11)	61.85 (11.98)	17.70	C < A < S*	−1.61	−.59
MAN	72.14 (9.59)	65.82 (11.59)	63.34 (10.58)	10.04	C, A < S	−.59	−.87
PAR	72.15 (14.30)	52.41 (8.67)	54.07 (11.58)	35.08	C, A < S	−1.67	−1.39
SCZ	74.86 (13.82)	58.50 (10.02)	64.05 (9.39)	19.84	C, A < S	−1.36	−.91
BOR	74.20 (11.66)	57.50 (10.03)	63.39 (11.72)	22.33	C, A < S	−1.54	−.92
ANT	81.26 (13.07)	65.59 (13.08)	62.24 (12.02)	32.07	C, A < S	−1.20	−1.51
ALC	63.05 (16.17)	53.50 (11.14)	52.61 (10.72)	8.63	C, A < S	−.69	−.76
DRG	64.24 (18.22)	50.00 (8.23)	46.93 (5.27)	22.57	C, A < S	−1.01	−1.29
RXR	44.80 (8.81)	51.18 (9.55)	43.07 (9.21)	5.97	C > A, S	.69	−.19

Note. SOM = Somatic Complaints; ANX = Anxiety; ARD = Anxiety-Related Disorders; DEP = Depression; MAN = Mania; PAR = Paranoia; SCZ = Schizophrenia; BOR = Borderline Features; ANT = Antisocial Features; ALC = Alcohol Problems; DRG = Drug Problems; RXR = Treatment Rejection.
* Cohen's *d* effect size for C vs. A on DEP = −0.85.

SIRS is conflated with genuine psychopathology, it would be expected that RDF would evince less predictive ability identifying malingered profiles where SIRS scores were used as the basis for malingering designations. Therefore, we conclude that further investigation of RDF's ability to identify feigned ADHD is warranted.

PIM scores were also significantly lower in the sample instructed to feign than in the clinical or archival comparison groups. This result is not surprising after considering that people attempting to exaggerate distress would not also be expected to simultaneously portray themselves as problem free and overly virtuous.

Finally, supplemental analyses were run investigating a scale independently developed to identify simulated ADHD on the PAI. AUC values obtained with the Musso ADHD symptom validity scale were comparable with the values obtained for RDF and NIM. Additionally, the Musso scale demonstrated moderate to large correlations with all symptom validity indices, especially NIM. An inspection of the items on the Musso scale revealed that the items tap primarily dysphoric affect which likely contributes to the high correlation with NIM. Additionally, the PAI contains no item overlap, and this was done intentionally to avoid psychometric problems (e.g., high scale intercorrelations). Given that the Musso scale is largely redundant with NIM, performs comparably to RDF and NIM, and would introduce an element of scale intercorrelation, our results do not suggest any particular advantage to pursuing the development of this scale as an addition to the inventory.

In contrast to results regarding the validity indicators, an inspection of the results for the clinical subscales of interest demonstrated a somewhat different presentation among the three samples. Results for the BOR-S subscale that taps impulsive, self-destructive behaviors demonstrated that the archival comparison group obtained elevations on this subscale significantly less than the simulation group, yet significantly higher than the clinical comparison group. The difference between the archival and clinical comparison groups may be due to the extent to which the ADHD symptoms were impacting their daily functioning. Specifically, the archival sample participants were presenting for a diagnostic evaluation, presumably as a precursor to treatment, suggesting their symptoms were likely untreated and interfering with their daily functioning. In contrast, the clinical comparison sample participants already had an ADHD diagnosis (based on their self-report and also verified via a semi-structured interview assessing past and current ADHD symptoms) and self-reported that they were either currently or previously treated for ADHD symptoms. Because a substantial portion of the clinical comparison sample reported during the interview that these participants were currently taking stimulant medication that directly addresses ADHD symptomatology, their self-reported disinhibited behaviors may be expected to appear less intense in comparison to their untreated counterparts.

Additionally, on two subscales that indicate heightened activity levels (MAN-A) and confusion and disturbances in thought process (SCZ-T), the archival comparison group did not significantly differ from either the simulation group or the clinical comparison group. Perhaps this nonsignificant difference in elevations on these latter two subscales between the simulation group and archival comparison group can be attributed to the treatment-seeking element reflecting psychological distress among the sample presenting for an assessment evaluation and that which might be exhibited by those attempting to exaggerate ADHD. Although, this result is difficult to interpret, given that the archival comparison group did not significantly differ from the clinical comparison group on these scales either. However, on the remaining two clinical subscales of interest (MAN-I and ANT-S) reflecting poor frustration tolerance and tendencies to engage in reckless, impulsive behaviors, a similar picture emerged as found with the validity scales (i.e., the clinical and archival comparison samples obtained significantly lower elevations on these scales than that obtained in the simulation sample).

The totality of these results suggests that heightened suspicion of symptom exaggeration may be warranted when the overall PAI clinical profile reflects (1) mild to moderate but not necessarily extreme elevations on symptom validity scales; (2) markedly below average scores on PIM; and (3) robust elevations on most or all of the subscales of interest investigated in this study, particularly SCZ-T which has the closest face validity approximation for ADHD symptomatology and MAN-A which is one of the lowest points on the PAI skyline reflecting 2 standard deviations above

the mean of the clinical normative sample (i.e., it is unusual to find extreme elevations on that subscale in clinical populations). When encountering such a profile, clinicians would certainly be prudent to proceed with administering performance validity instruments if they are not already regularly incorporated into the assessment battery.

Results from our exploratory analyses with the PAI clinical scales and Treatment Rejection scale were overall consistent with the findings previously discussed. Interestingly, results for the Treatment Rejection scale were in striking contrast with results for all other scales in that the clinical comparison sample mean was significantly higher on this scale than both the archival comparison and simulation samples. The Treatment Rejection Scale is interpreted such that higher scores reflect difficulty acknowledging shortcomings and problems and a tendency to externalize blame (Morey, 2007a). Although participants in the clinical comparison sample produced an average score of approximately 50*T* on this scale, scores in this range are interpreted clinically as indicating contentment with oneself and one's behaviors with no need to engage in treatment to effect change in one's life. In contrast, both the archival comparison sample and the simulation sample produced mean scores in a range that would be interpreted as reflecting treatment amenability (yet not overwhelming distress and unrealistic expectations for therapy). This scale may also capture the treatment-seeking element of the archival comparison sample comprised of individuals presenting for a psychological evaluation who ultimately received an ADHD diagnosis. The mean score on this scale for the simulation group is also consistent with the overall clinical picture portrayed by our participants instructed to feign ADHD who managed to endorse symptomatology relevant to ADHD without overtly appearing as though they were engaging in negative dissimulation according to traditionally recommended cut scores on the PAI symptom validity indices. On the Treatment Rejection scale, those in the simulation group appeared eager to participate in treatment but not as desperately seeking "immediate alleviation of his or her suffering" (Morey, 2007a, p. 46).

Further, five of the clinical scales encompassing subscales selected a priori for investigation (Anxiety, Mania, Schizophrenia, Borderline Features, and Antisocial Features) were elevated in the simulation sample at approximately two standard deviations above the mean. This result is not surprising given that these scales tap content related to symptomatology associated with ADHD. However, further investigation from the exploratory analyses of mean scale elevations in the simulation sample reflected that three additional scales were elevated approximately two standard deviations above the mean (Anxiety-Related Disorders, Depression, and Paranoia), one averaged approximately 1.5 standard deviations above the mean (Somatic Complaints), and two (Alcohol and Drug Problems) were on average one standard deviation above the mean. A cursory clinical interpretation of scores in this range might resemble the following: *The person is expressing self-doubt and pessimism (Depression) as well as concern about her health functioning (Somatic Complaints). She presents as skeptical and wary in her relationships (Paranoia). Her responses also reflect some moderate substance use from which she may have experienced some adverse consequences, although the substance use may be largely historical in nature and merits further inquiry.* (The Anxiety-Related Disorders elevation cannot be interpreted at the full scale level and would require further investigation of the subscales to determine which is driving the elevations.) It is concerning that most of these elevations in the simulation sample, in conjunction with the unremarkable elevations on the validity scales, could be interpreted by a clinician as consistent with problems stemming from or related to an undiagnosed, untreated ADHD diagnosis.

Although the findings of this study suggest the PAI demonstrates utility for identifying individuals malingering ADHD symptomatology, the data must be considered within the limitations of the study. The most notable limitation concerns the use of archival comparison data for which the motivation of the participant was unknown. Individuals presenting for an assessment evaluation for ADHD may consciously or unconsciously attempt to portray to the evaluator that they are experiencing significant problems with confused thoughts (e.g., attentional issues) and heightened activity levels and impulsivity (e.g., hyperactivity concerns). Specific to the present study, it is possible that individuals included in the archival comparison group exaggerated or even malingered ADHD symptoms for secondary gain (e.g., academic accommodations, stimulant medication prescription). Although diagnoses were made within the context of an entire test battery and an ADHD diagnosis was not made without multiple sources of supporting data, it is possible that individuals within this group successfully feigned ADHD without detection by the diagnosing clinician. Additionally, results obtained in this sample were impacted to some extent by criterion contamination since the PAI results would have been considered in addition to results from the entire assessment battery when determining the ADHD diagnosis. However, the PAI and CAARS results in the archival sample would not have been utilized in isolation to render a diagnosis. The entire assessment battery in this sample would also have included clinical interview data from a semi-structured interview designed to assess ADHD symptomatology as well as a thorough clinical intake, records review (where available), a computerized task designed to assess executive functioning (i.e., the Continuous Performance Test; Conners, 1995/2000), cognitive testing, other measures of executive functioning (e.g., the Delis Kaplan Executive Function System; Delis, Kaplan, & Kramer, 2001), and in some cases a performance validity task (e.g., the TOMM). Performance validity tests were administered in only a select few of the evaluations for the archival comparison group. Therefore, we were unable to investigate participants' performance on these measures in our data. Despite rendering ADHD diagnoses in the archival comparison sample from a thorough test battery, it is certainly still possible that participants in the archival sample were able to successfully exaggerate ADHD symptomatology without detection.

Separately, however, PAI results from the clinical comparison sample in this study were not used to confirm diagnoses, and participants in that sample would have had no known reason to distort their responses. Moreover, results for the clinical and archival comparison samples were remarkably similar on the validity scales. An additional limitation of the study design is that we cannot guarantee that participants in the clinical comparison sample were not feigning. This would seem like much less of a concern for this particular group, however, given the lack of any incentive for doing so. Importantly, the clinical sample was comprised of data from 22 participants. In comparison with larger samples, decreased confidence is inherent to interpretations gleaned from

results obtained from small samples such as this. It would be important to attempt to replicate this study with a larger sample of college-age individuals with verified ADHD diagnoses.

A notable limitation of previous research includes the large standard deviations between simulating and clinical groups (Frazier et al., 2008; Quinn, 2003) limiting the clinical utility of certain scales in differentiating genuine from simulated psychopathology. The data in this study also produced larger than average standard deviations (i.e., greater than 10*T*) on certain scales for certain groups. For example, the simulation sample obtained an average NIM score that was two standard deviations above the mean, and the corresponding standard deviation for that scale in this group was also approximately two standard deviations (i.e., approximately 20*T*). Thus, some participants who successfully simulated ADHD according to the present study criteria obtained NIM scores that were average (i.e., approximately 50*T*) and would present in ranges obtained in the clinical and archival samples. In contrast, the variance associated with NIM scores in the simulation group reflected that many participants also obtained higher elevations (i.e., approximately 90*T*) that would ordinarily trigger concerns about symptom exaggeration. Although in that scenario, without elevations on MAL and RDF, clinicians may interpret the elevated NIM score as largely due to psychological distress rather than an overt attempt to exaggerate symptomatology. Consequently, the overlap of standard deviations associated with mean scale elevations across the samples decreases confidence in the particular scale elevations obtained in this study and necessitates further investigation into this topic.

Future research should explore the utility of the PAI in symptom validity detection with larger samples, specifically with clinical samples of individuals with verified ADHD diagnoses as well as ADHD treatment seeking populations. Further, researchers may consider how the performance of symptom validity indicators included in other popular measures of personality and psychopathology (e.g., the MMPI-2-RF; Tellegen & Ben-Porath, 2011) compare with the PAI.

It is also important to consider that at the time the data were collected, the *DSM–5* had not yet been published. Given that the revised *DSM–5* ADHD diagnostic criteria are considered to be more lax than the *DSM–IV* criteria (e.g., by increasing the age of childhood symptoms onset symptoms to 12 from the previous benchmark of 7), particularly for diagnosing adults, it is unclear to what extent our results might have been different if diagnoses for our clinical and archival comparison groups were made under the new diagnostic criteria.

Finally, the retrospective nature of an archival study design limits the amount and type of data that researchers can gather. In this study, we were unable to obtain the item-level data for the archival comparison sample, making it impossible to investigate the Musso ADHD symptom validity scale (Musso, 2013). The Musso scale is relatively new and understudied; it is unclear how this formula may perform with various populations. As previously discussed, the utility of this scale is questionable given its high correlation with NIM.

These limitations notwithstanding, the present study contributes to our understanding of feigned ADHD symptomatology on a commonly administered, broad band self-report measure of personality and psychopathology. Given the rise in the number of college-aged individuals seeking an ADHD diagnosis, future research should focus on identifying the best assessment methods for identifying feigned symptomatology in this population. We hope that these results will stimulate further research on this important topic.

References

American Council on Education. (1995). *Getting ready for college: Advising high school students with learning disabilities*. Washington, DC: Heath Resource Centre.

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

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Pankratz, V. S., Weaver, A. L., Weber, K. J., . . . Jacobsen, S. J. (2002). How common is attention-deficit/hyperactivity disorder? Incidence in a population-based birth cohort in Rochester, Minn. *Archives of Pediatrics & Adolescent Medicine, 156,* 217–224. http://dx.doi.org/10.1001/archpedi.156.3.217

Barkley, R. A., & Murphy, K. R. (2005). *Attention-Deficit Hyperactivity Disorder: A Clinical Workbook* (2nd ed.). New York, NY: The Guilford Press.

Brown, T. E. (1996). *Brown ADD Scales*. San Antonio, TX: The Psychological Corporation, Harcourt Brace & Company.

Conners, C. K. (1995). *Conners' continuous performance test*. Toronto, Canada: Multi-Health Systems.

Conners, C. K. (2000). *Conners' continuous performance test-II (CPT-II) computer program for windows technical guide and software manual*. Toronto, Canada: Multi-Health Systems Inc.

Conners, C. K., Erhardt, D., & Sparrow, E. (1998). *Conners' adult ADHD rating scales (CAARS): ADHD across the life span*. Toronto, Canada: Multi-Health Systems.

Conners, C. K., Erhardt, D., & Sparrow, E. (1999). *Conners' adult ADHD rating scales (CAARS) technical manual*. Toronto, Canada: Multi-Health Systems.

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

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis Kaplan executive function system: Examiner's manual*. San Antonio, TX: Psychological Corporation.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics, 44,* 837–845.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD rating scale IV: Checklists, norms, and clinical interpretation*. New York, NY: Guilford Press.

Edens, J. F., Poythress, N. G., & Watkins-Clay, M. M. (2007). Detection of malingering in psychiatric unit and general population prison inmates: A comparison of the PAI, SIMS, and SIRS. *Journal of Personality Assessment, 88,* 33–42. http://dx.doi.org/10.1080/00223890709336832

Fisher, A. B., & Watkins, M. W. (2008). ADHD rating scales' susceptibility to faking in a college student sample. *Journal of Postsecondary Education and Disability, 20,* 81–92.

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

Glutting, J. J., Sheslow, D., & Adams, W. (2002). *CARE: College ADHD response evaluation manual*. Wilmington, DE: Wide Range.

Hagar, K. S., & Goldstein, S. (2001). Case study: Diagnosing adult ADHD: Symptoms versus impairment. *The ADHD Report, 9,* 11–15. http://dx.doi.org/10.1521/adhd.9.3.11.19073

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. (2004). An investigation of reported symptoms of ADHD in a university population. *The ADHD Report, 12,* 8–11. http://dx.doi.org/10.1521/adhd.2006.14.4.1

Harrison, A. G. (2006). Adults faking ADHD: You must be kidding! *The ADHD Report, 14,* 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

Hawes, S. W., & Boccaccini, M. T. (2009). Detection of overreporting of psychopathology on the Personality Assessment Inventory: A meta-analytic review. *Psychological Assessment, 21,* 112–124. http://dx.doi.org/10.1037/a0015036

Jasinski, L. J., Harp, J. P., Berry, D. T., Shandera-Ochsner, A. L., Mason, L. H., & Ranseen, J. D. (2011). Using symptom validity tests to detect malingered ADHD in college students. *The Clinical Neuropsychologist, 25,* 1415–1428. http://dx.doi.org/10.1080/13854046.2011.630024

McCabe, S. E., Knight, J. R., Teter, C. J., & Wechsler, H. (2005). Non-medical use of prescription stimulants among U.S. college students: Prevalence and correlates from a national survey. *Addiction, 100,* 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

McGuire, J. M. (1998). Education accommodations: A university administrator's view. In M. Gordon & S. Keiser (Eds.), *Accommodations in higher education under the Americans with Disabilities ACT: A non-nonsense guide for clinicians, educators, administrators, and lawyers* (pp. 20–45). DeWitt, NY: GSI Publications.

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

Morey, L. C. (1996). *An interpretive guide to the Personality Assessment Inventory (PAI)*. Odessa, FL: Psychological Assessment Resources.

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

Morey, L. C. (2007b). *Personality Assessment Inventory-Adolescent professional manual*. Lutz, FL: Psychological Assessment Resources.

Morey, L. C., & Lanier, V. W. (1998). Operating characteristics of six response distortion indicators for the personality assessment inventory. *Assessment, 5,* 203–214. http://dx.doi.org/10.1177/107319119800500301

Murphy, K., & Barkley, R. A. (1995). Preliminary normative data on *DSM–IV* criteria for adults. *The ADHD Report, 3,* 6–7, 16.

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

Musso, M. W. (2013). *Detecting malingered ADHD using the Personality Assessment Inventory: An exploratory analysis in college students* (Unpublished dissertation). Louisiana State University, Baton Rouge, Louisiana.

Musso, M., Hill, B., Barker, A., Pella, R., Jones, G., Proto, D., & Gouvier, W. (2011, November). *Thought disorder subscale of the Personality Assessment Inventory as a measure of self-reported ADHD symptoms in college students*. Poster presented at the 31st Annual Meeting of the National Academy of Neuropsychology in Marco Island, FL.

Quinn, C. A. (2003). Detection of malingering in assessment of adult ADHD. *Archives of Clinical Neuropsychology, 18,* 379–395. http://dx.doi.org/10.1093/arclin/18.4.379

Reidy, T. J., Sorensen, J. R., & Davidson, M. (2015). Testing the predictive validity of the Personality Assessment Inventory (PAI) in relation to inmate misconduct and violence. *Psychological Assessment, 28,* 871–884. http://dx.doi.org/10.1037/pas0000224

Reynolds, W. M. (1998). *Adolescent Psychopathology Scale*. Odessa, FL: Psychological Assessment Resources.

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. http://dx.doi.org/10.1080/00223891.2013.821071

Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *The Structured Interview of Reported Symptoms (SIRS) professional manual*. Odessa, FL: Psychological Assessment Resources.

Rogers, R., Sewell, K. W., Morey, L. C., & Ulstad, 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

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

Sandford, J. A., & Turner, A. (1995). *Intermediate visual and auditory Continuous Performance Test interpretation manual*. Richmond, VA: Braintrain.

Slick, D., Hopp, G., Strauss, E., & Thompson, G. B. (1997). *VSVT: Victoria Symptom Validity Test, Version 1.0, professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Solanto, M. V., Wasserstein, J., Marks, D. J., & Mitchell, K. J. (2012). Diagnosis of ADHD in adults: What is the appropriate *DSM–5* symptom threshold for hyperactivity-impulsivity? *Journal of Attention Disorders, 16,* 631–634. http://dx.doi.org/10.1177/1087054711416910

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

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson.

Tellegen, A., & Ben-Porath, Y. S. (2008). *The Minnesota Multiphasic Personality Inventory-2 restructured form: Technical manual*. Minneapolis, MN: University of Minnesota Press.

Tellegen, A., & Ben-Porath, Y. S. (2011). *MMPI–2–RF (Minnesota Multiphasic Personality Inventory–2 Restructured Form): Technical manual*. Minneapolis, MN: University of Minnesota Press.

Tombaugh, T. (1997). *Test of memory malingering*. Toronto, Canada: Multi-Health Systems.

Ward, M. F., Wender, P. H., & Reimherr, F. W. (1993). The Wender Utah Rating Scale: An aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. *The American Journal of Psychiatry, 150,* 885–890.

Weyandt, L. L., & DuPaul, G. (2006). ADHD in college students. *Journal of Attention Disorders, 10,* 9–19. http://dx.doi.org/10.1177/1087054705286061

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

Received April 28, 2016
Revision received November 5, 2016
Accepted November 30, 2016 ■