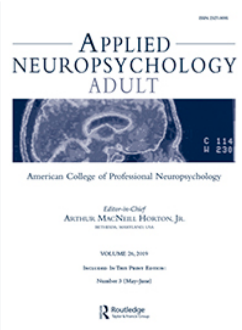

Applied Neuropsychology: Adult

ISSN: 2327-9095 (Print) 2327-9109 (Online) Journal homepage: www.tandfonline.com/journals/hapn21

Examining a novel performance validity task for the detection of feigned attentional problems

Leslie C. Morey

To cite this article: Leslie C. Morey (2019) Examining a novel performance validity task for the detection of feigned attentional problems, Applied Neuropsychology: Adult, 26:3, 255-267, DOI: 10.1080/23279095.2017.1409749

To link to this article: https://doi.org/10.1080/23279095.2017.1409749

Published online: 18 Dec 2017.

APPLIED NEUROPSYCHOLOGY: ADULT
2019, VOL. 26, NO. 3, 255–267
https://doi.org/10.1080/23279095.2017.1409749

Examining a novel performance validity task for the detection of feigned attentional problems

Leslie C. Morey

Department of Psychology, Texas A&M University, College Station, Texas, USA

ABSTRACT

Attention Deficit Hyperactivity Disorder (ADHD) is a diagnosis of particular significance for college students, as when present it can significantly compromise academic achievement. However, because high-stakes decisions may be riding on the outcome of an ADHD evaluation (such as access to educational accommodations or stimulant medication), the diagnosis is vulnerable to exaggeration or feigning of symptoms or impairment. This study evaluates a novel procedure, the Tests of Attentional Distraction (TOAD), which is a computer-based performance validity measure involving a low difficulty continuous performance task that directly and obviously targets attentional function. A group of college student simulators ($n = 115$) feigning ADHD were compared to 32 individuals diagnosed with ADHD and 221 control participants on the TOAD, as well as on symptom validity indicators from the Personality Assessment Inventory (PAI). Moderate to large effects differentiating the feigning group from control participants, both ADHD and non-ADHD, were observed for both the TOAD and PAI indicators. Incremental validity analyses indicated that the two approaches to validity assessment contributed independently to the detection of suspect responding.

KEYWORDS
Attention Deficit Disorder; feigning; performance validity; symptom validity

Attention Deficit Hyperactivity Disorder (ADHD) is a diagnosis with an estimated prevalence in college students of between 2% and 8% (DuPaul, Weyandt, O'Dell, & Varejao, 2009) that when present can significantly compromise academic achievement (Frazier, Youngstrom, Glutting, & Watkins, 2007). As such, requests for evaluation, intervention, or accommodation of this condition appear to be steadily increasing (Pella, Hill, Shelton, Elliott, & Gouvier, 2011; Weyandt & DuPaul, 2008; Wolf, 2001). However, accurate diagnosis in this setting poses a number of formidable challenges. First, the college setting demands a level of self-sufficiency and cognitive aptitude that typically exceeds the student's prior educational experiences, and as such many students present with complaints of attentional difficulties but have no prior diagnosis of ADHD. Second, important information about clinical history regarding the presence of ADHD symptoms during childhood must often rely exclusively on the retrospective report of the students themselves, in contrast to assessments conducted in primary and secondary schools where records of previous testing and symptoms are often available, and it may be difficult for the college-aged examinee to objectively evaluate the abnormality of any feature or associated impairment experienced in childhood. Third, establishing the presence of current ADHD symptoms relies heavily upon client self-report, as objective cognitive measures of attention tend to be inconsistently related to the diagnosis (Schoechlin & Engel, 2005; Weyandt & DuPaul, 2006). However, high-stakes decisions may be riding on the outcome of an ADHD evaluation such as access to educational accommodations (such as extra time on important examinations) or stimulant medication (such as stimulants with high potential for abuse). Because of the possibility of secondary gain in such situations, self-report assessment may be vulnerable to exaggeration or feigning of symptoms or impairment (Musso & Gouvier, 2014). Indeed, estimated rates for feigned or grossly exaggerated symptoms in college students presenting for clinical evaluation of ADHD have ranged from 22% to 48% (Marshall et al., 2010; Sullivan, May, & Galbally, 2007).

Numerous studies have demonstrated that it is relatively simple to feign symptoms of ADHD on self-report checklists, particularly following a cursory review of the DSM criteria for the disorder. In one of the earliest studies on this topic, Quinn (2003) found that simulated malingerers and a genuine ADHD group could not be differentiated by the ADHD Behavior Checklist (Murphy & Barkley, 1996). Jachimowicz and Geiselman (2004) found that most students were able falsify a

CONTACT Leslie C. Morey morey@tamu.edu Department of Psychology, Texas A&M University, College Station, TX 77843–4235, USA.

positive ADHD diagnosis on four different measures: 75% of students taking the ADHD Rating Scale (DuPaul, Power, Anastopoulos, & Reid, 1998), 95% of students taking the Brown (1996) Adult ADHD Scale, 90% of students taking the Conners Adult ADHD Rating Scale (Conners, Erhardt, & Sparrow, 1998), and 65% of students taking the Wender Utah Rating Scale (Ward, Wender, & Reimherr, 1993). Fisher and Watkins (2008) reported that 93% of students who completed the College ADHD Response Evaluation (CARE; Glutting, Monaghan, Adams, & Sheslow, 2002) were able to fake the appropriate number and pattern of ADHD symptoms need to test positive for the diagnosis. However, it should be noted that many of these symptom checklists involve meeting a requisite number of symptoms to receive the diagnosis, rather than being derived from scores based upon a normative sample. Because the symptoms of ADHD reflect relatively common experiences in the general population (Goldstein, 2006; Harrison, 2004), achieving significantly elevated scores on norm-based instruments may be somewhat more challenging. For example, Harrison, Edwards, and Parker (2007) and Rios and Morey (2013) both found that only roughly 45% of feigners achieved scores on the Conners Adult ADHD Rating Scale (CAARS: Conners et al., 1998) that were two or more standard deviations above age means.

Because of the susceptibility of self-report to such distortion, many evaluators use cognitive and neuropsychological testing as an adjunct to the diagnostic process, typically to evaluate the impact of the disorder on objectively measured cognitive performance. A variety of studies have identified ADHD-related deficits in cognitive test performance that are independent of intelligence or education (Boonstra, Oosterlaan, Sergeant, & Buitelaar, 2005; Lange et al., 2014; Schoechlin & Engel, 2005; Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005), particularly in tasks tapping attention, behavioral inhibition/impulsivity, and cognitive flexibility. However, the results of such studies are often inconsistent, and such measures tend to have sensitivity and specificity that is too limited to serve as a reliable diagnostic marker of the disorder (Gualtieri & Johnson, 2005).

One of the more promising measures in this area involves the use of continuous performance tests (CPTs) that assess sustained attention and response inhibition that require vigilance for an infrequently occurring stimulus. The basic metrics derived from CPTs involve two types of errors in performance: omission errors, which reflect missed responses presumably arising from inattention, and commission errors, which are "false positive" responses that presumably reflect impulsivity. CPT performance tends to demonstrate relatively consistent differences between ADHD samples and control samples, with heightened rates of both omission and commission errors (Huang-Pollock, Karalunas, Tam, & Moore, 2012; Losier, McGrath, & Klein, 1996), although the correlation between CPT performance and self-reported ADHD symptoms tends to be modest (e.g., Hopwood & Morey, 2008). However, other CPT-metrics, such as mean reaction time, or within-subject variability, may also hold some promise as diagnostic markers, and because these variables are not transparent to the examinee, the CPT may be of particular use in cases where there is suspicion of malingering. For example, Quinn (2003) reported that a CPT impairment index could distinguish simulated malingerers from genuine ADHD patients, while self-report could not.

However, there are also consistent data demonstrating that individuals can feign poor performance on these cognitive measures. For example, Harrison et al. (2007) found that a feigning group performed much more poorly on measures of thinking speed than normal controls, and even more poorly than ADHD patients. Other investigators have similarly found that impaired performance on attentional measures can be feigned (Henry, 2005; Leark, Dixon, Hoffman, & Huynh, 2002). Booksh, Pella, Singh, and Drew Gouvier (2010) found that simulated malingerers were able to provide impaired scores on the CPT relative to normal controls across an array of CPT metrics; although Booksh et al. found that the feigning group performed significantly worse than ADHD patients, other studies (Suhr, Sullivan, & Rodriguez, 2011) obtained results indicating that feigned and genuine ADHD could not be distinguished with the CPT. Because of this susceptibility to suspect effort, investigators have proposed certain combinations of CPT metrics for use as embedded validity indicators (e.g., Erdodi, Pelletier, & Roth, 2016; Erdodi, Roth, Kirsch, Lajiness-O'Neill, & Medoff, 2014). As such, there is reason to be concerned that the objective nature of cognitive performance measures renders them no less susceptible to malingering efforts than are self-report measures.

Given such results, the field has begun to evaluate the utility of specific tests or procedures designed to raise suspicion regarding malingering in contexts where secondary gain may be an issue. One approach has involved the use of "validity scales" or markers that are included on broadband self-report measures such as the Minnesota Minnesota Multiphasic Personality Inventory–2 (MMPI–2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) or the Personality Assessment Inventory (PAI: Morey, 1991), although these measures tend to be most useful in evaluating the extent to which an individual is exaggerating psychiatric (rather than cognitive) symptoms. For example, Harp, Jasinski, Shandera-Ochsner, Mason, and Berry (2011) examined the validity scales from the MMPI-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008) and found that the Fp-r validity scale, a measure of infrequent psychopathology responses, demonstrated the greatest sensitivity (sensitivity of .636 to feigning, specificity of .900 to genuine ADHD) among the MMPI-2-RF validity scales in identifying feigned ADHD, although it was necessary to employ a cutting score (T = 77) lower than that recommended in the MMPI-2-RF manual to achieve this sensitivity. A different MMPI-2 scale, the Response Bias Scale (RBS: Gervais, Ben-Porath, Wygant & Green, 2007) has demonstrated some promise in detecting cognitive response bias, although this scale has not been investigated with respect to feigned attentional problems.

The validity indicators of the PAI (Morey, 1991, 2007a) have also shown promise in identifying feigned ADHD. Rios and Morey (2013) found that college students completing the Adolescent version of the PAI (PAI-A; Morey, 2007b) under ADHD feigning instructions could be discriminated from genuine ADHD patients on three commonly used PAI indicators of negative distortion, the Negative Impression (NIM) scale, the Malingering Index (MAL; Morey, 1996) and the Rogers Discriminant Function (RDF; Rogers, Sewell, Morey, & Ustad, 1996). Of these indicators, the RDF demonstrated the largest differentiation between the simulation and clinical comparison groups, with a Cohen's d effect size of 2.26. Musso, Hill, Barker, Pella, and Gouvier (2016) also found that the RDF demonstrated the best combination of sensitivity and specificity (sensitivity of .66 to feigning, specificity of .79 to genuine ADHD) of various PAI indicators of response distortion. Finally, Smith, Cox, Mowle, & Edens (2017) replicated these results, finding that RDF demonstrated the largest effect sizes (*d* effect sizes of 1.36 and 1.71) in distinguishing feigned ADHD from two genuine ADHD groups, although large effects were also noted for the PAI NIM and Infrequency (INF) scales.

Although the aforementioned self-report validity indicators have demonstrated utility in evaluating possible malingering, as noted previously they tend to focus on exaggeration of psychiatric symptoms, and their utility for assessing feigned cognitive impairment is less clear (Haggerty, Frazier, Busch, & Naugle, 2007; Zakzanis, Gammada, & Jeffay, 2012). Within the field of neuropsychology, a number of procedures (typically described as "performance validity" tests) have been developed to identify the degree to which test performance accurately reflects examinee ability, which are often used to identify suspect effort or malingering on performance-based cognitive tests. It has been estimated that roughly one-half of the variance in scores from a neuropsychological test battery is explained by effort and examinee cooperation (Green, Rohling, Lees-Haley, & Allen, 2001). Furthermore, a survey of clinical neuropsychologists by Mittenberg, Patton, Canyock, and Condit (2002) reported that malingering or symptom exaggeration is estimated to occur in about 39% of mild head injury cases, 30% of disability assessments, and 29% of personal injury cases. As a result, use of some form of validity testing, including symptom and/or performance validity indicators, has become the norm (Sharland & Gfeller, 2007) in cases where secondary gain may be a factor motivating client performance.

Among the most commonly used free-standing performance validity procedures used by clinicians to help establish the validity of observed impairment are tests such as the Test of Memory Malingering (TOMM; Tombaugh, 1996), the Validity Indicator Profile (VIP; Frederick & Crosby, 2000), the Word Memory Test (WMT; Green, Lees-Haley, & Allen, 2003), Victoria Symptom Validity Test (Slick, Hopp, Strauss, & Spellacy, 1996), and the Rey Fifteen Item Test (Frederick, 2003). Although there is supportive evidence of the potential utility of such measures, the large majority of this research has been done in the context of evaluating personal injury, often involving litigation (Bianchini, Mathias, & Greve, 2001; Iverson, 2006). Furthermore, while some individuals may demonstrate a pattern of exaggerated impairment across a wide array of cognitive functions (Frazier et al, 2007), symptom exaggeration is often more specific to the type of symptoms or condition being simulated (Tan, Slick, Strauss, & Hultsch, 2002). Because most of the aforementioned performance validity procedures were developed to detect feigned memory problems, it is not clear whether this approach is optimal for the evaluation of feigned attentional problems such as might be found in ADHD. Nonetheless, preliminary findings suggest that these performance validity measures might play an important role in ADHD assessment. For example, Sullivan et al. (2007) found that 47.6% of individuals seeking a diagnosis of ADHD demonstrated sub-optimal effort on at least one of the WMT performance validity markers. Booksh (2005) found that the WMT distinguished feigning from normal participants demonstrating a large effect (*d* effect size of 1.60) by Cohen's (1992) conventions, but the ability to distinguish feigned from genuine ADHD could not be evaluated in that sample. Similar results were reported by Frazier, Frazier, Busch, Kerwood, and Demaree (2008), who found that the VIP and VSVT were reasonably accurate in differentiating feigned ADHD from normal controls, but as with Booksh no clinical ADHD comparison sample was included in the Frazier et al. study. However, studies involving feigned vs. genuine ADHD comparisons, such as Marshall et al. (2010) and Sollman, Ranseen, and Berry (2010), support the contention these procedures initially designed to detect memory problems may also be of use in identifying feigned ADHD presentations. However, it remains to be determined if a validity procedure that more directly and obviously targets attentional function might provide a useful addition to these tools for assessing symptom exaggeration.

The goal of the present study is to provide a description and initial validation of an attention-based performance validity measure, the Tests of Attentional Distraction (TOAD: Morey, 2016), as a potential indicator of feigned ADHD. An individual who might be interested in malingering or exaggerating attention problems can easily become familiar with the diagnostic features of ADHD, and can also become acquainted with techniques used to evaluate this disorder, including self-report questionnaires as well as performance measures such as the CPT procedure. The TOAD was designed to mimic a CPT procedure in that it requires sustained attention to the presence or absence of a particular stimulus. However, while an actual CPT provides an attentional challenge that focuses upon response inhibition requiring vigilance for an infrequently occurring stimulus, the TOAD conversely requires response performance in the presence of a stimulus that occurs 50% of the time; a task that is far less demanding of attentional resources. Because of the resemblance of the TOAD to an actual CPT, it is anticipated that individuals motivated to perform poorly on the CPT will show similar decrements in TOAD performance; a pattern not expected in those with genuine attention problems.

Rogers (1997) notes that malingerers will tend to search for cues within the evaluation environment as a guide for how they should respond in order to best simulate impairments associated with the condition. The TOAD has a number of features that were designed to provide such cues to individuals motivated to exaggerate or malinger attentional deficits, cues that would be expected to lead to decreased performance. First, the procedure is clearly labeled as a measure of "attentional distraction," with the name of the procedure prominently displayed to the participant on the initial display screen, and thus it should be obvious to the respondent that poor performance on this measure should be a sign of attention problems and distractibility. Second, the TOAD, like many other performance validity measures but unlike an actual CPT, provides clear feedback about the accuracy of each response. This makes the test similar to a Forced Choice paradigm where participants receive feedback about the accuracy of their choices. Because these tasks tend to be relatively simple, the majority of feedback received by the examinee is that their response is "correct," providing potential feigners with the suggestion that their performance is too good to reflect attentional impairment. To further support this tendency, on the corner of each stimulus screen the TOAD displays ongoing progress bars that reflect poorly disguised tracking of number of correct and incorrect responses. However, unknown to the examinee, the scaling of these progress bars are manipulated such that the bar reflecting the number of correct responses is expanded and that for the number of incorrect responses is compressed, a manipulation which becomes stronger as the tasks allegedly become more difficult.

Thus, this study has a variety of aims. First, it seeks to add to the existing literature that examines the extent to which self-report and performance (i.e., CPT) measures of attention problems can be manipulated by simulated malingerers. Second, it expands upon prior work examining the utility of selected PAI-based self-report indicators of response distortion for detecting feigning of ADHD. Finally, it examines a novel performance-based measure, the TOAD, as a potential indicator of invalid performance, and explores whether the use of such a performance-based measure can increment the diagnostic utility of the aforementioned self-report indicators of feigning.

Method

Participants

Participants in the study included 368 undergraduate students selected from an introductory psychology participant pool. The mean age of these students was 18.9 years ($SD = 1.38$ years); the majority were Caucasian (55.2%; 22.3% were Hispanic, 7.6% Asian-American, 6.3% African-American) and slightly more than half were women (56.3%). All participants were initially queried about any history or presence of diagnosis of ADHD/ADD; 32 participants (8.7%) answered this question affirmatively, with 29 indicating that they had received medication at some point for ADHD, 27 indicating that the diagnosis was current, and 21 noting that they held a current prescription for ADHD medication. This group of 32 participants constituted the ADHD comparison group for this study. The remaining 336 participants were randomly assigned (using a 67% to 33% randomization) to either a standard condition group ($n = 221$) or a feigning group ($n = 115$).

Instruments

Conners Adult Attention Rating Scale-Self-Report Short Version (CAARS-SS; Conners et al., 1998). The CAARS-SS is a self-report instrument that is used to identify ADHD symptoms of attention deficit, hyperactivity, and impulsivity with respect to criteria; the Short Version used in this study is highly correlated (Pearson correlations =.92–.98; Conners et al. 1998) with the full version of this measure. The CAARS-S includes a variety of subscales measuring different aspects of phenomena related to ADHD; in this study, the primary score of interest was the "ADHD Index," a composite score designed to identify individuals in a population that are likely to be diagnosed with ADHD. Although there is some variability in this measure across age and gender (Conners et al., 1998), a raw score of 21 was used (as was done in Rios & Morey, 2013) to infer clinically significant attention problems as this score tends to fall 1.5 to 2.0 standard deviations above the mean across different demographic groups. In this total sample, the internal consistency alpha of the CAARS subscales were: ADHD Index = .872, Inattention/Memory Problems = .822, Hyperactivity/Restlessness = .801, Impulsivity/Emotional Lability = .853, and Problems with Self-Concept = .862.

Personality Assessment Inventory (PAI; Morey, 1991*) scales.* The PAI is a self-administered, multi-scale personality inventory that measures various constructs relevant to clinical diagnosis and decision making; it is widely used in mental health, forensic/correctional, screening, and training contexts (e.g., Archer, Buffington-Vollum, Stredy, & Handel, 2006). In this study, three PAI validity scales were investigated with respect to their ability to differentiate simulated and standard instruction profiles. These scales were the Negative Impression (NIM) scale, found in multiple studies (Rios & Morey, 2013; Smith et al., 2017) to distinguish genuine and feigned ADHD; the Infrequency (INF) scale, identified by Smith et al. (2017) as another potentially useful indicator of feigned ADHD; and the Positive Impression (PIM) scale, which has been considered as a potentially useful supplement to NIM scores (Hopwood et al., 2008). In this total sample, the internal consistency alpha of the NIM scale was .795 and PIM was .794; because the INF scale was designed as a measure of random measurement error rather than as a substantive construct, it is not expected to demonstrate internal consistency (Morey, 2007a).

Psychology Experiment Building Language (PEBL: Mueller & Piper, 2014*) Continuous Performance Test.* The PEBL is an open source software system that is used to design and run psychological experiments. It includes a variety of standard applications in its test battery, including a version of the Continuous Performance Test (pCPT). The pCPT procedure is consistent with variants of this task that are widely used in ADHD assessment (e.g., the Connors CPT-II: Conners, 2000); it is a sustained attention task about 14 minutes in length that requires responses to 324 target letters (which vary) as well as non-responses to 36 foil stimuli (an "X"), which are presented with an inter-stimulus interval of either 1, 2, or 4 s. Piper, Mueller, Talebzadeh, and Ki (2016) found that the correlations between the pCPT and Conners' CPT-II on key measures such as omission and commission errors, or variability in reaction time, compared favorably to the test–retest correspondence scores on CPT tests reported previously in the literature. In addition, the within-test correlations among the different indicators were similar across the CPT-II and the pCPT (Piper et al., 2016). In this study, pCPT administrations with rates of combined omission/commission error above 50% were excluded from the analyses as invalid data to eliminate outliers; this excluded 7 participants from the control condition who were unlikely to have understood and/or complied with the task (e.g., Farmer et al., 2017), and 3 participants from the feigning condition, who although perhaps attempting to feign did not produce believable results, as emphasized in the feigning instructions. No participants in the ADHD group obtained pCPT error rates in this range.

Tests of Attentional Distraction (TOAD: Morey, 2016*).* The TOAD is a computer-administered continuous performance task that differs from traditional CPT tasks in a number of ways to maximize its sensitivity to suspect effort on tasks of this nature. First, it requires response *performance* rather than response inhibition, which requires less attentional vigilance. Second, it provides a 50/50 distribution of targets and foils, making correct response much less difficult resulting in a relatively low ceiling for performance. Third, unlike traditional CPT tasks, each trial provides feedback about the accuracy of the trial to provide obvious cues about task performance.

The TOAD, like the CPT, takes roughly 15 minutes to complete. It is comprised of three subtests, each with two sample items, and it is emphasized to the participant that the three subtests involve tasks that are increasingly difficult, although in fact most participants miss few if any items on any of the tasks. The first subtest requires the examinee to press a key whenever an "X" appears as the single target stimulus letters. The second subtest requires the examinee to press a key whenever an "X" appears as the single target stimulus letter, but only if the "X" is located inside a circle. The third subtest presents three letters randomly located on the screen, and the respondent is to press a key whenever an "X" appears as one of the three stimulus letters. All trials for all three subtests have the same interstimulus interval of 1.8 s, with feedback screens presented for 500 ms following each trial. As with standard CPT tasks, key variables of interest involve the rate of omission and commission errors, as well as the standard deviations of reaction times with the three subtests.

Procedures and design

All participants completed all tasks in a computer laboratory in private work cubicles. All tasks and all instructions were presented by computer, with research assistants present primarily to obtain consent and address any questions. After providing informed consent, participants completed a few screening questions designed to determine whether they had a history of ADHD; all such participants completed subsequent tasks in the Standard Instruction conditions. Those participants not indicating attention problems were then randomized by the computer into a non-ADHD Standard Instruction group (roughly two third of the sample) and a Feigning Instruction group (roughly one third of the sample). Standard Instruction participants received the standard set of instructions associated with each instrument (CAARS, PAI, pCPT, TOAD).

Feigning Instruction participants were instructed to feign a diagnosis of ADHD that might lead to them receiving medication or academic accommodations. The importance of successful feigning was stressed to them, and they were warned to be convincing in their presentations so as to be "believable" by the clinician who would be the presumed recipient of the results, using the instructions presented in the following paragraphs.

> In today's study, you will be asked to complete two brief questionnaires and two brief computer performance tasks that all measure different aspects of attention. We would like you to try to complete ALL of these measures as if you were trying to respond like a person who had Attention-Deficit/Hyperactivity Disorder (ADHD); for example, imagine you needed to convince a psychiatrist that you had ADHD so that you could receive medication or academic accommodations, such as additional time on class tests. Your task is to simulate someone with ADHD in order to make sure the psychiatrist will diagnose you as having this disorder. You should simulate this performance on BOTH the computer tasks as well as the questionnaires. But remember, you want them to BELIEVE you. So, complete all of the tasks BELIEVABLY, but in a way that makes you look like you have ADHD. Here are the diagnostic criteria, or symptoms that someone with ADHD may experience:

The Feigning Instruction participants were then presented with the DSM-5 criteria for ADHD *verbatim,* to provide them with some guidance as to how this condition might present. It was emphasized that they would be completing four different types of tasks, and that they were to feign ADHD symptoms on all four tasks.

The four tasks (CAARS questionnaire, PAI questionnaire, pCPT computer task, TOAD computer task) were presented in counterbalanced, randomized order, with the restriction that the questionnaire tasks were always separated by a computer task and vice versa. Upon completing the four tasks, participants completed a brief demographic questionnaire as well as a multiple choice manipulation check item designed to ascertain if participants understood their instructions (i.e., to do their best, or to feign AHDH, or to feign some other condition not actually included in the study). A total of 38 participants (10.1%) failed this manipulation check item and were excluded from all analyses; the majority of these were from the feigning condition and who apparently failed to adequately review the study instructions before beginning the procedures.

Data analyses

The core data analyses involved one-way analyses of variance between study groups (feigning, ADHD, and standard instruction controls) on study outcome measures of interest, followed by Bonferroni-corrected pairwise group comparisons with Cohen's *d* effect sizes used to characterize specific group contrasts. To test for incremental validity of the different validity indicators, analyses of covariance were performed examining the effect of feigning upon performance validity indicators, controlling for self-report validity measures, and vice versa. Receiver Operating Characteristic (ROC) curves were also computed examining the functional relationship of sensitivity and specificity of these measures, with the Area Under the Curve (AUC) calculated as an indicator of the magnitude of these effects.

Results

Initial analyses evaluated the extent to which participants in the feigning condition were able to successfully feign clinically significant ADHD symptoms on the CAARS. As noted earlier, a raw score of 21 was used (as was done in Rios & Morey, 2013) to infer clinically significant attention problems as this score tends to fall 1.5 to 2.0 standard deviations above the mean across different demographic groups. Using this threshold, a total of 48.2% of the feigning participants successfully simulated ADHD on the CAARS, a number similar to the 45% obtained by Rios & Morey.

Preliminary analyses were also performed to evaluate possible order effects within the questionnaire and computer task administrations. Order effects had no statistically significant impact on any of the primary study dependent variables, including the CAARDs ADHD Index raw score (F 1, 301 = 1.77, $p < .19$), PAI NIM T-score (F 1, 301 = 2.68, $p < .11$), pCPT total errors (F 1, 291 = 0.027, $p < .705$) or TOAD total accuracy (F 1,302 = 2.389, $p < .123$). As such, all subsequent analyses were collapsed across administration order.

Statistical comparisons for the three study groups across the ADHD indicators (i.e., CAARS scores, pCPT performance indicators) are presented in Table 1. This table includes the F-test results from a one-way analysis of variance among the three groups (Standard Instruction condition, ADHD comparison group, and Feigning Instruction condition) the results of Bonferroni *post hoc* comparisons among the groups; and the Cohen's d effect sizes for the comparison of the various groups. The results in Table 1 show significant effects for all putative ADHD indicators, with the post-hoc contrasts indicating that the Feigning group scored significantly higher (i.e., more problematic) than the Standard Instruction group on every indicator, with the sole exception of the CAARS Problems with Self-Concept scale, an indicator which is not specifically tied to attention problems. These effects of the Feigning instructions were generally moderate to large by Cohen's (1992) conventions, supporting the results of prior research that individuals can readily feign attentional problems on both self-report as well as performance-based measures of attention deficit. The results indicated that the ADHD comparison group also scored higher (i.e., more problematic) than the Standard Instruction control group on the various CAARS self-report measures, confirming the presence of reported attention problems in this group. However, the ADHD group did not significantly differ from controls on any of the pCPT measures of error or reaction time variability, a result consistent with prior research indicating that performance on cognitive measures of sustained attention are not consistently associated with ADHD.

Comparisons for the three study groups across validity indicators (i.e., TOAD performance indicators, PAI scores, and the CAARS Inconsistency scale) are presented in Table 2. As with Table 1, included are the F-test results from a one-way analysis of variance

Table 1. Comparison of study groups on measures of attentional problems.

	Standard Instruction (S)	ADHD (A)	Feign Instruction (F)	F value	Effect size F vs. A	Effect size F vs. S	Effect size A vs. S
CAARS ADHD Index, raw							
M	11.79	15.90	19.51	45.75**	0.54*	1.17*	0.72*
SD	5.53	5.84	7.63				
CAARS Inattention-Memory, raw							
M	5.37	8.07	9.14	42.53**	0.32	1.13*	0.87*
SD	3.03	3.15	3.62				
CAARS Hyperactivity-Restlessness, raw							
M	6.75	8.67	10.54	37.95**	0.51*	1.11*	0.58*
SD	3.06	3.51	3.78				
CAARS Impulsivity-Emot. Lability, raw							
M	3.50	4.50	6.92	34.06**	0.67*	1.02*	0.34
SD	2.66	3.24	4.03				
CAARS Self-Concept, raw							
M	5.18	6.80	6.36	4.38*	−0.12	0.32*	0.41*
SD	3.81	4.05	3.51				
pCPT Omission Errors							
M	6.34	7.70	15.04	6.75**	0.34	0.47*	0.07
SD	15.52	21.73	21.46				
pCPT Commission Errors							
M	18.23	17.40	21.85	10.29**	0.73*	0.56*	−0.14
SD	6.37	5.59	6.53				
pCPT React Time *SD*							
M	112.90	102.22	157.18	9.85**	0.69*	0.50*	−0.17
SD	70.10	53.56	106.67				

Note. Effect sizes reflecting significant Bonferroni comparisons ($p < .05$) indicated with asterisk. CAARS = Conners Adult Attention Rating Scale.

Table 2. Comparison of study groups on symptom validity measures, full sample.

	Standard Instruction	ADHD	Feign Instruction	*F*	Effect size	Effect size	Effect size	AUC
	(S)	*(A)*	*(F)*	value	*F* vs. *A*	*F* vs. *S*	*A* vs. *S*	S.E.
TOAD Omission Errors								
M	0.34	0.37	2.55	45.75**	0.59*	0.61*	0.04	.662
SD	0.63	0.76	6.58					.039
TOAD Commission Errors								
M	1.89	2.20	7.13	42.53**	0.93*	1.01*	0.15	.723
SD	1.88	2.17	8.49					.038
TOAD Total Errors								
M	0.01	0.02	0.06	37.95**	0.95*	1.02*	0.14	.755
SD	0.01	0.02	0.08					.036
TOAD React Time *SD*								
M	97.30	103.82	133.48	34.06**	0.65*	0.77*	0.20	.708
SD	34.56	31.78	58.96					.036
PAI NIM *T-score*								
M	51.10	53.23	60.46	16.54**	0.50*	0.69*	0.20	.675
SD	9.59	11.76	17.41					.037
PAI INF T-score								
M	52.28	53.43	59.08	15.41**	0.55*	0.69*	0.13	.667
SD	8.21	8.95	11.49					.037
PAI NIM - PIM								
M	2.49	9.73	17.60	16.88**	0.35	0.71*	0.39	.675
SD	16.94	19.95	25.47					.036
CAARS Inconsistency								
M	4.78	5.07	4.55	0.780	−0.26	−0.11	0.15	.453
SD	1.99	1.76	2.16					.039

Note. Effect sizes reflecting significant Bonferroni comparisons ($p < .05$) indicated with asterisk. TOAD = Tests of Attentional Distraction; PAI = Personality Assessment Inventory; NIM = Negative Impression; INF = Infrequency; CAARS = Conners Adult Attention Rating Scale.

among the three groups (Standard Instruction condition, ADHD comparison group, and Feigning Instruction condition) the results of Bonferroni *post hoc* comparisons among the groups; and the Cohen's *d* effect sizes for the comparison of the various groups. Also presented are the Area Under the Curve (AUC) estimates and standard errors from Receiver Operating Characteristic (ROC) curve analyses for the identification of Feigning participants as compared to other participants. These results demonstrate significant effects for all TOAD and PAI validity indicators, with the post-hoc contrasts indicating that the Feigning group scored significantly higher (i.e., more problematic) than both the ADHD group and the Standard Instruction group on every indicator. Furthermore, on all TOAD and PAI validity indicators, no differences were obtained between the ADHD group and the Standard Instruction group, supporting the conclusion that atypical scores on these measures are related to feigning but not to ADHD. The CAARS Inconsistency scale, however, did not differentiate between study groups and was not able to distinguish feigning from standard responding.

Examining the ROC curves could identify potential cutting scores that yielded a minimum of 90% specificity for identifying feigned attentional problems. These cutting scores and the associated sensitivity and specificity values are presented in Table 3. At these cutting scores, relatively few *bona fide* responders would be misidentified as feigning, but they demonstrated moderate sensitivities to feigning at these thresholds.

Given that both the TOAD and the PAI validity indicators appeared to demonstrate utility for identifying feigned attention problems, it is important to determine if these measures are providing independent information about validity. Thus, analyses were conducted to evaluate whether the self-report PAI and performance-based TOAD would demonstrate incremental utility in the identification of feigning. Results indicated that each contributed independently. First, it was determined that the correlations between the methods were quite modest; for example, the TOAD Total Error rate correlated .160 with PAI NIM and .212 with PAI INF, values that were positive and statistically significant, but represent very small effects. Next, analyses were conducting to evaluate Receiver Operating Characteristic curve AUC results for these indicators

Table 3. Sensitivity and specificity of cutting scores on symptom validity indicators for identifying feigning.

Symptom validity indicator	Cutting score	Sensitivity	Specificity
TOAD Omission Errors	≥2	.280	.939
TOAD Commission Errors	≥5	.440	.900
TOAD Total Error rate	≥3.67%	.480	.909
TOAD React Time SD	>136.99.ms	.400	.900
PAI NIM	≥64 T	.295	.905
PAI INF	≥65 T	.333	.931
PAI NIM - PIM	≥27	.308	.901

Note. Effect sizes reflecting significant Bonferroni comparisons ($p < .05$) indicated with asterisk. TOAD = Tests of Attentional Distraction; PAI = Personality Assessment Inventory; NIM = Negative Impression; INF = Infrequency.

Table 4. Comparison of successful and unsuccessful feigners on symptom validity measures.

	Successful Feign (SF)	Unsuccessful Feign (UF)	*F* value	Effect size SF vs. *A*	Effect size SF vs. *S*	Effect size UF vs. *A*	Effect size UF vs. *S*	AUC S.E.
TOAD Omission Errors								
M	3.55	1.51	10.99**	0.65*	0.67*	0.87	0.94	.639
SD	9.00	1.88						.054
TOAD Commission Errors								
M	9.37	4.84	33.25**	1.18*	1.26*	0.65	0.75*	.750
SD	9.95	5.97						.052
TOAD Total Errors								
M	8.61%	4.23%	33.63**	1.14*	1.20*	0.84	0.95*	.755
SD	10.42%	4.33%						.036
TOAD React Time *SD*								
M	133.71	133.24	13.72**	0.58*	0.69*	0.78*	0.92*	.663
SD	71.65	43.25						.055
PAI NIM *T-score*								
M	64.59	56.33	14.36**	0.76*	0.97*	0.22	0.41	.749
SD	18.12	15.83						.046
PAI INF T-score								
M	58.36	59.79	10.41**	0.49	0.63*	0.61*	0.74*	.655
SD	11.06	12.01						.051
PAI NIM - PIM								
M	24.59	10.62	14.95**	0.64*	1.02*	0.04	0.41	.757
SD	26.24	22.93						.044
CAARS Inconsistency								
M	4.28	4.82	0.980	−0.42	−0.25	−0.12	0.02	.422
SD	1.96	2.34						.051

Note. Effect sizes reflecting significant Bonferroni comparisons ($p < .05$) indicated with asterisk.

after controlling for the contribution of the other method. Thus, the TOAD, after partialling out the association with NIM, continued to demonstrate a significant AUC of .711 ($SE = .040$).. Similarly, NIM also continued to demonstrate a significant AUC of .607 ($SE = .042$) after controlling for TOAD error rate. However, the NIM-PIM T-score difference demonstrated a non-significant AUC value of .538 ($SE = .037$) after controlling for the NIM score in isolation. This latter result replicates the previous findings reported by Hopwood, Talbert, Rogers, and Morey (2008) that the difference between NIM and PIM scores adds no predictive variance for the identification of problematic responding, beyond the validity of the base NIM score itself.

Because of the incremental information provided by the self-report NIM score and the performance based TOAD error rate, the utility of the two in conjunction for identifying feigning participants was explored. Using a decision rule based upon the cutting scores presented in Table 3 of NIM T-score ≥ 64 or TOAD total error rate ≥ 3.67%, this combination yielded a sensitivity of 59.2% for identifying participants in the Feigning group, with a specificity of 81.7%. Thus, use of this disjunction rule enhances sensitivity beyond that of the individual procedures observed in Table 3, at the expense of a decrease in specificity. A conjunctive decision rule that required both of these indicators to exceed these cuttings scores demonstrated high specificity (99.6%) with limited sensitivity (18.2%).

Of particular interest is the ability of the validity indicators to detect those participants in the Feigning condition who were able to successfully simulate ADHD symptoms as reflected on the CAARS ADHD Index score. Results for the feigning condition participants, divided into successful and unsuccessful feigners, are provided in Table 4. Comparing Tables 2 and 4 shows that successful ADHD simulators consistently scored higher on the validity measures than did other study participants, while unsuccessful feigners in several instances were not significantly different from control participants. The diagnostic efficiency (as represented by the AUC values) of the self-report symptom validity measures, particularly PAI NIM, were larger for identifying the successful feigning group than for the simulation sample as a whole, although the performance-based TOAD yielded similar results for the Feigning group regardless of the obtained CAARS-ADHD Index score. Use of the TOAD + NIM disjunction rule described above yielded a sensitivity of 71.1% for identifying these successful feigners, with the conjunctive rule demonstrating 23.1% sensitivity, with specificity values unchanged in this subsample. As was the case with the feigning group as a whole, the CAARS Inconsistency scale was not able to differentiate between successful feigners and standard responding.

Discussion

This study examined college students' ability to simulate ADHD on self-report and performance measures of attentional problems, and the efficiency of self-report and a novel performance-based validity test to detect such efforts at feigning. With respect to ADHD simulation, participants in the Feigning condition were able to simulate problematic responding, obtaining higher scores than Standard Instructions control participants on both self-report and cognitive performance indicators of attention problems, with moderate to large effect sizes observed on all measures. The finding that attentional problems can be simulated on such widely-used diagnostic measures underscores the need for validity assessment in this population. As hypothesized, these Feigning participants were significantly different from both Standard Instruction controls as well as from participants with a history of ADHD on both the performance-based (TOAD) and self-report (PAI) validity indicators examined in this study.

In this study, results demonstrated that both the TOAD and the PAI approaches demonstrated some success in the identification of feigned attention problems. Of the various indicators examined, the TOAD Total Error rate demonstrated the greatest success identifying feigning, resulting in the largest effect sizes ($d = 1.02$ and AUC = .755, both representing large effects by Cohen's 1992 conventions) of any of the validity indicators studied. The success of the TOAD Total Error rate as an indicator of questionable effort, like other performance validity measures such as the Rey 15-Item Test, appears to be due to the low ceiling of the procedure, making it easy for examinees to obtain near-perfect scores. For example, 79% of participants in the Standard Instruction condition and 77% of ADHD participants were able to get more than 98% of the TOAD items correct; none of these participants had less than 90% correct, although 20% of participants in the Feigning condition had scores below 90%. Such results support the notion that poor performance on easy tasks that are strongly cued as "attention measures" can be used to identify questionable effort in ADHD assessment.

With respect to the self-report symptom validity measures from the PAI, both the NIM scale and the INF scale demonstrated significant differences between Feigning and other participants, with greater sensitivity in identifying feigning participants who were able to produce clinically elevated scores on the CAARS ADHD Index. The diagnostic efficiency values obtained, particularly among successful feigners, were similar to those reported in other studies of these indicators by Smith et al. (2017) and Rios and Morey (2013), with AUC values in the moderate to large range. Interestingly, the cutting score identified in this study of T = 64 is similar to values reported by Smith et al. (2017) and Rios and Morey for detecting feigned ADHD, in that this NIM cutting score is comparatively low when compared to optimal cut scores obtained in most studies examining feigned psychopathology (Hawes & Boccaccini, 2009). Because NIM was constructed to assess negative distortions in psychopathology, rather than in cognitive disturbance, it does not tend to demonstrate marked elevations when participants are attempting to simulate rather circumscribed cognitive problems, such as memory difficulties (e.g., Armistead-Jehle & Buican, 2012; Brennan et al., 2009). However, in contexts where demands to simulate cognitive problems are prominent and emotional or behavioral problems are less relevant, then use of cutting scores on NIM that are lower may be advisable.

The sensitivity results for the various performance validity measures obtained in this study that correspond to a specificity of 90% are relatively modest, suggesting that they might be of greatest use as a rule-out of true ADHD, perhaps confirming other evidence of feigning but not highly sensitive of efforts to feign this disorder. However, it should be noted that the self-report and performance based validity indicators made largely independent contributions to the identification of feigning. As such, using a disjunctive "or" rule, where suspicions would be raised if either self-report or performance indicators were elevated, resulted in appreciably better sensitivities, while still demonstrating a specificity of over 80%. Such a rule might be considered in situations where a false-negative decision around feigning might be costly, for example, in evaluating an ADHD medication decision in an individual with an extensive history of substance abuse. Conversely, in situations where a false-positive identification of feigning might be costly (e.g., a decision around educational accommodations that might be sorely needed by a student), a conjunctive "and" rule requiring questionable performance on both self-report and performance indicators might be merit. Such a multimodal approach to validity testing might be considered in high-stakes testing situations.

Although this study provides an important first step in evaluating the detecting of feigned ADHD with the TOAD, there is considerable need for additional work in this area. An important limitation of the present study is that the ADHD comparison group involved individuals who self-identified as such, and although they obtained CAARS scores that were significantly higher than those of standard controls, there was no careful evaluation of an ADHD diagnosis in these participants. In addition, there was no assessment of other conditions (e.g., history of head injury or other psychopathology) that might have impacted performance. Another limitation of the current study is that it only studied college students. Although college is a context in which possible secondary gains associated with an ADHD are salient, ADHD diagnoses are often first assigned in childhood, and thus an important extension of the results obtained here would involve examining the efficiency of these validity indicators in younger participants. Such investigations would provide additional information salient to the important issue of detecting feigning, exaggeration, or malingering of ADHD symptoms to achieve various types of secondary gains. Furthermore, aside from the manipulation check employed to exclude participants, there were no other assessments of validity administered to this sample, and as such there was no independent criterion to ensure good effort among the ADHD and Standard Instructions groups, and poor effort among the Feigning Instructions group. Finally, there are limits to the generalizability of analogue-simulation design as used here. For example, performance in these experimental simulations may result in findings that are stronger than those obtained in naturalistic settings; on the other hand, individuals participating in actual clinical assessments often have much larger incentives riding on the outcome of such an evaluation (Rogers, 1997) and, thus, could be motivated to perform particularly poorly relative to simulation participants. It is hoped that future research can expand upon these initial findings to help refine such procedures for addressing the important clinical issue of feigned ADHD, with a particular need for research studying college students undergoing clinical ADHD assessment for academic accommodations or access to stimulant medication.

References

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

Armistead-Jehle, P., & Buican, B. (2012). Evaluation context and Symptom Validity Test performances in a US Military sample. *Archives of Clinical Neuropsychology*, *27*(8), 828–839. doi:10.1093/arclin/acs086

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

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist*, *15*(1), 19–45. doi:10.1076/clin.15.1.19.1907

Booksh, R. L. (2005). Ability of college students to simulate ADHD on objective measures of attention. Retrieved June 23, 2006, from Louisiana State University Web site: http://etd.lsu.edu/docs/available/etd-11162005111135/unrestricted/Booksh_dis.pdf.

Booksh, R. L., Pella, R. D., Singh, A. N., & Drew Gouvier, W. (2010). Ability of college students to simulate ADHD on objective measures of attention. *Journal of Attention Disorders*, *13*(4), 325–338. doi: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*(8), 1097–1108. doi:10.1017/s003329170500499x

Brennan, A. M., Meyer, S., David, E., Pella, R., Hill, B. D., & Gouvier, W. G. (2009). The vulnerability to coaching across measures of effort. *Clinical Neuropsychologist*, *23*, 314–328. doi:10.1080/13854040802054151

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

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI–2: Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. doi:10.1037//0033-2909.112.1.155

Conners, C. (2000). *Conners Continuous Performance Test II*. Tonawanda, NY: Multi-Health Systems, Inc.

Conners, C. K., Erhardt, D., & Sparrow, E. (1998). *Conners' Adult ADHD Rating Scales (CAARS): ADHD across the life span*. Toronto, ON: Multi-Health Systems.

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.

DuPaul, G. J., Weyandt, L. L., O'Dell, S. M., & Varejao, M. (2009). College students with ADHD: Current status and future directions. *Journal of Attention Disorders*, *13*(3), 234–250. doi:10.1177/1087054709340650

Erdodi, L. A., Pelletier, C. L., & Roth, R. M. (2016). Elevations on select Conners' CPT-II scales indicate noncredible responding in adults with traumatic brain injury. *Applied Neuropsychology: Adult*, *1–10*. doi:10.1080/23279095.2016.1232262

Erdodi, L. A., Roth, R. M., Kirsch, N. L., Lajiness-O'Neill, R., & Medoff, B. (2014). Aggregating validity indicators embedded in Conners' CPT-II outperforms individual cutoffs at separating valid from invalid performance in adults with traumatic brain injury. *Archives of Clinical Neuropsychology*, *29*(5), 456–466. doi:10.1093/arclin/acu026

Farmer, C. A., Epstein, J. N., Findling, R. L., Gadow, K. D., Arnold, L. E., Kipp, H., … McNamara, N. K. (2017). Risperidone added to psychostimulant in children with severe aggression and attention-deficit/hyperactivity disorder: lack of effect on attention and short-term memory. *Journal of Child and Adolescent Psychopharmacology*, *27*(2), 117–124. doi:10.1089/cap.2016.0040

Fisher, A., & Watkins, M. (2008). ADHD rating scales' susceptibility to faking in a college student sample. *Journal of Postsecondary Education and Disability*, *20*(2), 81–92.

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

Frazier, T. W., Youngstrom, E. A., Glutting, J. J., & Watkins, M. W. (2007). ADHD and achievement: Meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. *Journal of Learning Disabilities*, *40*(1), 49–65. doi:10.1177/00222194070400010401

Frederick, R. I. (2003). A review of Rey's strategies for detecting malingered neuropsychological impairment. *Journal of Forensic Neuropsychology*, *2*(3–4), 1–25. doi:10.1300/j151v02n03_01

Frederick, R. I., & Crosby, R. D. (2000). Development and validation of the Validity Indicator Profile. *Law and Human Behavior*, *24*(1), 59–82. doi:10.1023/a:1005426803586

Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2007). Development and validation of a Response Bias Scale (RBS) for the MMPI-2. *Assessment*, *14*, 196–208. doi:10.1177/1073191106295861

Glutting, J., Monaghan, M., Adams, W., & Sheslow, D. (2002). Some psychometric properties of a system to measure ADHD among college students: Factor pattern, reliability, and one-year predictive validity. *Measurement and Evaluation in Counseling and Development*, *34*(4), 194–209.

Goldstein, S. (2006). Is ADHD a growth industry? *Journal of Attention Disorders*, *9*, 461–464. doi:10.1177/1087054705283678

Green, P., Lees-Haley, P. R., & Allen III, L. M. (2003). The Word Memory Test and the validity of neuropsychological test scores. *Journal of Forensic Neuropsychology*, *2*(3–4), 97–124. doi:10.1300/j151v02n03_05

Green, P., Rohling, M. L., Lees-Haley, P. R. & Allen, L. M., III. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury*, *15*(12), 1045–1060. doi:10.1080/02699050110088254

Gualtieri, C. T., & Johnson, L. G. (2005). ADHD: Is objective diagnosis possible? *Psychiatry (Edgmont)*, *2*(11), 44.

Haggerty, K. A., Frazier, T. W., Busch, R. M., & Naugle, R. I. (2007). Relationships among Victoria symptom validity test indices and personality assessment inventory validity scales in a large clinical sample. *The Clinical Neuropsychologist*, *21*, 917–928. doi:10.1080/13854040600899724

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. doi: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. doi:10.1521/adhd.12.6.8.55256

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

Hawes, S. W., & Boccaccini, M. T. (2009). Detection of over reporting of psychopathology on the Personality Assessment Inventory: A meta-analytic review. *Psychological Assessment*, *21*(1), 112–124. doi:10.1037/a0015036

Henry, G. (2005). Probable malingering and performance on the Test of Variables of Attention. *Clinical Neuropsychologist*, *19*, 121–129. doi:10.1080/13854040490516604

Hopwood, C. J., & Morey, L. C. (2008). Emotional problems suppress the relation between continuous performance and adult ADHD. *Journal of Psychopathology and Behavioral Assessment*, *30*, 204–210. doi:10.1007/s10862-007-9067-8

Hopwood, C. J., Talbert, C. A., Rogers, R., & Morey, L. C. (2008). Testing the incremental utility of the NIM/PIM differential in detecting simulated PAI profiles. *Journal of Clinical Psychology*, *64*, 338–343. doi:10.1002/jclp.20439

Huang-Pollock, C. L., Karalunas, S. L., Tam, H., & Moore, A. N. (2012). Evaluating vigilance deficits in ADHD: A meta-analysis of CPT performance. *Journal of Abnormal Psychology*, *121*(2), 360–371. doi:10.1037/a0027205

Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology*, *13*(2), 77–90. doi:10.1207/s15324826an1302_3

Jachimowicz, G., & Geiselman, R. E. (2004). Comparison of ease of falsification of attention deficit hyperactivity disorder diagnosis using standard behavioral rating scales. *Cognitive Science Online*, *2*(1), 6–20.

Lange, K. W., Hauser, J., Lange, K. M., Makulska-Gertruda, E., Takano, T., Takeuchi, Y., … Tucha, O. (2014). Utility of cognitive neuropsychological assessment in attention-deficit/hyperactivity disorder. *ADHD Attention deficit and hyperactivity disorders*, *6*(4), 241–248. doi:10.1007/s12402-014-0132-3

Leark, R. A., Dixon, D., Hoffman, T., & Huynh, D. (2002). Fake bad test response bias effects on the Test of Variables of Attention. *Archives of Clinical Neuropsychology*, *17*, 335–342. doi:10.1093/arclin/17.4.335

Losier, B. J., McGrath, P. J., & Klein, R. M. (1996). Error patterns on the continuous performance test in non-medicated and medicated samples of children with and without ADHD: A meta-analytic review. *Journal of Child Psychology and Psychiatry*, *37*(8), 971–987. doi:10.1111/j.1469-7610.1996.tb01494.x

Marshall, P., Schroeder, R., O'Brien, J., Fischer, R., Ries, A., Blesi, B., & Barker, J. (2010). Effectiveness of symptom validity measures in identifying cognitive and behavioral symptom exaggeration in adult attention deficit hyperactivity disorder. *Clinical Neuropsychologist*, *24*(7), 1204–1237. doi:10.1080/13854046.2010.514290

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, *24*(8), 1094–1102. doi:10.1076/jcen.24.8.1094.8379

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

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

Morey, L. C. (2007a). *Personality Assessment Inventory 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. (2016). *Tests of attentional distraction*. College Station, TX: Morey Consulting.

Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, *222*, 250–259. doi:10.1016/j.jneumeth.2013.10.024

Murphy, K., & Barkley, R. A. (1996). Prevalence of DSM-IV symptoms of ADHD in adult licensed drivers: Implications for clinical diagnosis. *Journal of Attention Disorders*, *1*, 147–161. doi:10.1177/108705479600100303

Musso, M. W., & Gouvier, W. D. (2014). "Why is this so hard?" A review of detection of malingered ADHD in college students. *Journal of Attention Disorders*, *18*(3), 186–201. doi:10.1177/1087054712441970

Musso, M. W., Hill, B. D., Barker, A. A., Pella, R. D., & Gouvier, W. D. (2016). Utility of the personality assessment inventory for detecting malingered ADHD in college students. *Journal of Attention Disorders*, *20*(9), 763–774. doi:10.1177/1087054714548031

Pella, R. D., Hill, B. D., Shelton, J. T., Elliott, E., & Gouvier, W. D. (2011). Evaluation of embedded malingering indices in a non-litigating clinical sample using control, clinical, and derived groups. *Archives of Clinical Neuropsychology*, *27*(1), 45–57. doi:10.1093/arclin/acr090

Piper, B., Mueller, S. T., Talebzadeh, S., & Ki, M. J. (2016). Evaluation of the validity of the Psychology Experiment Building Language tests of vigilance, auditory memory, and decision making. *PeerJ*, *4*, e1772. doi:10.7717/peerj.1772

Quinn, C. A. (2003). Detection of malingering in assessment of adult ADHD. *Archives of Clinical Neuropsychology*, *18*(4), 379–395. doi:10.1016/s0887-6177(02)00150-6

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. doi:10.1080/00223891.2013.821071

Rogers, R. (1997). Malingering of posttraumatic disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 398–426). New York, NY: Guilford.

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. doi: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*(6), 727–744. doi:10.1016/j.acn.2005.04.005

Sharland, M. J., & Gfeller, J. D. (2007). A survey of neuropsychologists' beliefs and practices with respect to the assessment of effort. *Archives of Clinical Neuropsychology*, *22*(2), 213–223. doi:10.1016/j.acn.2006.12.004

Slick, D. J., Hopp, G., Strauss, E., & Spellacy, F. J. (1996). Victoria symptom validity test: Efficiency for detecting feigned memory impairment and relationship to neuropsychological tests and MMPI-2 validity scales. *Journal of Clinical and Experimental Neuropsychology*, *18*(6), 911–922. doi:10.1080/01688639608408313

Smith, S. T., Cox, J., Mowle, E. N., & Edens, J. F. (2017). Intentional inattention: Detecting feigned attention-deficit/hyperactivity disorder on the personality assessment inventory. *Psychological Assessment*, *29*(12), 1447–1457. https://doi.org/10.1037/pas0000435

Sollman, M. J., Ranseen, J. D., & Berry, D. T. (2010). Detection of feigned ADHD in college students. *Psychological Assessment*, *22*(2), 325–335. doi:10.1037/a0018857

Suhr, J. A., Sullivan, B. K., & Rodriguez, J. L. (2011). The relationship of noncredible performance to continuous performance test scores in adults referred for attention deficit/hyperactivity disorder evaluation. *Archives of Clinical Neuropsychology*, *26*, 1–7. doi:10.1093/arclin/acq094

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*(3), 189–207. doi:10.1080/09084280701509083

Tan, J. E., Slick, D. J., Strauss, E., & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist*, *16*(4), 495–505. doi:10.1076/clin.16.4.495.13909

Tombaugh, T. N. (1996). *Test of memory malingering: TOMM*. North Tonawanda, NY: 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. *American Journal of Psychiatry*, *150*, 885–889.

Weyandt, L. L., & DuPaul, G. (2006). ADHD in college students. *Journal of Attention Disorders*, *10*(1), 9–19. doi:10.1177/1087054705286061

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

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, *57*(11), 1336–1346. doi:10.1016/j.biopsych.2005.02.006

Wolf, L. E. (2001). College students with ADHD and other hidden disabilities. *Annals of the New York Academy of Sciences*, *931*(1), 385–395. doi:10.1111/j.1749-6632.2001.tb05792.x

Zakzanis, K. K., Gammada, E., & Jeffay, E. (2012). The predictive utility of neuropsychological symptom validity testing as it relates to psychological presentation. *Applied Neuropsychology: Adult*, *19*(2), 98–107. doi:10.1080/09084282.2011.644099