Article

Journal of Attention Disorders
2022, Vol. 26(12) 1622–1630

DOI: 10.1177/10870547221092095
journals.sagepub.com/home/jad
SAGE

Identifying Feigned ADHD in College Students: Comparing the Multidimensional ADHD Rating Scale to Established Validity Measures

Heather E. Potts[1], Lawrence J. Lewandowski[2], and Benjamin J. Lovett[3]

Abstract

Objective: There is increased concern for malingering in ADHD evaluations due to presence of incentives such as accommodations and medications. Although several validity tests are available to classify malingering in non-ADHD populations, there is no test with proven effectiveness to detect feigned ADHD. This study investigated the ability of the MARS Symptom Validity Index 4 (MARS SV-index 4) and two published validity tests (the Word Memory Test [WMT] and Clinical Assessment of Attention Deficit-Adult [CAT-A] Infrequency scale) to detect malingered ADHD. **Method:** The participants consisted of 68 young adults, 34 with ADHD. Participants completed the MARS, CAT-A, and WMT validity measures. **Results:** The MARS SV index-4 demonstrated higher sensitivity rates for simulated malingering (61.8%) at close to optimal specificity (88.2%) compared to two published tests (which had sensitivity <42% at specificity >90%). **Conclusion:** The MARS shows good ability to detect feigned ADHD and appears to be useful for adult ADHD assessments. *(J. of Att. Dis. 2022; 26(12) 1622-1630)*

Keywords

ADHD, malingering, diagnostic validity, assessment

Since the 1990s, there has been increased recognition that ADHD can be diagnosed for the first time in adulthood (Barkley et al., 2008; Davidson, 2008). Research-based guidance for making such diagnoses is now common (e.g., Sibley, 2021), and this guidance typically acknowledges the possibility of overdiagnosis. Indeed, with an increasing number of young adults self-referring for ADHD evaluations on college campuses (Barkley et al., 2008; Harrison, 2004; Weyandt & DuPaul, 2006), some have questioned whether current assessment practices are adequate, or if they instead facilitate many false positive diagnoses (Harrison, 2017; Musso & Gouvier, 2014).

One source of overdiagnosis is symptom exaggeration, and clients being evaluated for ADHD often have substantial incentives to exaggerate. Individuals with ADHD are frequently provided academic accommodations and stimulant medications, which non-ADHD college students also view as beneficial (Benson et al., 2015; Jasinski & Ranseen, 2011; Lewandowski et al., 2014). Unfortunately, some degree of exaggeration appears to occur in approximately 15% to 50% of adult ADHD evaluations (Harrison & Edwards, 2010; Sullivan et al., 2007). Individuals who falsely receive an ADHD diagnosis could be provided with accommodations that result in an unfair academic advantage (Jasinski & Ranseen, 2011), as well as prescribed stimulants that place them at increased medical risk (Benson et al., 2015; Park & Haning, 2016). Thus, there is a clear need to refine our diagnostic practices to accurately differentiate true cases of ADHD from those malingering the disorder (Musso & Gouvier, 2014).

Research indicates that standard ADHD assessments and practices have limited ability to detect malingering. With little preparation, motivated individuals can easily fake an ADHD profile by under-performing on cognitive/neuropsychological measures and over-reporting on rating scales (e.g., Marshall et al., 2010, 2016; Tucha et al., 2015). Although malingerers tend to *over*exaggerate (i.e., they perform worse and report worse symptoms than individuals

[1]Boston Children's Hospital, MA, USA
[2]Syracuse University, NY, USA
[3]Columbia University, New York, NY, USA

Corresponding Author:
Heather E. Potts, Boston Children's Hospital, Outpatient Psychiatry Service, 300 Longwood Avenue, BCH 3222, Boston, MA 02115, USA.
Email: heather.potts@childrens.harvard.edu

with actual ADHD), feigned data often still fall at levels that are comparable to a typical ADHD profile. For example, the Conners' Adult ADHD Rating Scales (Conners et al., 1999) is a self-report measure that includes embedded items that assess for consistency of reported symptoms. Although individuals feigning ADHD do tend to overreport symptoms, self-report measures are unable to reliably discriminate feigned ADHD from the genuine disorder (Cook et al., 2016; Fuermaier et al., 2017). Because the malingering profile is not easily differentiated from ADHD on standard diagnostic measures (e.g., symptom rating scales and continuous performance tests), diagnosticians are at increased risk of rendering false ADHD diagnoses when relying upon these measures (Musso & Gouvier, 2014; Tucha et al., 2015).

Measures of symptom validity (overreporting of psychiatric symptoms) and performance validity (effort on cognitive tests) are recommended to help address overdiagnosis concerns (Musso & Gouvier, 2014; Sagar et al., 2017). Symptom validity tests (SVTs) would be applicable in evaluations in which ADHD is diagnosed on observations or reports of symptoms. Though cognitive testing is not a strictly necessary part of ADHD assessment (Lovett & Nelson, 2018), assessment of ADHD typically can involve testing of intelligence and cognitive processing, at least when college students are evaluated (Nelson et al., 2019). In these evaluations, PVTs should be considered to assess for effort on these cognitive measures.

Researchers and clinicians have considered the issue of symptom over-report as a potential indicator of malingering (e.g., Rogers, 2008). Reporting extremely high levels of ADHD symptoms could, theoretically, be a "red flag" for malingered ADHD. However, since individuals with genuine and severe ADHD would also be expected to endorse elevated symptom levels, this is not recommended as a sole technique to detect malingering (Musso & Gouvier, 2014). A better detection method may be the use of symptom items that individuals with ADHD rarely endorse, yet malingerers often endorse, essentially catching those that do not have experience with the disorder.

The only SVT developed to assess response bias for ADHD comes from the Clinical Assessment of Attention Deficit-Adult (CAT-A) rating scale (Bracken & Boatwright, 2005). The CAT-A includes an "Infrequency scale" consisting of items that individuals with ADHD in the normative sample endorsed at low rates, and although not validated for these purposes, the items were presumed to be adequate tools to assess for response bias. Since publication of the CAT-A, Marshall et al. (2010, 2016) conducted two archival studies finding that the CAT-A Infrequency scale had 57% to 58% sensitivity to malingering at an optimal specificity of above 90%. However, a limitation of archival designs is that researchers rely upon validity tests to identify "suspected" cases of malingering to form the groups that are subsequently used to assess the ability of other validity tests to detect malingering, and there is no assurance that a "suspected" case actually is a "true" case of malingering (Rogers, 2008). Given the inherent limitations of archival research that can only suspect malingering (and cannot prove malingering), additional research appears needed to validate the utility of the CAT-A to detect feigned cases of ADHD.

In addition to a lack of specifically developed symptom validity scales, there is no available PVT designed specifically to detect feigned ADHD. However, because commercially available PVTs demonstrate high detection accuracy for feigning within their intended populations (e.g., feigned memory impairments), researchers have sought to determine whether these measures could also detect feigned ADHD as well. The Word Memory Test (Green, 2003) is used frequently in feigned ADHD research and ADHD evaluations, presumably because there are few options available and there is some research to support its use (e.g., Booksh et al., 2010; Edmundson et al., 2017; Marshall et al., 2010; Suhr et al., 2008). Specifically, one archival study suggested that the WMT demonstrated moderate sensitivity (63%) to detect a group of individuals classified with suspected poor effort during a psychological evaluation (Marshall et al., 2010). The WMT may be fairly sensitive at detecting uncoached, or unprepared malingerers (sensitivity = 57%–74%; Edmundson et al., 2017). However, sensitivity rates dropped significantly when participants were coached on the symptom criteria of the disorder (30%; Edmundson et al., 2017). Savvy (or coached) students feigning ADHD may know that poor memory is not a core ADHD symptom, and so they would have no reason to fail the WMT, which is an obvious test of memory. The different findings between coached and uncoached malingerers suggest that additional research is needed to examine the utility of the WMT as a measure of *coached* feigned ADHD.

In summary, the field of ADHD assessment has lacked a symptom or performance validity test that is specifically designed to detect feigned ADHD and that has been shown to be effective at doing so. However, one measure developed recently to detect feigned ADHD in adults is the Multidimensional ADHD Rating Scale (MARS; Potts et al., 2021). The MARS contains four embedded items describing symptoms that young adults with ADHD endorse at low rates, but that malingerers endorse at much higher rates. The total score across these four items, known as the symptom validity index 4 (SV-index 4), demonstrated good sensitivity for detecting malingering (64.7%) with optimal specificity (91.8%). Potts et al. (2021) recommended that the MARS be used in adult ADHD evaluations in part because of the SV-index 4s ability to detect malingered ADHD.

The purposes of the present study were to confirm the diagnostic efficiency of the four-item MARS Symptom Validity Index 4 (SV-4) to classify simulated malingering, and to compare the classification accuracy of the MARS

SV-4 to that of the Word Memory Test (WMT) and the Clinical Assessment of Attention Deficit-Adult (CAT-A) Infrequency Scale.

Method

Participants

After receiving approval from the university Institutional Review Board, we recruited participants from undergraduate psychology classes and publicly posted flyers. Students who reported having an ADHD diagnosis completed an interview to confirm the diagnosis before participating in the study.

This study sought to include ADHD participants with a verified diagnosis, along with a Malingering group comprised of individuals who could feign ADHD to such a degree that they could be considered for the diagnosis. Therefore, both groups included participants who reported elevated symptom reports (T-score ≥60) on both the CAT-A Clinical Index and at least one CAT-A Current Symptom Scale (Inattention, Hyperactivity, or Impulsivity). All participants who failed to meet these criteria were removed from the study.

Participants consisted of 68 young adults, 34 who met criteria for ADHD and 34 individuals instructed to malinger ADHD. The two groups were deliberately matched on age, sex, and ethnicity. Each group included 16 males (47.1%) and 18 females (52.9%). The majority of each group identified as Caucasian (n = 30; 88.2% of each group). Two participants in each group identified as Asian (5.9%). Two Malingering participants identified as African American/Black (5.9%), and two ADHD participants identified as Biracial (5.9%). The mean age of the ADHD group was 18.82 years (SD = 1.51) and the Malingering group's mean age was 18.76 years (SD = 0.86). All ADHD participants self-reported a primary diagnosis of ADHD. Fourteen reported at least one comorbid disorder, which included learning disorder (n = 8, 23.5%), depression (n = 3, 8.8%), and anxiety (n = 3, 8.8%). The majority of Malingering participants did not report a disability. However, one Malingering participant reported a diagnosis of depression (2.9%) and another reported anxiety (2.9%). Among the ADHD group, 30 participants (88.2%) reported being prescribed medication, and 28 of those participants reported regular use of stimulants. Three Malingering participants disclosed regular use of an unspecified type of medication.

Measures

ADHD interview. A semi-structured interview was created for this study to corroborate self-reported diagnosis in ADHD participants. The interview included a series of questions regarding past and current symptoms, functional impairment, and history of diagnosis and treatment. The interview was conducted by a licensed psychologist or trained advanced doctoral student in psychology. Individuals were retained for analyses if they met the following criteria: (a) a diagnosis of ADHD received from a qualified professional (e.g., psychologist, physician), (b) symptom onset before the age of 12, (c) at least two CAT-A current symptom indexes above the mild range (T-score >60), and (d) reported impairment in at least one area (i.e., academic, occupational, or social) described in detail and judged by the interviewer to be substantial in severity.

Multidimensional ADHD Rating Scale (MARS). The MARS (Potts et al., 2021) consisted of 18 symptom, 22 impairment, and symptom validity items. All items were scaled from 0 to 8 (Symptom scale: 0 = "Never" to 8 = "Very often"; Impairment scale: 0 = "Not at all" to 8 = "Severe"). In the present study, only the four-item Symptom Validity Index was used. In prior work it had demonstrated moderate sensitivity to malingering (64.7%) with optimal specificity (91.8%). Similarly, ROC analyses had found that the MARS clinical indexes (symptom and impairment) yielded strong discrimination of ADHD from Controls (AUC range 0.85 to 0.88), and the SV-4 index was a good discriminator of ADHD from Malingering (AUC = 0.77). The MARS also contained three embedded "catch" items (e.g., "respond 3 if you are still reading this survey") to assess for effort and attention to the survey. Individuals who responded incorrectly to any catch item were removed from analyses.

Clinical Assessment of Attention Deficit-Adult. The Clinical Assessment of Attention Deficit-Adult (CAT-A; Bracken & Boatwright, 2005) is a 108-item adult ADHD self-report measure that includes a Childhood Memories section (54 items) and Current Symptoms section (54 items). This study used the Current Symptom Inattentive, Hyperactivity, and Impulsivity Scales and the CAT-A Clinical Index (derived from both childhood and current symptom sections) to select participants for analyses. A T-score ≥60 falls within clinical risk ranges. In a validation study, the authors reported the CAT-A scales had a correct classification accuracy between 79% and 88% for the Symptom scales to differentiate clinical ADHD from learning disabilities and from non-ADHD, non-disabled controls (Bracken & Boatwright, 2005).

The current study used the CAT-A Infrequency scale to detect feigned ADHD and compare its effectiveness to WMT and MARS validity indicators. The Infrequency scale consists of 10 embedded items that were endorsed infrequently by both the clinical ADHD group (≤6%) and non-clinical general population (≤1%) during the standardization process. The manual indicates that high endorsement (strongly agree) on ≥4 items may indicate noncredible responding.

Word Memory Test. The Word Memory Test (WMT; Green, 2003) is a performance validity test consisting of recognition memory activities. After being presented twice with a set of pairs of words, clients are given an Immediate Recognition (IR) memory task. After a 30-minutes delay, they are administered a Delayed Recognition (DR) memory task. Consistency (CNS) is the calculation of reliability between responses on IR and DR subtests. Following the two primary subtests (IR and DR), the present study also administered a supplemental Multiple Choice (MC) memory task, which is considered to be slightly more challenging compared to the forced-choice trials due to the possibility of semantic interference with the six response options. As such, the MC subtest tends to have more variable specificity rates and lower overall classification accuracy to differentiate malingering from those with significant cognitive impairments (Green, 2003; Strauss et al., 2006).

Validation of the WMT has generally involved the test's ability to distinguish between malingered and genuine cases of cognitive impairment. Research has found that the WMT demonstrates high specificity (> 95%) to individuals with genuine cognitive impairment at a cut score of 82.5% correct (Green, 2003). The WMT subtests have high internal consistency and high intercorrelations (r = .80). The WMT displays good convergent validity with other measures of effort, such as the Test of Memory Malingering (r = .68; Strauss et al., 2006).

As a feigned ADHD detection measure, Marshall et al. (2010) found that the WMT manual's cut scores of ≤82.5% on the IR or CNS each demonstrated 63% sensitivity, while the DR subtest had 18% sensitivity, to detect suspected cases at optimal specificity (>90%). A simulation study found the WMT may be less effective at detecting a malingering group coached on ADHD symptoms, with the following reported sensitivity rates: WMT IR = 43%, WMT DR = 30%, and WMT CNS = 52%, and specificity rates of 95%, 95%, and 86%, respectively (Edmundson et al., 2017). The current study used the four WMT scores as separate validity indicators to detect feigned ADHD.

Procedure

ADHD participants completed the in-person screening interview before completing the computerized/electronic measures in a private room with an experimenter. Non-ADHD participants completed the online survey in small groups (<10 people) in a university computer lab. After electronically signing an informed consent form, all participants answered basic demographic questions. ADHD participants were instructed to complete all the measures honestly and based upon an unmedicated symptom presentation.

The Malingering group was provided with instructions to feign ADHD, including information about diagnosis and evaluation of ADHD adapted from the WebMD ADD and ADHD Health Center website (Centers for Disease Control and Prevention, 2014). Participants were informed that those who adhered to the assignment, faking ADHD without detection would be entered into a raffle drawing for a $100 gift card. Next, both groups were administered the WMT IR subtest, followed by the online rating scales in counterbalanced fashion: (a) MARS symptom items, (b) MARS symptom validity items, (c) MARS functional impairment items, and (d) the CAT-A. Thirty minutes after the completion of the IR subtest, participants were administered the WMT DR and MC subtests.

Results

Group means, standard deviations, *t*-tests, and effect sizes are reported in Table 1. The *t*-test results were statistically significant for the SV-index 4 and WMT DR, Consistency, and MC subtests using a Bonferroni correction (p ≤ .008). As expected, large effect sizes were found between ADHD and Malingering groups on the SV-Index 4 (d = 1.12), with higher scores for those in the Malingering group. Medium to large effect sizes were obtained on the WMT DR (d = 0.85), Consistency (d = 0.74), and MC subtests (d = 0.71), and small effect sizes were found for the WMT IR (d = 0.25) and CAT-A Infrequency scale (d = 0.24), all in the same direction.

With regard to detection of malingering, the AUC statistics, standard error, and confidence intervals are presented in Table 2, and classification accuracy calculations for the index to discriminate cases of malingering can be found in Table 3. ROC analyses found that the SV-index 4 was a good discriminator between Malingering and ADHD (AUC = 0.79). AUC values were lower for the WMT (AUC = 0.53–0.67) and CAT-A Infrequency scale (AUC = 0.55). The SV-index 4 performed better than all other validity measures, producing high classification accuracy (75%), with good sensitivity (61.8%) and specificity (88.2%).

Classification accuracy calculations indicated that while the CAT-A item count (≥4 items) had high specificity (94.1%) to not identify ADHD, this cut score yielded very weak sensitivity for malingering detection (17.6%) and an overall classification accuracy of 55.9%. Using a lower threshold (≥3 items) suggested by prior research (Marshall et al., 2010), resulted in a slight increase in sensitivity (32.4%), yet at a cost to specificity (85.3%), and modest overall classification accuracy (58.8%).

Although the WMT IR, DR, and CNS scores were effective at ruling out ADHD, with specificity between 94.1 and 97.1%, these validity tests had low sensitivity (14.7%–41.2%) to detect simulated malingering. The DR subtest had slightly better classification accuracy overall (69.1%) compared to the other WMT subtests, yet still had sensitivity rates that were under 50%. Though the WMT

Table 1. Descriptive Statistics.

	ADHD group (n = 34)		Malingering group (n = 34)				
Index	*M*	*SD*	*M*	*SD*	*t*	*p*	Cohen's *d*
MARS							
Symptom validity index 4	9.1	4.6	15.7	6.9	4.61	<.001	1.12
CAT-A infrequency scale							
Item count	1.2	1.3	1.6	1.6	1.01	.318	0.94
WMT							
Immediate recognition	95.8	5.7	93.8	9.9	1.03	.310	0.25
Delayed recognition	96.6	6.2	87.1	14.6	3.50	.001	0.85
Consistency	94.5	7.2	85.2	16.3	3.04	.003	0.74
Multiple choice	89.1	13.2	76.6	20.9	2.95	.005	0.71

Table 2. Area Under the Curve from Receiver Operating Characteristic Analyses to Classify Malingering between ADHD and Malingering Groups.

				Asymptotic 95% confidence interval	
Index	AUC	Standard error	*p*-Value	Lower	Upper
MARS					
Symptom validity index 4	0.79	0.06	<.001	0.68	0.90
CAT-A infrequency scale					
Item count	0.55	0.07	.52	0.41	0.68
WMT					
Immediate recognition	0.53	0.07	.69	0.39	0.67
Delayed recognition	0.67	0.07	.02	0.54	0.80
Consistency	0.65	0.07	.04	0.51	0.78
Multiple choice	0.65	0.07	.04	0.52	0.78

Table 3. Classification Accuracy Calculations to Classify Malingering between ADHD and Malingering Groups.

							Current study base rate (50%)		Estimated base rate (25%)	
Measure	Cut score	Sensitivity	Specificity	Accuracy	LR+	LR−	PPP	NPP	PPP	NPP
MARS										
SV-Index 4	≥14.5	61.8%	88.2%	75.0%	5.2	0.4	84.0%	69.8%	63.6%	87.4%
CAT-A infrequency scale										
Marshall et al. (2010)	≥3	32.4%	85.3%	58.8%	2.2	0.8	68.8%	55.8%	42.4%	79.1%
Test manual	≥4	17.6%	94.1%	55.9%	3.0	0.9	75.0%	53.3%	49.9%	77.4%
WMT										
Immediate recognition	≤ 82.5	14.7%	97.1%	55.9%	5.1	0.9	83.3%	53.2%	62.8%	77.3%
Delayed recognition	≤ 82.5	41.2%	97.1%	69.1%	14.2	0.6	93.3%	62.3%	82.6%	83.2%
Consistency	≤ 82.5	41.2%	94.1%	67.6%	7.0	0.6	87.5%	61.5%	69.9%	82.8%
Multiple choice	≤ 82.5	52.9%	79.4%	66.2%	2.6	0.6	72.00%	62.8%	46.1%	83.5%
WMT pass/fail (IR/DR/Con)		41.2%	94.1%	67.6%	6.9	0.6	87.50%	61.5%	69.9%	82.8%

Note. WMT pass/fail = fail on one subtest.

MC subtest demonstrated higher sensitivity (52.9%), this subtest had less than optimal specificity (79.4%), with seven ADHD participants' scores falling in the "fail" range on the MC subtest. The creation of a global WMT fail variable (failure on either the IR, DR, or CNS) also did not improve sensitivity (41.2%) or specificity (94.1%). Additionally, ROC curves were examined to determine whether an alternate cut score could increase detection of malingering on any of the scales. Unfortunately, review of the data indicated that sensitivity to malingering could not be increased, while maintaining optimal specificity (~90%).

Predictive power calculations also support the SV-index relative to the two published tests. Though true prevalence of malingered ADHD is unknown, research estimates a range of 20% to 30% in clinical evaluations (e.g., Harrison & Edwards, 2010; Marshall et al., 2010, 2016) Using a 25% estimate of the base rate of malingering, the SV-index had the highest negative predictive power (87.4%), indicating high accuracy to identify true negatives. The SV-index also had one of the highest positive predictive powers (63.6%). Only the Consistency index had a higher PPP calculation, related to the high specificity of this measure to confidently rule out malingering.

Discussion

This study examined the ability of a new Symptom Validity Index within a comprehensive rating scale (MARS) to detect feigned ADHD from genuine ADHD. The SV-index 4 outperformed two published validity tests. Specifically, the SV-index 4 had good sensitivity (61.8%) at close to optimal specificity (88.2%), and an overall classification accuracy of 75.0%. Both the WMT and CAT-A Infrequency scale demonstrated modest sensitivity (<50%) and/or specificity rates (≤85%) that were not acceptable for clinical practice. Overall, the results suggest that the MARS SV-index 4 is a promising symptom validity test for the detection of feigned ADHD. This study demonstrated that a validity test specifically designed to detect feigned ADHD outperformed validity tests designed for other populations.

The SV index is intended to lead malingerers to endorse items that they think reflect characteristics of ADHD, but that are not actually commonly reported by individuals with ADHD. Symptom validity items often tap into false perceptions, stereotypes, and misconceptions that the general population may have about a disorder (Rogers, 2008). Research has highlighted that the general public is aware of ADHD, but this awareness primarily extends to basic knowledge about externalizing behaviors, such as excessive movement (McLeod et al., 2007).

The present study found that the WMT was relatively ineffective at detecting simulated cases of malingered ADHD. On the positive side, the WMT IR, DR, and CNS had high specificity to identify true negative cases of ADHD, yet these measures had low to moderate sensitivity to detect malingering. The WMT MC subtest did not demonstrate optimal specificity required of effective validity tests, perhaps because this subtest is slightly more challenging compared to the easier forced-choice subtests, and consequently is less able to discriminate malingering from those with true cognitive impairments (Green, 2003; Strauss et al., 2006). We should note that the classification accuracy of the WMT for detect malingered cognitive/memory impairments is excellent (Sollman & Berry, 2011). However, the WMT is a PVT that is not designed to identify overreporting of symptoms. The current findings suggest that this widely used PVT designed to detect other feigned disorders (e.g., traumatic brain injury) is not particularly effective at detecting feigned ADHD.

With regard to the CAT-A, our results did not support the ability of the Infrequency scale to detect cases of malingering, regardless of the cut score or the calculation method (item count or total score). The use of the manual's recommended cut score (≥4 items) effectively ruled out ADHD (specificity = 94.1%) but demonstrated weak sensitivity (17.6%). Use of the lower cut point proposed by research (≥3 items; Marshall et al., 2010), raised sensitivity levels (32.4%), but resulted in poor specificity (85.3%). The CAT-A was less sensitive to malingering in the present study compared to previous archival research (sensitivity = 58%; Marshall et al., 2010). The discrepant findings are perhaps best explained by the differences between the studies' research designs. Marshall et al. (2010) utilized an archival study design, in which individuals were assessed for ADHD and suspected of malingering based upon scores on validity tests. In contrast, the present study employed a simulation research design that instructed non-ADHD participants to perform like a person with ADHD. Simulation designs tend to have higher internal validity as there is experimental control over group assignment, but simulated cases of malingering often tend to overestimate deficits compared to suspected archival cases (Brennan & Gouvier, 2006; Rogers, 2008). The relative limitations of these studies suggest that additional research may be warranted to corroborate the classification accuracy of this symptom validity test. At this time, the obtained sensitivity rates are too low and variable (19%–58%) to indicate that this measure could be relied upon as an individual detector of those faking ADHD.

Implications for Clinical Practice

Validity tests have become an increasingly important part of psychodiagnostic evaluations. The MARS SV-index 4 is one such measure that seems to offer clinicians a metric for detection of feigned ADHD. Though such a measure may be useful for detecting feigned ADHD, scores on validity tests alone are insufficient evidence of malingering. Data from well-established and research-supported validity tests should be considered within the context of a comprehensive evaluation. This would include consideration of scores on all diagnostic assessments (e.g., symptoms, impairment, cognitive testing), and information in the clinical interview, as well as from collateral reports (Chafetz et al., 2015; Heilbronner et al., 2009; Iverson, 2006; Lovett et al., 2020; Sibley, 2021). Similar to other diagnostic evaluations, clinicians are urged to consider alternative explanations for the obtained scores on all tests, including validity tests. A clinician also could document any plausible evidence that an individual was manipulating his/her performance intentionally to obtain external incentives (Chafetz et al., 2015; Heilbronner et al., 2009; Iverson, 2006). In conclusion, the SV-index 4 could be a helpful tool as part of a larger effort to detect malingering.

Limitations and Directions for Future Research

This study is limited by the use of a simulation design, in that such studies cannot fully replicate real-world situations (i.e., malingering on a high-stakes exam). Although this study utilized a modest incentive to promote instruction following, a gift card is less motivating than other natural incentives for feigning ADHD (e.g., medication, test accommodations). External validity is partially improved by coaching participants on the diagnosis. However, while there is some consistency in coaching instructions across studies (e.g., use of scenario and ADHD diagnostic information), differences in coaching instructions could limit generalizability across these studies (Brennan & Gouvier, 2006; Rogers, 2008; Rogers & Gillard, 2011). Future research on validity measures is needed to further examine their effectiveness as an important component of adult ADHD evaluations.

Another limitation is the reliance on a self-report interview to determine diagnosis. Though this study also utilized a screening form and used the CAT-A clinical scales to verify diagnosis, it is possible that a portion of the ADHD sample included individuals that do not meet all *DSM-5* criteria for ADHD diagnosis, and equally possible that some individuals with ADHD were removed because they underreported symptoms. The selection procedures did not include collateral symptom reports or neuropsychological testing. Consequently, study findings do not necessarily generalize to all ADHD populations.

Although this study provides support for the use of the SV-index to detect simulated cases of malingering, it is unknown whether this measure will remain effective in other samples. Additional research using both simulation designs and clinical studies of those referred for ADHD assessment is warranted, as the strengths of both research methodologies should help to provide the best support for a validity test's ability to detect malingering (Rogers, 2008). Although the assumption of validity tests is that they can be easily passed by most, regardless of participant characteristics (e.g., age, gender, ethnicity, cognitive impairment), it is possible that differences across studies may be attributed, even in some small part, to differences in participant characteristics. With this in mind, future research with the MARS should extend to diverse groups, including both college and working adult populations. This research could identify different malingering response styles across groups, and if present, what measures/methods can best detect them.

Acknowledgments

The authors would like to thank the following individuals for their contribution to this study: Kevin Antshel, Whitney Wood, Jillian Berfield, Kaya Wechsler, and Katie Herer.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Heather E. Potts https://orcid.org/0000-0002-5597-7594
Benjamin J. Lovett https://orcid.org/0000-0003-4785-3070

References

Barkley, R., Murphy, K., & Fischer, M. (2008). *ADHD in adults: What the science says*. Guilford Press.

Benson, K., Flory, K., Humphreys, K. L., & Lee, S. S. (2015). Misuse of stimulant medication among college students: A comprehensive review and meta-analysis. *Clinical Child and Family Psychology Review*, *18*, 50–76. https://doi.org/10.1007/s10567-014-0177-z

Booksh, R. L., Pella, R. D., Singh, A. N., & Gouvier, W. D. (2010). Ability of college students to simulate ADHD on objective measures of attention. *Journal of Attention Disorders*, *13*, 325–338. https://doi.org/10.1177/1087054708329927

Bracken, B., & Boatwright, B. (2005). *CAT-A, clinical assessment of attention deficit-adult: Professional manual*. Psychological Assessment Resources.

Brennan, A. M., & Gouvier, W. D. (2006). Are we honestly studying malingering? A profile and comparison of simulated and suspected malingerers. *Applied Neuropsychology*, *13*(1), 1–11. https://doi.org/10.1207/s15324826an1301_1

Centers for Disease Control and Prevention. (2014). *Attention deficit/hyperactivity disorder*. https://www.cdc.gov/ncbddd/adhd/index.html

Chafetz, M. D., Williams, M. A., Ben-Porath, Y. S., Bianchini, K. J., Boone, K. B., Kirkwood, M. W., Larrabee, G. J., & Ord, J. S. (2015). Official position of the American Academy of Clinical Neuropsychology: Social Security Administration policy on validity testing: Guidance and recommendations for change. *Clinical Neuropsychologist*, *29*(6), 723–740. https://doi.org/10.1080/13854046.2015.1099738

Conners, C. K., Erhardt, D., & Sparrow, E. (1999). *Conners' adult ADHD rating scales*. Multi-Health Systems.

Cook, C. M., Bolinger, E., & Suhr, J. A. (2016). Further validation of the Conner's Adult Attention Deficit/Hyperactivity Rating Scale Infrequency Index (CII) for detection of non-credible report of Attention Deficit/Hyperactivity Disorder symptoms.

Archives of Clinical Neuropsychology, *31*, 358–364. https://doi.org/10.1093/arclin/acw015
Davidson, M. A. (2008). ADHD in adults: A review of the literature. *Journal of Attention Disorders*, *11*(6), 628–641. https://doi.org/10.1177/1087054707310878
Edmundson, M., Berry, D. T. R., Combs, H. L., Brothers, S. L., Harp, J. P., Williams, A., Rojas, S. L., Saleh, A. K., & Scott, A. B. (2017). The effects of symptom information coaching on the feigning of adult ADHD. *Psychological Assessment*, *29*, 1429–1436. https://doi.org/10.1037/pas0000478
Fuermaier, A. B. M., Tucha, O., Koerts, J., Lange, K. W., Weisbrod, M., Aschenbrenner, S., & Tucha, L. (2017). Noncredible cognitive performance at clinical evaluation of adult ADHD: An embedded validity indicator in a visuospatial working memory test. *Psychological Assessment*, *29*, 1466–1479. https://doi.org/10.1037/pas0000534
Green, P. (2003). *The Word Memory Test*. Green's Publishing.
Harrison, A. G. (2004). An investigation of reported symptoms of ADHD in a university population. *The ADHD Report*, *12*(6), 8–11. https://doi.org/10.1521/adhd.12.6.8.55256
Harrison, A. G. (2017). Clinical, ethical, and forensic implications of a flexible threshold for LD and ADHD in postsecondary settings. *Psychological Injury and Law*, *10*(2), 138–150. https://doi.org/10.1007/s12207-017-9291-9
Harrison, A. G., & Edwards, M. J. (2010). Symptom exaggeration in post-secondary students: Preliminary base rates in a Canadian sample. *Applied Neuropsychology*, *17*, 135–143. https://doi.org/10.1080/09084281003715642
Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., & Millis, S. R. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. *Clinical Neuropsychologist*, *23*(7), 1093–1129. https://doi.org/10.1080/13854040903155063
Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology*, *13*(2), 77–90. https://doi.org/10.1207/s15324826an1302_3
Jasinski, L. J., & Ranseen, J. D. (2011). Malingered ADHD evaluations: A further complication for accommodations reviews. *The Bar Examiner*, *80*, 6–16.
Lewandowski, L., Lambert, T. L., Lovett, B. J., Panahon, C. J., & Sytsma, M. R. (2014). College students' preferences for test accommodations. *Canadian Journal of School Psychology*, *29*, 116–126. https://doi.org/10.1177/0829573514522116
Lovett, B. J., & Nelson, J. M. (2018). Assessing adults for ADHD: A systematic, evidence-based protocol. *Journal of Health Service Psychology*, *44*, 48–52.
Lovett, B. J., Spenceley, L. M., & Lewandowski, L. J. (2020). Response validity in psychoeducational assessment: A primer for school psychologists. *Contemporary School Psychology*. Advance online publication. https://doi.org/10.1007/s40688-020-00334-5
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*, 1204–1237. https://doi.org/10.1080/13854046.2010.514290
Marshall, P. S., Hoelzle, J. B., Heyerdahl, D., & Nelson, N. W. (2016). The impact of failing to identify suspect effort in patients undergoing adult attention-deficit/hyperactivity disorder (ADHD) assessment. *Psychological Assessment*, *28*, 1290–1302. https://doi.org/10.1037/pas0000247
McLeod, J. D., Fettes, D. L., Jensen, P. S., Pescosolido, B. A., & Martin, J. K. (2007). Public knowledge, beliefs, and treatment preferences concerning attention-deficit hyperactivity disorder. *Psychiatric Services*, *58*, 626–631. https://doi.org/10.1176/appi.ps.58.5.626
Musso, M. W., & Gouvier, W. D. (2014). "Why is this so hard?": A review of detection of malingered ADHD in college students. *Journal of Attention Disorders*, *18*, 186–201. https://doi.org/10.1177/1087054712441970
Nelson, J. M., Whipple, B., Lindstrom, W., & Foels, P. A. (2019). How is ADHD assessed and documented? Examination of psychological reports submitted to determine eligibility for postsecondary disability. *Journal of Attention Disorders*, *23*(14), 1780–1791.
Park, T. M., & Haning, W. F., 3rd. (2016). Stimulant use disorders. *Child and Adolescent Psychiatric Clinics of North America*, *25*, 461–471. https://doi.org/10.1016/j.chc.2016.02.004
Potts, H. E., Lewandowski, L. J., & Lovett, B. J. (2021). The multidimensional ADHD Rating Scale: A measure of symptoms, impairment, and symptom validity. *Journal of Clinical and Experimental Neuropsychology*, *43*(4), 426–436. https://doi.org/10.1080/13803395.2021.1942795
Rogers, R. (2008). *Clinical assessment of malingering and deception* (3rd ed.). Guilford Press.
Rogers, R., & Gillard, N. D. (2011). Research methods for the assessment of malingering. In B. Rosenfeld & S. D. Penrod (Eds.), *Research methods in forensic psychology* (pp. 174–188). John Wiley & Sons Inc.
Sagar, S., Miller, C. J., & Erdodi, L. A. (2017). Detecting feigned attention-deficit/hyperactivity disorder (ADHD): Current methods and future directions. *Psychological Injury and Law*, *10*, 105–113. https://doi.org/10.1007/s12207-017-9286-6
Sibley, M. H. (2021). Empirically-informed guidelines for first-time adult ADHD diagnosis. *Journal of Clinical and Experimental Neuropsychology*, *43*(4), 340–351. https://doi.org/10.1080/13803395.2021.1923665
Sollman, M. J., & Berry, D. T. (2011). Detection of inadequate effort on neuropsychological testing: A meta-analytic update and extension. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *26*, 774–789. https://doi.org/10.1093/arclin/acr066
Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford University Press.
Suhr, J., Hammers, D., Dobbins-Buckland, K., Zimak, E., & Hughes, C. (2008). The relationship of malingering test failure to self-reported symptoms and neuropsychological findings in adults referred for ADHD evaluation. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *23*(5), 521–530. https://doi.org/10.1016/j.acn.2008.05.003
Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assessments. *Applied Neuropsychology*, *14*, 189–207. https://doi.org/10.1080/09084280701509083

Tucha, L., Fuermaier, A. B., Koerts, J., Groen, Y., & Thome, J. (2015). Detection of feigned attention deficit hyperactivity disorder. *Journal of Neural Transmission*, *122*, S123–S134. https://doi.org/10.1007/s00702-014-1274-3

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

Author Biographies

Heather E. Potts, Ph.D. is an attending psychologist in the Outpatient Psychiatry Service at Boston Children's Hospital and Instructor of Psychology in the Department of Psychiatry at Harvard Medical School. Her clinical specialization and research interests are in the assessment and treatment of Complex ADHD.

Lawrence J. Lewandowski, Ph.D. is Emeritus Professor of Psychology and Meredith Professor for Teaching Excellence at Syracuse University. His research involves assessment and interventions for neurodevelopmental disorders with an emphasis on testing accommodations.

Benjamin J. Lovett, Ph.D., is Associate Professor of Psychology and Education at Teachers College, Columbia University. His research areas include the diagnosis of learning and attention problems, the provision of disability accommodations, and the management of test anxiety.