University of Groningen

Noncredible cognitive performance at clinical evaluation of adult ADHD

Fuermaier, Anselm B M; Tucha, Oliver; Koerts, Janneke; Lange, Klaus W; Weisbrod, Matthias; Aschenbrenner, Steffen; Tucha, Lara

Published in:
Psychological Assessment

DOI:
[10.1037/pas0000534](https://doi.org/10.1037/pas0000534)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

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(12), 1466-1479. <https://doi.org/10.1037/pas0000534>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Noncredible Cognitive Performance at Clinical Evaluation of Adult ADHD: An Embedded Validity Indicator in a Visuospatial Working Memory Test

Anselm B. M. Fuermaier, Oliver Tucha,
and Janneke Koerts
University of Groningen

Klaus W. Lange
University of Regensburg

Matthias Weisbrod
SRH Clinic Karlsbad-Langensteinbach, Karlsbad, Germany, and
University of Heidelberg

Steffen Aschenbrenner
SRH Clinic Karlsbad-Langensteinbach, Karlsbad, Germany

Lara Tucha
University of Groningen

The assessment of performance validity is an essential part of the neuropsychological evaluation of adults with attention-deficit/hyperactivity disorder (ADHD). Most available tools, however, are inaccurate regarding the identification of noncredible performance. This study describes the development of a visuospatial working memory test, including a validity indicator for noncredible cognitive performance of adults with ADHD. Visuospatial working memory of adults with ADHD ($n = 48$) was first compared to the test performance of healthy individuals ($n = 48$). Furthermore, a simulation design was performed including 252 individuals who were randomly assigned to either a control group ($n = 48$) or to 1 of 3 simulation groups who were requested to feign ADHD ($n = 204$). Additional samples of 27 adults with ADHD and 69 instructed simulators were included to cross-validate findings from the first samples. Adults with ADHD showed impaired visuospatial working memory performance of medium size as compared to healthy individuals. Simulation groups committed significantly more errors and had shorter response times as compared to patients with ADHD. Moreover, binary logistic regression analysis was carried out to derive a validity index that optimally differentiates between true and feigned ADHD. ROC analysis demonstrated high classification rates of the validity index, as shown in excellent specificity (95.8%) and adequate sensitivity (60.3%). The visuospatial working memory test as presented in this study therefore appears sensitive in indicating cognitive impairment of adults with ADHD. Furthermore, the embedded validity index revealed promising results concerning the detection of noncredible cognitive performance of adults with ADHD.

Public Significance Statement

The clinical evaluation of adults with attention-deficit/hyperactivity disorder (ADHD) is complicated by the fact that a sizable number of individuals deliberately feign symptoms and impairments. This study contributes to solving this issue by the development of a test that is helpful in the assessment of both, function and effort of individuals.

Keywords: adult ADHD, noncredible, performance validity, feigning, faking

Anselm B. M. Fuermaier, Oliver Tucha, and Janneke Koerts, Department of Clinical and Developmental Neuropsychology, University of Groningen; Klaus W. Lange, Department of Experimental Psychology, University of Regensburg; Matthias Weisbrod, Department of Psychiatry and Psychotherapy, SRH Clinic Karlsbad-Langensteinbach, Karlsbad, Germany, and Department of General Psychiatry, Center of Psychosocial Medicine, University of Heidelberg; Steffen Aschenbrenner, Department of Clinical Psychology and Neuropsychology, SRH Clinic

Karlsbad-Langensteinbach; Lara Tucha, Department of Clinical and Developmental Neuropsychology, University of Groningen.

Matthias Weisbrod and Steffen Aschenbrenner have contracts for the development of neuropsychological diagnostic and training tools with Schuhfried GmbH.

Correspondence concerning this article should be addressed to Anselm B. M. Fuermaier, Faculty of Behavioural and Social Sciences, Department of Clinical and Developmental Neuropsychology, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, the Netherlands. E-mail: a.b.m.fuermaier@rug.nl

In the last decade, a growing body of literature has emphasized that the diagnostic assessment of attention-deficit/hyperactivity disorder (ADHD) in adulthood is complicated by a considerable number of individuals showing noncredible effort and symptom reporting at clinical evaluation (Fuermaier, Tucha, Koerts, Weisbrod, et al., 2016; Harrison & Edwards, 2010; Harrison, Edwards, & Parker, 2007; L. Tucha, Sontag, Walitza, & Lange, 2009; Pella, Hill, Shelton, Elliott, & Gouvier, 2012; Suhr, Hammers, Dobbins-Buckland, Zimak, & Hughes, 2008; Sullivan, May, & Galbally, 2007). A proportion of these individuals may exaggerate or deliberately feign symptoms and impairments to benefit from being diagnosed with ADHD. Possible benefits of being diagnosed with this disorder are having access to stimulant medication, either as a cognitive enhancer or for recreational purposes, or to receive special accommodations in the academic context, such as being awarded extra time for assignments and exams (Lensing, Zeiner, Sandvik, & Opjordsmoen, 2013; Rabiner, 2013). Thus, there is broad agreement that the assessment of symptom and performance validity should become an essential part of the clinical evaluation of adults with ADHD (Musso & Gouvier, 2014; L. Tucha, Fuermaier, Koerts, Groen, & Thome, 2015). In this respect, several tools have been examined for their utility to detect noncredible performance and symptom reporting of individuals evaluated for ADHD, comprising personality inventories, self-report rating scales of ADHD symptom severity, neuropsychological tests routinely used in clinical practice for the assessment of cognition, as well as tests that were specifically designed for the detection of feigned cognitive dysfunctions (including SVTs; for comprehensive reviews, see Musso & Gouvier, 2014; L. Tucha et al., 2015). Even though some of these tools showed promising classification rates in terms of sensitivity and specificity, there is no consensus yet on a standardized procedure for the assessment of noncredible performance and symptom reporting of individuals at clinical evaluation of ADHD.

With regard to neuropsychological functioning of adults with ADHD, a large number of studies described ADHD as a disorder of attention and executive control (Boonstra, Oosterlaan, Sergeant, & Buitelaar, 2005; Boonstra, Kooij, Oosterlaan, Sergeant, & Buitelaar, 2010; Fuermaier et al., 2015; Lange et al., 2014; Mostert et al., 2015; Schoechlin & Engel, 2005; L. Tucha et al., 2008; O. Tucha et al., 2006). In this context, findings indicating working memory deficits are very robust among both children (Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005) and adults with ADHD (Alderson, Kasper, Hudec, & Patros, 2013), so that working memory deficits consequently have been proposed as a core feature of, and potential endophenotype for, ADHD (Alderson et al., 2013; Castellanos & Tannock, 2002; Rapport, VanVoorhis, Tzelepis, & Friedman, 2001; van Ewijk et al., 2014). Given the common use of neuropsychological test batteries, including working memory tests, in the assessment of adults with ADHD, several studies examined the utility of standard measures of cognition for detecting noncredible cognitive performance at clinical evaluation (Musso & Gouvier, 2014; L. Tucha et al., 2015). Among these studies, three studies also included working memory paradigms (Booksh, Pella, Singh, & Gouvier, 2010; Harrison, Rosenblum, & Currie, 2010; Suhr et al., 2008), such as subtests of the Wechsler Adult Intelligence Scale (e.g., the Digit Span as a measure of verbal working memory; Wechsler, 1997). Although Suhr and colleagues (2008) found significant differences between a non-

credible and a credible group of patients with ADHD, Harrison et al. (2010) as well as Booksh et al. (2010) failed to adequately differentiate these groups from one other on the basis of these measures. Harrison et al. (2010) showed that Digit Span performance was helpful in identifying credible cognitive performance of individuals with ADHD with high specificity, however, sensitivity toward noncredible performance was low. Conclusive findings on the usefulness of working memory paradigms for the detection of noncredible performance of adults with ADHD could therefore not be drawn from these studies.

Although previous studies using working memory tests for the detection of noncredible cognitive performance of adults with ADHD included paradigms of verbal working memory, the present study aimed to explore the potential value of a test measuring visuospatial working memory (VSWM). There are reasons to assume that a test for VSWM may represent a promising candidate for the simultaneous assessment of both function and noncredible cognitive performance at clinical evaluation of ADHD. First, VSWM deficits were suggested as one of the most consistently impaired executive functions in patients with ADHD (Alderson et al., 2013; Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005; van Ewijk et al., 2014). Second, we assume that such a test might also be helpful for the assessment of credibility of clients, as most cognitive tasks in daily life somehow involve the processing of verbal information. Thus, people appear to be very experienced with regard to the level of verbal functions, which may facilitate them in feigning cognitive dysfunctions on tests using verbal material (such as verbal working memory). However, as people have less experience in dealing with nonverbal (e.g., spatial) information in daily life, it can be assumed that individuals feigning ADHD may have more difficulties in mimicking the level of performance that one would expect from genuine ADHD on tests using visuospatial information.

For the purpose of the present study, a VSWM test (VSWMT) was developed that followed the principles of the task, as devised by Brooks (1968), in requiring participants to develop, inspect, and navigate through mental images (e.g., a geometric figure shaping the letter "F"). Healthy individuals were allocated to either a control condition which requires participants to perform the VSWMT to the best of their abilities, or to one of three simulation conditions asking participants to perform the VSWMT as if they suffered from ADHD. VSWMT performance of the simulation groups were compared to VSWMT performance of a group of adults with ADHD as well as a healthy control group to explore the utility of the VSWMT in detecting noncredible cognitive performance.

When compared to a healthy comparison group, patients with ADHD were expected to show impairments in VSWM. Further, participants of the simulation conditions were expected to show overly poor performance on the VSWMT as compared to both healthy comparison participants and patients with ADHD. A valid indicator of noncredible cognitive performance of adults with ADHD should achieve specificity of at least 90% (Boone, 2007; Marshall et al., 2010). Additional analyses were conducted to identify cut-off scores for the VSWMT that would result in the optimal combination of sensitivity and specificity rates (Fuermaier, Tucha, Koerts, Grabski, et al., 2016).

Method

Design and Procedure

The present study employed a simulation design in which the test performance of adults with ADHD was compared to test performances of several groups of healthy individuals. Although some of the healthy individuals (i.e., the healthy comparison group [HCG] and the control group [CG]) were instructed to show normal behavior, others (i.e., the naïve simulation group [NSG], symptom-coached simulation group [SSG], test-coached simulation group [TSG], and symptom- and test-coached simulation group [STSG]) were instructed to assume the role of someone who aims to feign ADHD. Table 1 presents an overview of instructions given per group, including the type of information participants received, characteristics of groups are presented in Table 2. The utility of the newly developed VSWMT for the detection of feigned ADHD will be determined by comparing the performances of patients with ADHD with the performances of a collapsed group of all individuals instructed to feign ADHD. An effort index (VSWMT index) will be calculated via a weighted sum of all variables of the VSWMT. The weights will be determined via regression coefficients of a logistic regression model for the prediction of feigned ADHD relative to true ADHD. In addition, the utility of the effort index will be cross-validated on two additional and independent samples, that is, an additional sample of patients with ADHD as well as another group of instructed simulators.

Assessment of patients with ADHD. All patients with ADHD were tested individually and received no reward for their participation. Written informed consent was sought from all participants prior to the assessment. The first sample of patients with ADHD (ADHD1, $n = 48$) was assessed with a comprehensive battery consisting of self-report questionnaires, interviews, the VSWMT, and routine measures of cognition (i.e., tests for vigilance and verbal working memory). The second and independent sample of patients with ADHD (ADHD2, $n = 27$) was assessed by requesting them to perform the VSWMT in addition to three established effort tests, in this case the TOMM, the DCT, and the b Test. Patients were only included in the present study if they ‘passed’ all three effort tests to have sufficient evidence that all patients of this sample showed credible effort at neuropsychological assessment. This additional sample of patients with ADHD was recruited to examine whether VSWMT performance, as obtained from the

patients of ADHD1, can be replicated on an independent sample of patients. The total duration of the assessment of patients with ADHD was about two hours. The study was conducted in compliance with ethical standards of the Helsinki Declaration and was approved by the local institutional ethical committee (Medical Faculty of the University of Heidelberg, Germany).

Assessment of healthy participants. All healthy participants were tested individually in a quiet laboratory. At the beginning of the experiment, descriptive information was obtained including age, sex, vocabulary skills (intellectual functions), and self-reported ADHD symptom severity. Further, participants were asked for any history of psychiatric or neurological diseases as well as pharmacological treatment. The assessment of healthy individuals was approved by the Ethical Committee Psychology affiliated with the University of Groningen, the Netherlands. All participants gave written informed consent prior to participation and were debriefed at the end of the assessment.

Healthy comparison group (HCG) and control group (CG). Participants of the HCG and CG were asked to perform all tests applied to the best of their abilities. Participants received a notification by email on the day before the assessment in which the clinical significance of the study was outlined, but which did not contain information on the aim of the study. The duration of the assessment of these groups was about 50 min.

Simulation groups. Participants of the simulation groups (NSG, SSG, TSG, and STSG) were asked to perform the VSWMT while pretending to be affected with ADHD (feigning ADHD). For this purpose, participants were presented with a vignette, describing reasons why someone would be motivated to feign ADHD. Several benefits that may be associated with a diagnosis of ADHD were introduced in the vignette; for example, financial accommodations, more flexibility and freedom concerning working hours and deadlines, and the prescription of stimulant medication. Participants were further explicitly instructed to make their feigning of ADHD seem realistic, that is, by avoiding a very obvious exaggeration. To encourage participants to feign ADHD in a believable and realistic manner, participants were informed that the participant who feigns the condition best would be awarded with a top of the range tablet PC. However, due to ethical reasons, the tablet PC was in fact assigned randomly to one of the participants across all conditions, independently of the participants’ test performance. The NSG received no further information and no

Table 1
Type of Information and Instruction Given per Group

Group	Scenario to feign ADHD (vignette)	Information about ADHD symptoms (symptom-coaching)	Information about assessment (test-coaching)	Instruction
Patients with ADHD (ADHD1, ADHD2)	No	No	No	Normal behavior
Healthy comparison group	No	No	No	Normal behavior
Control group	No	No	No	Normal behavior
Naïve simulation group	Yes	No	No	Feign ADHD
Symptom-coached simulation group	Yes	Yes	No	Feign ADHD
Test-coached simulation group	Yes	No	Yes	Feign ADHD
Symptom- and test-coached simulation group	Yes	Yes	Yes	Feign ADHD

Note. ADHD = attention-deficit/hyperactivity disorder; ADHD1 = first sample of patients with ADHD; ADHD2 = second sample of patients with ADHD.

Table 2
Characteristics of Participants

Characteristic	ADHD1	ADHD2	HCG	CG	NSG	SSG	TSG	SG2
<i>N</i>	48	27	48	48	68	68	68	69
Age (in years)	34.2 ± 11.3	33.5 ± 11.7	34.1 ± 11.6	20.5 ± 2.2	21.1 ± 2.3	21.3 ± 2.3	21.5 ± 2.1	20.0 ± 2.2
Gender (female/male)	21/27	10/17	26/22	31/17	54/14	49/19	41/27	54 ± 15
Vocabulary skills ^a	102.2 ± 11.4	105.0 ± 9.8	104.3 ± 11.8	98.3 ± 9.2	99.5 ± 9.5	100.4 ± 8.6	102.0 ± 10.1	
ADHD subtype ^b	20/0/28	8/0/19						
Comorbidity ^c	18/5/4/3/2/1/1	4/0/0/0/0/1/0						
Medication ^d	13/0	8/0						
ADHD symptom severity								
Childhood symptoms ^e	45.1 ± 14.3	34.9 ± 13.9	15.3 ± 8.7	15.8 ± 11.1	18.1 ± 11.9	15.6 ± 11.2	16.5 ± 10.3	14.4 ± 8.4
Current symptoms ^f	34.0 ± 9.2	28.8 ± 9.1	11.6 ± 7.3	11.3 ± 6.9	11.5 ± 7.9	10.4 ± 6.2	11.4 ± 7.0	11.8 ± 9.5

Note. ADHD1 = first sample of patients with attention-deficit/hyperactivity disorder (ADHD); ADHD2 = second sample of patients with ADHD; CG = control group; NSG = naive simulation group; SSG = symptom-coached simulation group; TSG = test-coached simulation group; SG2 = Simulation Group 2, consisting of naive simulators ($n = 24$), symptom-coached simulators ($n = 19$) and symptom- and test-coached simulators ($n = 26$).

^a Multiple Choice Vocabulary Test. ^b Number of patients with specific subtype of ADHD: inattentive/hyperactive-impulsive/combined. ^c Number of patients affected with psychiatric comorbid disorders (some patients affected with more than one): mood disorders/personality disorders/anxiety disorders/eating disorders/posttraumatic stress disorders/substance abuse disorders (not in previous 6 months)/somatoform disorders. ^d Number of participants taking medication at the time of the assessment: Antidepressants/stimulants. ^e Wender Utah Rating Scale—short version. ^f ADHD Self-Report Scale.

suggestions of how to fake ADHD. The SSG received a description of the diagnostic criteria for ADHD as outlined in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM-IV*; American Psychiatric Association, 1994). Participants of the TSG were informed about how a diagnostic clinical examination of ADHD is commonly performed. Participants of the TSG were also informed about typical behaviors of patients with ADHD when they are being tested. Participants of the STSG received both a description of the diagnostic criteria for ADHD and information on how a diagnostic clinical examination of ADHD is usually performed. Participants of all simulation groups received the respective information by the experimenter on two occasions; that is, by email on the day before the assessment and at the beginning of the assessment. Prior to assessments, participants of the SSG, TSG, and STSG were requested to respond to a number of questions on the content of information they were provided with, to ascertain that they indeed read but also understood the information. Finally, participants were requested to start feigning ADHD and to perform the assessment as if they would seek a diagnosis of ADHD. At the end of the assessment, all participants of simulation groups indicated that they followed instructions, including instructions to feign ADHD. The assessment of participants of the simulation groups took about 70 min.

Participants

Patients with ADHD. The sample of ADHD1 included 48 adults diagnosed with ADHD. Current treatment with ADHD medication (i.e., stimulant medication) was an exclusion criterion as individuals of simulation groups were trying to simulate test performance of patients with ADHD *not* being treated with such medication. This seems relevant considering that those attempting to feign ADHD at clinical evaluation do so to get access to certain incentives, such as stimulant medication. Thus, the level of performance that individuals feigning ADHD try to mimic at clinical evaluation is the performance of (still untreated) genuine patients with ADHD. Because stimulant medication is known to affect cognitive performance, we only included patients with ADHD if

they were not treated with stimulants at the time of assessments. Patients were referred from local psychiatrists or neurologists to the Department of Psychiatry and Psychotherapy of the SRH (Stiftung Rehabilitation Heidelberg) Clinic Karlsbad-Langensteinbach, Germany. A diagnostic assessment of adult ADHD as well as participation in the research study was offered to all participants. Diagnostic assessments were performed by experienced clinicians associated to the Department of Psychiatry and Psychotherapy and involved a clinical psychiatric interview based on *DSM-IV* criteria for ADHD (American Psychiatric Association, 1994; Barkley & Murphy, 1998), including both the retrospective assessment of childhood symptoms as well as current symptoms. All diagnoses were made by mutual agreement between at least two clinicians who were part of a diagnostic team and experienced in the assessment and treatment of adults with ADHD. The diagnostic assessment also included identifying objective impairments supporting the diagnosis of ADHD (e.g., evidence derived from school reports, failure in academic and/or occupational achievement) and comprised multiple informants, such as employer evaluation and partner- or parent-reports. Moreover, all participants completed two standardized self-report rating scales designed to quantify current and retrospective ADHD symptoms (Wender Utah Rating Scale (German: WURS-K) and ADHD Self-Report Scale (ASR); Rösler, Retz-Junginger, Retz, & Stieglitz, 2008; Ward, Wender, & Reimherr, 1993). After completion of the recruitment and assessment of the patients of ADHD1 as described above, an independent second sample of 27 adults with ADHD (ADHD2) was recruited and assessed.

Healthy individuals. The HCG ($n = 48$) was recruited via public announcements and word-of-mouth. Healthy individuals were selected according to age, gender and vocabulary skills (intellectual functions), resulting in comparable characteristics to the patients of ADHD1.

Furthermore, 252 first-year psychology students of the University of Groningen, the Netherlands, took part in the study in exchange for course credits. Prior to assessment, these participants were randomly assigned to one of four conditions; the CG, the

NSG, the SSG, or the TSG. All three simulation groups were collapsed and formed *Simulation Group 1* (SG1).

After completion of the recruitment and assessment of participants described above, an independent sample of first-year psychology students of the University of Groningen, the Netherlands, were recruited and formed *Simulation Group 2* (SG2, $n = 69$). Participants of the SG2 were randomly assigned to one of three conditions, that is, a naïve simulation condition ($n = 24$), a symptom-coached simulation condition ($n = 19$), or a symptom- and test-coached simulation condition ($n = 26$).

Materials

Vocabulary skills (intellectual functions). Vocabulary skills (intellectual functions) were measured using the Multiple Choice Vocabulary Test (MWT-B; Lehl, 1995). This test consists of 37 lines, each comprising one authentic word and four fictitious words. The participants were required to find the authentic word by underlining it. The MWT-B was suggested to provide a useful measure for premorbid IQ that is relatively insensitive to cerebral dysfunction of individuals (Lehl, 1995; Lehl, Triebig, & Fischer, 1995). In studies on typically developing individuals, the MWT-B was shown to correlate fairly well with global IQ, yielding a median correlation of $r = .72$ in 22 samples (Lehl et al., 1995).

Clinical assessment of ADHD symptom severity. Childhood ADHD symptoms were self-rated on the short version of the Wender Utah Rating Scale (WURS-K), which includes 25 items rated on a five-point scale (Ward et al., 1993). Severity of current ADHD symptoms was self-rated with the ADHD Self-Report Scale (ASR; Rösler et al., 2008) consisting of 18 items rated on a four-point scale corresponding to the diagnostic criteria of *DSM-IV* (American Psychiatric Association, 1994; Rösler et al., 2008). A sum score was calculated for each rating scale.

Visuospatial working memory test (VSWMT). A VSWMT was developed that required participants to navigate through a mental image and to count the number of corners this image contained from memory (based on the principles of the task as devised by Brooks, 1968). The VSWMT comprised 60 trials, divided into three blocks of 20 trials each. A fixation cross appeared in the middle of a computer screen at the beginning of each trial, followed by the presentation of a geometric figure (for 2,000 ms). Subsequently, a visual mask (unstructured black and white dots) was displayed for 500 ms to prevent aftereffects. Directly after the presentation of the visual mask, participants were requested to navigate through the mental image from memory and to count all corners the geometric figure contained. The response had to be given verbally and was noted by the experimenter. Furthermore, the response time was registered by measuring the time from the moment the visual mask disappeared to the participant's response. The subsequent trial was initiated by the experimenter. Block 1 contained 8 trials presenting letters, and 12 trials presenting nonletters (straight lines forming geometric figures, see Figure 1 for examples). Block 2 contained the same figures as Block 1, but presented the figures in a view rotated by either 90 or 180 degrees. Block 3 was a repetition of Block 1, containing the same figures in the same spatial orientation. Presentation order of trials was randomized in each block. Two example trials were administered before testing started, to allow participants to familiarize themselves with the task. The VSWMT was conducted on a

Figure 1. Example items of the visuospatial working memory test. Left: Example item of a letter (top: normal view as used for Blocks 1 and 3; bottom: rotated view as used for Block 2); Right: Example item of a nonletter (top: normal view as used for Blocks 1 and 3; bottom: rotated view as used for Block 2).

computer of 15.4 in. screen size, using the presentation software E-Prime 2.0. The number of errors on letters and nonletters were counted for each block. Furthermore, the mean response times for correct responses were measured for letters and nonletters per block. The response times for incorrect responses were not considered for analysis, as no clear instruction was given to participants on how to respond if they failed to keep the geometric figure as a mental image; for example giving no response, giving a best guess as quickly as possible, or giving a best guess after careful consideration. The response time of incorrect responses would therefore not provide a useful measure of task performance. As the VSWMT was specifically designed for the present study and has not been used before, its psychometric properties, such as reliability of scores, are not known yet. In the present study, internal consistency of scores of the VSWMT (Cronbach's alpha) ranged from 0.81 to 0.84 for patients with ADHD, healthy control groups, and instructed simulators.

Routine measures of cognition. Vigilance was measured with the computerized VIGIL test of the Vienna Test System VTS (Schuhfried, 2013). In this test, a white dot moved along a circular path (resembling the second hand of an analog clock) in small regular jumps. Occasionally (1 out of 10 jumps), the white dot made a double jump (critical stimulus). The participants were requested to react to this infrequent event of a double jump as quickly as possible with a button press on a response panel. The number of correct responses (to critical stimuli), the number of incorrect responses (to noncritical stimuli, i.e., to events without a double jump), and the mean reaction time (RT; ms) for correct responses were measured.

The N-Back Verbal (NBV) of the VTS (Schuhfried, 2013) was applied as a test for verbal working memory. In this test, letter sequences are presented on a computer screen and participants are required to respond by pressing a button when the letter displayed at that time is identical to the second to last letter. The number of errors, as well as the mean response time to correct and incorrect responses were registered.

Tests for performance validity. Three established effort tests were applied, that is the Test of Memory Malingering (TOMM; Tombaugh, 1996), the Dot Counting Test (DCT; Boone, Lu, & Herzberg, 2002b), and the b Test (Boone, Lu, & Herzberg, 2002a). The TOMM is a visual recognition test consisting of two learning trials and a retention trial. The DCT is a short test requesting participants to visually perceive patterns of dots and use elementary counting and multiplication skills. The b Test is a letter-recognition and discrimination task that asks participants to circle, as quickly as possible, all the b's that appear on several pages of a stimulus booklet. All three tests measure skills that are assumed to be preserved in most individuals with cognitive dysfunction. By employing recommended cutoffs of the TOMM (suspect effort if Trial 2 or Retention score <45), the DCT (suspect effort if E-Score >13), and the b Test (suspect effort if E-Score >119) as specified in the test manuals (Boone et al., 2002a, 2002b; Tombaugh, 1996), these tests can be applied to distinguish between credible and suspect effort among the validation sample of patients with ADHD.

Statistical Analysis

Neuropsychological test performance on vigilance and verbal working memory of the patients of ADHD1 was analyzed using descriptive statistics. A neuropsychological impairment was determined on the basis of a commonly accepted categorization of ability levels using age-based test norms as provided by the test publisher (percentile ≤ 10 ; Lezak, Howieson, & Loring, 2004). VSWMT scores (total number of errors and total mean response time of correct responses) of the patients of ADHD1 and the HCG were compared using multivariate analysis of variance (MANOVA). Effect sizes of univariate group differences were indicated by Cohen's d and were classified, based on Cohen's classification, into negligible effects ($d < 0.20$), small effects ($0.20 \leq d < 0.50$), medium effects ($0.50 \leq d < 0.80$), and large effects ($d \geq 0.80$; Cohen, 1988). The association between visuospatial working memory performance (VSWMT) and (a) verbal working memory (NBV) and (b) vigilance (VIGIL) was explored by Pearson correlations. It is expected that visuospatial working memory is significantly associated with verbal working memory in both error variables and RT measures. However, visuospatial working memory is expected not to be significantly associated with vigilance performance, that is, neither in accuracy nor RTs. Correlation coefficients were interpreted as negligible ($r < .1$), small ($0.1 \leq r < .3$), medium ($0.3 \leq r < .5$), and large ($r \geq .5$; Cohen, 1988).

Furthermore, MANOVA was applied to compare VSWMT performance between the patients of ADHD1, the CG, and all simulation groups. Statistical significance was calculated for the multivariate comparison as well as for univariate comparisons of each variable of the VSWMT. Pairwise group comparisons were performed by using Tukey's Honest Significant Difference (HSD) post hoc tests. A rigorous alpha level was set at .01 to control for

the problem of multiple comparisons. However, interpretations were largely based on effect sizes, as effect sizes indicate the magnitude of an effect independently from the significance and are also considered to be more informative in malingering research compared to statistical significance (Rogers, 2008). Effect sizes were indicated by Cohen's d and η^2 . The index η^2 was used as an effect size estimator for (M)ANOVAs, providing information about the proportion of variance which is accounted for by the factor group membership. According to Cohen (1988), a small effect size corresponds to an $\eta^2 = .0099$, a medium effect size to an $\eta^2 = .0588$, and a large effect size to an $\eta^2 = .1379$. The effect size Cohen's d was computed for pairwise comparisons between simulation groups and patients with ADHD. As Cohen's d (and its categorization) was designed to consider relatively small effects as relevant to research (Cohen, 1988), it was stressed that more rigorous standards are needed for the assessment of malingering (Rogers, 2008). Rogers (2008), therefore, introduced a categorization of effects sizes (Cohen's d) suitable for malingering research and distinguished between moderate effects ($0.75 \leq d < 1.25$), large effects ($1.25 \leq d < 1.50$), and very large effects ($d \geq 1.50$).

In addition, binary logistic regression analysis was carried out to determine the utility of the VSWMT in predicting feigned ADHD relative to true ADHD (ADHD1). All simulation conditions (NSG, SSG, and TSG) were collapsed for the purpose of this analysis into one feigning group (i.e., SG1) as it likely remains unknown in any given assessment context if and how individuals had prepared themselves prior to the diagnostic evaluation. The VSWMT index was calculated for each participant by summing the VSWMT scores that were individually weighted with their logistic regression coefficients, to achieve the best combination of VSWMT scores to predict feigned ADHD. The accuracy of the VSWMT index in detecting individuals feigning ADHD (i.e., SG1, $n = 204$) relative to patients with ADHD (ADHD1, $n = 48$) was explored in receiver operating characteristics (ROC). A ROC curve plots the sensitivity against '1—specificity' at each level of the VSWMT index to predict the criterion (feigned ADHD). ROC analysis allows for determination of diagnostic accuracy as measured by the area under the curve (AUC), indicating the probability that a score drawn at random from the first sample is higher than a score drawn at random from the second sample (Rice & Harris, 2005). For further examination of the internal validity of the interpretations of scores from the prediction model, a bootstrap resampling procedure was applied. A bootstrap resampling procedure allows assigning measures of accuracy (e.g., confidence intervals) to AUCs as derived from ROC analyses. From the original dataset of 252 participants (204 instructed simulators of SG1 and 48 patients with ADHD), 3000 random samples were drawn with replacement for bootstrap analysis. Bootstrap AUC and its associated 90% confidence interval were estimated for the prediction model using Skalská & Freylich's bootstrap application (Skalská & Freylich, 2006).

Finally, results as derived from the analyses described above were cross-validated on independent samples of patients with ADHD (i.e., ADHD2, $n = 27$) and instructed simulators (SG2, $n = 69$). VSWMT performance of the independent group of patients with ADHD (validation sample, $n = 27$) was compared to VSWMT performance of the patients of ADHD1 ($n = 48$) by means of multivariate and univariate ANOVA. In addition, cutoff values of the VSWMT, as derived from the patients of ADHD1,

were also applied to the validation sample of patients. A comparable performance and similar classification rates (i.e., specificity) of both groups would support the credibility of test performance of the patients of ADHD1. Moreover, for a complete cross-validation, ROC analysis was also performed using SG2. Furthermore, cutoff values of the VSWMT as derived from SG1 were applied on SG2 to determine sensitivity rates on a new and independent sample of individuals instructed to feign ADHD.

Results

Characteristics of Participants

Characteristics of participants are presented in Table 2. An official diagnosis of ADHD in childhood had been reported by some of the patients with ADHD, however, diagnostic veracity of ADHD in childhood could not be confirmed in the majority of cases due to incomplete medical records. Therefore, presence of childhood ADHD was assessed retrospectively in all cases. Patients with ADHD suffering from comorbid psychiatric disorders were not excluded because comorbidity is very prevalent among patients with ADHD and is therefore representative for the clinical picture of this condition (Biederman et al., 1993). Statistics as presented in Table 3 reveal that participants of ADHD1 and the HCG did not differ significantly with regard to age, gender, and vocabulary skills. However, as expected, patients with ADHD scored significantly higher on self-report scales for both current ADHD symptoms and retrospective ADHD symptoms.

Healthy participants of SG1 and CG (175 female, 77 male) had a mean age of 21.1 years ($SD = 2.2$ years) and a mean score on vocabulary skills of 100.3 ($SD = 9.4$). After group allocation (CG, NSG, SSG, TSG), descriptive variables did not differ significantly between groups, neither with regard to age, gender, vocabulary skills, nor ADHD symptom severity (see Table 3). None of the healthy individuals reported to have a history of neurological or psychiatric diseases and none were taking any medication known to affect the central nervous system.

Furthermore, characteristics of SG2 were comparable to characteristics of other simulation groups, although it must be noted that the test for vocabulary skills was not applied in this group (see Table 2). However, given the similar process of recruitment among the same population (first-year psychology students), it can be assumed that educational level and vocabulary skills of participants of SG2 were comparable to participants of SG1.

Neuropsychological Functions of Patients With ADHD

Neuropsychological test performance of patients with ADHD (ADHD1, $n = 48$) as assessed with routine measures of cognition are presented in Table 4. Cognitive impairments of patients with ADHD were most prominent in vigilance, with more than half of the patients with ADHD exhibiting impairments. In contrast, only a small number of patients with ADHD displayed impairments in verbal working memory (% of patients with impairment <10).

A correlation analysis between patients' visuospatial working memory (as assessed by the VSWMT) and verbal working memory (as assessed by the NVB) revealed a significant association, of medium size, between the error variables of both tests (see Table 5). Nonsignificant correlations of negligible to small size were found between the pairs of variables representing response times. Moreover, nonsignificant associations of small size were found between the VSWMT and vigilance test (VIGIL), both with regard to accuracy of responses and RTs (see Table 5).

The statistical comparison of visuospatial working memory performance of patients with ADHD (ADHD1, $n = 48$) with that of the healthy comparison group revealed a significantly lower performance of large size among patients with ADHD, Wilk's $\lambda = 0.825$, $F(2, 93) = 9.847$, $p < .001$, $\eta^2 = .175$ (see Table 6). Univariate comparisons demonstrated medium and significant effects regarding errors, $F(1, 94) = 8.170$, $p = .005$, $d = 0.74$, and total mean response time of correct responses, $F(1, 94) = 12.617$, $p = .001$, $d = 0.72$, indicating a poorer visuospatial working memory performance in patients with ADHD.

Utility of the VSWMT for the Detection of Feigned Adult ADHD

Group comparisons between patients with ADHD and instructed simulators. Table 6 presents VSWMT performances of all given variables per group. A multivariate comparison (MANOVA) of VSWMT performance between patients of ADHD1, the control group, and all simulation groups revealed a large significant effect (Wilk's $\lambda = .420$, $F(65, 1369.1) = 5.088$, $p < .001$, $\eta^2 = .159$). Univariate comparisons indicated significant differences between groups on each of the variables of the VSWMT, including *B1-L-Errors* ($F(5, 358) = 14.132$, $p < .001$, $\eta^2 = .165$), *B1-L-RT* ($F(5, 358) = 9.336$, $p < .001$, $\eta^2 = .115$), *B1-NL-Errors* ($F(5, 358) = 14.144$, $p < .001$, $\eta^2 = .165$), *B1-NL-RT* ($F(5, 358) = 13.442$, $p < .001$, $\eta^2 = .158$), *B2-L-Errors* ($F(5, 358) = 18.361$, $p < .001$, $\eta^2 = .204$), *B2-L-RT* ($F(5, 358) = 9.673$, $p < .001$, $\eta^2 = .119$),

Table 3
Statistical Comparisons of Characteristics Between Groups

Characteristic	Comparison of first sample of patients with ADHD (ADHD1) with healthy comparison group (HCG)	Comparison of groups in simulation design (CG, NSG, SSG, TSG)
Age	$t(94) = .027$; $p = .979$	$F(3, 248) = 2.047$; $p = .108$
Gender	$\chi^2(1) = 1.042$; $p = .307$	$\chi^2(3) = 6.621$; $p = .085$
Vocabulary skills	$t(86) = -.856$; $p = .394$	$F(3, 248) = 1.320$; $p = .269$
ADHD childhood symptoms	$t(93) = 12.278$; $p < .001$	$F(3, 248) = .700$; $p = .553$
ADHD current symptoms	$t(94) = 13.252$; $p < .001$	$F(3, 248) = .436$; $p = .727$

Note. ADHD = attention-deficit/hyperactivity disorder; ADHD1 = first sample of patients with ADHD; CG = control group; NSG = naive simulation group; SSG = symptom-coached simulation group; TSG = test-coached simulation group.

Table 4

Routine Measures of Cognition as Performed by Patients With ADHD (ADHD1; n = 48)

Test measure	<i>M</i> ± <i>SD</i>	Percentile ≤ 10
Vigilance		
VIGIL—Correct reactions	83.7 ± 15.8	65.9
VIGIL—Incorrect reactions	6.1 ± 5.8	50.0
VIGIL—RT (ms)	646 ± 114	54.5
Verbal working memory		
NBV—Errors	3.57 ± 3.07	4.5
NBV—RT correct (ms)	763 ± 178	6.8
NBV—RT errors (ms)	906 ± 294	7.5

Note. Cognitive performance is reported in *M* ± *SD*; For each test variable, the percentages of patients scoring within the lowest 10% of test norms are displayed; VIGIL = vigilance test; NBV = N-Back Verbal; RT = response time; ADHD1 = first sample of patients with attention-deficit/hyperactivity disorder (ADHD).

B2-NL-Errors ($F(5, 358) = 21.222, p < .001, \eta^2 = .229$), *B2-NL-RT* ($F(5, 358) = 11.480, p < .001, \eta^2 = .138$), *B3-L-Errors* ($F(5, 358) = 18.967, p < .001, \eta^2 = .209$), *B3-L-RT* ($F(5, 358) = 6.464, p < .001, \eta^2 = .083$), *B3-NL-Errors* ($F(5, 358) = 19.927, p < .001, \eta^2 = .218$), *B3-NL-RT* ($F(5, 358) = 13.192, p < .001, \eta^2 = .156$), *VSWMT Errors* ($F(5, 358) = 30.883, p < .001, \eta^2 = .301$), *VSWMT RT Correct* ($F(5, 358) = 16.845, p < .001, \eta^2 = .190$), and *VSWMT index* ($F(5, 358) = 38.449, p < .001, \eta^2 = .349$). Effect sizes ranged from medium to large. Post hoc pairwise comparisons (Tukey's HSD) denoted that simulation groups committed significantly more errors and had shorter response times on the VSWMT compared to patients with ADHD. When compared to control participants, simulation groups committed significantly more errors, but did not differ in response times. As a consequence, participants of simulation groups had significantly larger scores on the VSWMT index than patients with ADHD and the control group.

Logistic regression analysis for the prediction of feigned ADHD. The utility of the VSWMT to predict feigned ADHD relative to true ADHD was examined by means of a binary logistic regression model. Based on all variables of the VSWMT as predictors, a model was derived that significantly predicts feigned ADHD (collapsed group of SG1, $n = 204$) relative to true ADHD (ADHD1, $n = 48$), $\chi^2(12, N = 252) = 107.174, p < .001$, explaining 34.6% of the variance (Cox & Snell R^2).

Parameters of the logistic regression model are presented in Table 7. To determine the best combination of VSWMT scores to predict feigned ADHD for each individual, unstandardized regression coefficients as derived from the model were used to weight VSWMT scores of each individual. A VSWMT index was calculated by summing up these weighted VSWMT scores with a greater total value indicating a greater likelihood for feigning ADHD.

Classification accuracies. A ROC analysis was performed to determine the accuracy of the VSWMT index in the identification of feigned ADHD (SG1, $n = 204$) relative to ADHD (ADHD1, $n = 48$). Data analysis supported the utility of the VSWMT index in predicting feigned ADHD, $AUC = 0.902, SE = 0.024, CI = 0.855; 0.949, p < .001$. Internal validity of test score interpretation derived from the prediction model was supported by the bootstrap

resampling procedure, which yielded high diagnostic accuracy close to the original AUC estimation, bootstrap $AUC = 0.893, 90\%-CI = 0.851; 0.932$. An inspection of classification statistics for various cutoffs of the VSWMT index as presented in Table 8 reveals adequate diagnostic accuracy at a cutoff of 2.17, reaching excellent specificity toward true ADHD (95.8%) and adequate sensitivity toward feigned ADHD (60.3%). When giving equal weight to false positives and false negatives, similar levels of sensitivity (80.4%) and specificity (81.2%) are obtained at a cutoff of 1.04. Figure 2 shows a graphical plot of the ROC curve, representing a visualization of the diagnostic accuracy of the VSWMT index for the identification of individuals feigning ADHD relative to patients with ADHD (see Figure 2).

Cross-validation of the VSWMT for the prediction of feigned ADHD. To replicate VSWMT performance of the patients of ADHD1, a validation sample of patients with ADHD was recruited, including patients who all 'passed' three established effort tests for performance validity. A multivariate statistical comparison (MANOVA) between both groups revealed a nonsignificant difference in the VSWMT performance between both samples of patients with ADHD, Wilk's $\lambda = .821, F(12, 62) = 1.128, p = .355, \eta^2 = .179$ (see Table 6). When the cutoff values of the VSWMT as derived on the patients of ADHD1 are applied on the validation sample of patients, similar rates of specificity were reached, such as 77.8% on a cutoff of 1.04, and 96.3% on a cutoff of 2.17 (see Table 8).

Furthermore, sensitivity to detect feigned ADHD was compared between the original sample of instructed simulators (SG1) and SG2. A comparison of sensitivity rates as presented in Table 8 reveals that the VSWMT index still achieved satisfactory accuracy rates in detecting noncredible cognitive performance of SG2, although sensitivity rates were lower compared to the sensitivity rates of the original sample (SG1). Similarly, a ROC analysis

Table 5

Pearson Correlations Among Patients With ADHD (ADHD1; n = 48) Between Visuospatial Working Memory Performance (VSWM) and Verbal Working Memory Performance (NBV), as Well as Between VSWM Performance and Vigilance (VIGIL) Performance

Test measure	VSWMT Errors	VSWMT RT correct
Verbal working memory		
NBV Errors	.438*	
NBV RT correct		.224
NBV RT incorrect		-.007
Vigilance		
VIGIL—Correct reactions	-.148	
VIGIL—Incorrect reactions	.181	
VIGIL—RT		-.214

Note. ADHD1 = First sample of patients with attention-deficit/hyperactivity disorder (ADHD); visuospatial working memory performance test (VSWMT) performance is indicated as the total number of errors, as well as the total mean response time of correct responses; N-Back Verbal (NBV) performance is indicated as the total number of errors as well as the mean response time to correct and incorrect responses; vigilance test (VIGIL) performance is indicated as the number of correct responses, the number of incorrect responses, as well as the mean response time (RT) to correct responses.

* $p < .01$.

Table 6
VSWMT Performance of Patients With ADHD, Control Participants, and Simulation Groups

Test variable	HCG (<i>n</i> = 48) <i>M</i> ± <i>SD</i> (Range)	ADHD2 (<i>n</i> = 48) <i>M</i> ± <i>SD</i> (Range)	ADHD2 (<i>n</i> = 27) <i>M</i> ± <i>SD</i> (Range)	CG (<i>n</i> = 48) <i>M</i> ± <i>SD</i>	NSG (<i>n</i> = 68) <i>M</i> ± <i>SD</i>	SSG (<i>n</i> = 68) <i>M</i> ± <i>SD</i>	TSG (<i>n</i> = 68) <i>M</i> ± <i>SD</i>	SG2 (<i>n</i> = 69) <i>M</i> ± <i>SD</i>	ES
31-L-Errors	.81 ± .84 (0–3)	1.62 ± 1.47 (0–6)	1.33 ± 1.30 (0–5)	.83 ± 1.16	2.81 ± 1.97 ^{ab}	.68	.93 [*]	5.88 ± 2.11 ^{ab}	.89 [*]
31-L-RT	4.62 ± 3.35 (.57–15.10)	7.17 ± 3.21 (1.51–13.99)	6.85 ± 4.38 (1.30–21.10)	4.23 ± 2.36 ^a	3.54 ± 2.87 ^a	1.20 [*]	.58	4.15 ± 2.72 ^a	1.03 [*]
31-NL-Errors	3.79 ± 1.50 (1–7)	4.43 ± 2.00 (1–10)	4.19 ± 1.88 (1–8)	3.54 ± 1.69	5.69 ± 2.21 ^{ab}	.59	.85 [*]	6.10 ± 1.90 ^{ab}	.86 [*]
31-NL-RT	3.96 ± 2.92 (.72–13.08)	6.55 ± 2.97 (1.38–14.45)	5.79 ± 3.61 (1.48–16.36)	4.02 ± 2.41 ^a	3.05 ± 2.12 ^a	1.40 ^{**}	.81 [*]	3.42 ± 1.90 ^{ab}	1.30 ^{**}
32-L-Errors	1.06 ± 1.10 (0–4)	1.40 ± 1.41 (0–6)	1.04 ± 1.40 (0–5)	.71 ± 1.10	2.88 ± 2.18 ^{ab}	.78 [*]	1.17 [*]	3.39 ± 2.79 ^{ab}	1.09 [*]
32-L-RT	4.45 ± 3.22 (.48–14.04)	6.36 ± 2.84 (1.39–13.91)	7.45 ± 4.87 (1.44–21.70)	4.09 ± 2.67 ^a	3.25 ± 2.58 ^a	1.16 [*]	.77 [*]	3.52 ± 1.92 ^a	1.17 [*]
32-NL-Errors	3.71 ± 1.43 (1–7)	4.31 ± 2.15 (0–8)	3.81 ± 2.00 (1–9)	3.17 ± 1.86	6.16 ± 2.50 ^{ab}	.78 [*]	1.08 [*]	6.35 ± 2.26 ^{ab}	.92 [*]
32-NL-RT	4.25 ± 3.26 (.68–12.80)	6.05 ± 3.91 (1.31–22.43)	5.91 ± 3.01 (.87–11.34)	4.16 ± 3.16 ^a	2.82 ± 2.00 ^a	1.10 [*]	.79 [*]	3.16 ± 1.89 ^a	1.00 [*]
33-L-Errors	.81 ± 1.05 (0–3)	1.06 ± 1.49 (0–6)	.63 ± 1.12 (0–5)	.90 ± 1.21	2.59 ± 2.10 ^{ab}	.82 [*]	1.13 [*]	3.50 ± 2.08 ^{ab}	1.31 ^{**}
33-L-RT	3.97 ± 2.83 (.47–10.92)	5.26 ± 2.71 (1.00–12.95)	4.89 ± 2.19 (1.23–10.38)	3.34 ± 2.43	2.67 ± 2.21 ^a	1.07 [*]	.38	3.25 ± 2.45 ^a	.79 [*]
33-NL-Errors	2.96 ± 1.53 (0–7)	3.96 ± 2.15 (0–8)	3.52 ± 1.60 (0–7)	3.04 ± 1.96	5.84 ± 2.83 ^{ab}	.73	1.23 [*]	5.81 ± 2.05 ^{ab}	.88 [*]
33-NL-RT	3.87 ± 3.09 (.49–12.50)	5.47 ± 3.05 (.99–13.85)	6.20 ± 3.55 (.60–17.13)	3.58 ± 2.39 ^a	2.64 ± 1.99 ^a	1.14 [*]	.90 [*]	2.79 ± 1.73 ^a	1.13 [*]
VSWMt Errors	13.15 ± 5.00 ^a (4–23)	16.79 ± 7.30 (4–36)	14.52 ± 6.91 (4–32)	12.19 ± 6.16	25.97 ± 10.70 ^{ab}	.97 ^{**}	1.47 ^{**}	28.03 ± 9.41 ^{ab}	1.31 ^{**}
VSWMt RT Correct	4.19 ± 2.78 ^a (83–12.35)	6.14 ± 2.61 (1.44–13.30)	6.18 ± 3.21 (1.82–16.33)	3.85 ± 2.27 ^a	3.00 ± 2.07 ^a	1.36 ^{**}	.80 [*]	3.35 ± 1.90 ^a	1.26 ^{**}
VSWMt index	−42 ± 2.17 (−5.69–3.86)	−1.10 ± 2.38 (−7.94–2.95)	−1.56 ± 2.84 (−8.52–5.0)	−42 ± 2.17	2.77 ± 2.29 ^{ab}	.72	.85	2.98 ± 2.08 ^{ab}	.85
								2.17 ± 1.83 ^{ab}	.52

Note. Variable names of the visuospatial working memory performance test (VSWMT) represent the time blocks of stimulus presentation (B1, B2, and B3), the type of stimulus (letter [L] or nonletter [NL]), and the type of measure (number of errors and response time (in seconds) of correct responses). Moreover, the total number of errors, the total mean response time of correct responses, as well as the VSWMT index are presented. HCG = healthy comparison participants; ADHD1 = ffirst sample of patients with attention-deficit/hyperactivity disorder (ADHD); ADHD2 = second sample of patients with ADHD; CG = control group; NSG = naive simulation group; SSG = symptom-coached simulation group; TSG = Test-coached simulation group; SG2 = Simulation Group 2, consisting of naive simulators ($n = 24$), symptom-coached simulators ($n = 19$) and symptom- and test-coached simulators ($n = 26$). Superscripts are described as follows. The total number of errors and total mean response time of correct responses are statistically compared between the first sample of patients with ADHD (ADHD1) and healthy comparison participants (HCG);^a significant at $p < .01$ when compared to ADHD1. Simulation groups are compared to the first sample of patients with ADHD (ADHD1) and the control group;^a significant at $p < .01$ when compared to ADHD1;^b significant at $p < .01$ when compared to CG. Effect sizes (ES; Cohen's d) of group differences between patients with ADHD1 and simulation groups.

moderate effect according to Roger's categorization. ^{**} large effect according to Roger's categorization. ^{***} very large effect according to Roger's categorization.

moderate effect according to Roger's categorization.

* large effect according to Roger's categorization.

very large effect according to Roger's categorization.

Table 7

Summary Statistics of Binary Logistic Regression Model That Predicts Feigned ADHD (SG1; $n = 204$) Relative to ADHD (ADHD1; $n = 48$) Based on the VSWMT

Predictor variable	B	SE B	Wald	p	Odds ratio
B1-L-Errors	-.150	.191	.620	.431	.860
B1-L-RT	-.044	.115	.148	.701	.957
B1-NL-Errors	.111	.146	.576	.448	1.117
B1-NL-RT	-.324	.152	4.55	.033	.723
B2-L-Errors	.466	.221	4.46	.035	1.594
B2-L-RT	.002	.151	<.001	.989	1.002
B2-NL-Errors	.275	.139	3.943	.047	1.317
B2-NL-RT	-.158	.165	.920	.338	.854
B3-L-Errors	.138	.154	.804	.370	1.148
B3-L-RT	.222	.136	2.663	.103	1.249
B3-NL-Errors	.177	.147	1.443	.230	1.194
B3-NL-RT	-.332	.199	2.803	.094	.717

Note. SG1 = first simulation group, consisting of naive simulation group, symptom-coached simulation group, and test-coached simulation group; ADHD1 = first sample of patients with attention-deficit/hyperactivity disorder (ADHD); variable names of the visuospatial working memory performance test (VSWMT) represent the time blocks of stimulus presentation (B1, B2, and B3), the type of stimulus (letter [L] or nonletter [NL]), and the type of measure (number of errors and response time of correct responses).

showed that the VSWMT index was also successful in the identification of instructed simulators of SG2 relative to ADHD1, $AUC = 0.867$, $SE = 0.033$, $CI = 0.802;0.932$, $p < .001$ (see Figure 3). As compared to diagnostic accuracy of SG1 (90.2%), diagnostic accuracy of SG2 was lower, but still satisfactory.

Discussion

Cognitive Functions of Patients With ADHD

Neuropsychological measures of patients with ADHD largely confirm previous research in showing that the majority of patients

Table 8

Classification Statistics for the Identification of Instructed Simulators Relative to Patients With ADHD for Various Cutoffs of the VSWMT Index

Cutoff	Sensitivity		Specificity	
	SG1 ($n = 204$)	SG2 ($n = 69$)	ADHD1 ($n = 48$)	ADHD2 ($n = 27$)
-0.42	95.1	92.8	66.7	63.0
0.20	90.2	85.3	70.8	74.1
0.69	85.3	76.8	75.0	77.8
1.04	80.4	73.9	81.2	77.8
1.52	70.1	66.7	85.4	85.2
2.10	60.3	55.1	89.6	96.3
2.17	60.3	50.7	95.8	96.3

Note. SG1 = Simulation Group 1 ($n = 204$), consisting of naive simulation group, symptom-coached simulation group, and test-coached simulation group; SG2 = Simulation Group 2 ($n = 69$), consisting of naive simulators, symptom-coached simulators, and symptom- and test-coached simulators; ADHD1 = first sample of patients with attention-deficit/hyperactivity disorder (ADHD; $n = 48$); ADHD2 = second sample of patients with ADHD ($n = 27$).

Figure 2. Receiver operating characteristics (ROC) curve indicating diagnostic accuracy of the visuospatial working memory test (VSWMT) in identifying feigned ADHD (Simulation Group 1, $n = 204$) relative to ADHD (ADHD1, $n = 48$).

with ADHD demonstrate impairments in some, but not all aspects of attention and/or executive control (Mostert et al., 2015; Thome et al., 2012). With regard to working memory performance of patients with ADHD, impaired performance was found in the newly developed visuospatial working memory test, whereas intact abilities were observed in verbal working memory. This appears

Figure 3. Receiver operating characteristics (ROC) curve indicating diagnostic accuracy of the visuospatial working memory test (VSWMT) in identifying feigned ADHD (Simulation Group 2, $n = 69$) relative to ADHD (ADHD1, $n = 48$).

surprising in the light of the relatively robust findings indicating both verbal (phonological) and visuospatial working memory deficits in adults with ADHD, for example as revealed in the meta-analyses of [Alderson and colleagues \(2013\)](#) who demonstrated working memory deficits of medium size in adults with ADHD. However, even though working memory deficits seem to be very frequent among adults with ADHD, it has been noted that these deficits are not universal ([Alderson et al., 2013](#)) and are thus neither necessary nor sufficient for a diagnosis of ADHD ([Willcutt et al., 2005](#)). Support for the construct validity (convergent validity) of test score interpretation obtained from the newly developed visuospatial working memory test is given by the medium sized correlations between the error variables of both working memory tasks in patients with ADHD. The lack of association between variables representing speed of responses of both working memory tasks could be explained by the fact that the NBV requires fast responses from participants by a sequential stimuli presentation at a fixed pace that is independent from the response of the participants. The VSWMT, in contrast, does not require or instruct participants to respond quickly. Also conform our expectations, no significant association was found between the accuracy of responses as well as RT measures of the visuospatial working memory test and vigilance test, which gives further support to the construct validity (divergent validity). Finally, content validity has been established as the visuospatial working memory test represents all facets of a typical working memory task by requiring to develop, inspect, and navigate through a mental image, thus tapping a cardinal function of visuospatial working memory ([Brooks, 1968](#)).

VSWMT to Detect Feigned Adult ADHD

There is a good indication that group manipulations in the context of the simulation design were successful as all participants of simulation groups were able to respond to questions regarding simulation instructions and also indicated after completion of the study that they followed instructions. Further evidence is given by group statistics demonstrating significantly elevated error rates of all simulation groups compared to the CG.

A comparison of VSWMT performance between the different simulation groups revealed that neither information about the criteria for ADHD (symptom-coaching) as per the *DSM-IV* ([American Psychiatric Association, 1994](#)) nor knowledge about the procedure of a diagnostic clinical examination (test-coaching) benefitted instructed malingerers to feign ADHD more realistically. Similar results have been reported by [Dunn, Shear, Howe, and Ris \(2003\)](#), who found little evidence for the utility of symptom information and coaching in feigning cognitive dysfunctions after brain injury. To examine the robustness of tools against the preparation of individuals attempting to feign ADHD, it is suggested that future research includes simulation conditions that inform individuals about the specific tool that will be used, and present the conduct of this tool in short demonstrations to individuals instructed to feign prior to the assessment.

The utility of the VSWMT in detecting feigning was advocated by moderate to very large group differences (Roger's categorization) between individuals instructed to feign ADHD and patients of ADHD1. The utility of the VSWMT was further highlighted in the ROC analysis by high diagnostic accuracy ($AUC = 90.2\%$),

with excellent specificity (95.8%), and adequate sensitivity (60.3%). Results of internal validation (bootstrap resampling procedure) may be seen as an approximation to external validity (i.e., generalizability) and supports the utility of the derived prediction model on new samples ([Steyerberg et al., 2001](#)). Furthermore, cross-validation obtained stable specificity rates on the independent sample of patients with ADHD, whereas sensitivity dropped from the first to the second group of instructed simulators, though still reaching adequate diagnostic accuracy.

Previous research on patients with brain injury already considered embedded validity measures in working memory tests for the identification of noncredible cognitive performance, and suggested that the Digit Span test of the Wechsler Adult Intelligence Scale could provide such information. In this respect, cutoffs of the scaled Digit Span score and the Reliable Digit Span have been introduced as valid indicators of noncredible cognitive performance in patients with brain injury ([Etherton, Bianchini, Greve, & Heinly, 2005](#); [Greve et al., 2007](#); [Iverson & Franzen, 1994](#); [Mathias, Greve, Bianchini, Houston, & Crouch, 2002](#)). These scores have also been examined for their utility to detect noncredible cognitive performance in the clinical evaluation of adults with ADHD ([Booksh et al., 2010](#); [Harrison et al., 2010](#); [Suhr et al., 2008](#)), but were shown to be less successful as they have been initially proposed for brain injured patients ([Booksh et al., 2010](#); [Harrison et al., 2010](#); [Suhr et al., 2008](#)). Whereas specificity rates toward credible performance of adults with ADHD were high, sensitivity rates toward noncredible cognitive performance of individuals with ADHD were low ([Harrison et al., 2010](#)). Based on the previous literature using working memory tasks as described above, as well as findings of the present study, it may be concluded that visuospatial working memory measures, in contrast to verbal working memory measures, are more useful for the development of validity indicators of neuropsychological test performance of adults with ADHD. Although it remains speculation, one could explain this discrepancy by the rather complex material as used in visuospatial working memory tasks that may be difficult to judge and evaluate by individuals attempting to feign ADHD, in comparison to digits as used in verbal working memory tasks that most people are more familiar with. Individuals are naturally more acquainted with verbal information from many tasks of daily living and might find it relatively easy to judge and evaluate verbal information in cognitive tasks when attempting to feign ADHD.

The present study demonstrates that effort testing can be embedded in routine measures of cognition (here in a working memory test). This is a meaningful finding as effort testing is considered as an essential and routine part of clinical neuropsychological assessment ([Bush et al., 2005](#)). Because effort testing takes time, leading to costs that can largely be avoided when cleverly combining test principles of regular measures of cognition (i.e., measuring a particular cognitive function) with principles of effort testing (i.e., detecting feigning), allowing the clinician to make conclusions about both the integrity of function and effort during assessment. Because clinical services have to be prepared to face increasing numbers of requests for neuropsychological assessments due to the demographical changes of society with an increasing number of older people developing conditions, saving time while maintaining quality of assessments becomes pivotal.

Limitations and Future Directions

As a limitation of this analysis, visuospatial working memory of patients was compared to a matched group of healthy individuals, whereas verbal working memory of patients was evaluated based on test norms. More valid conclusions on verbal and visuospatial working memory performance of adults with ADHD can be drawn if test performances of patients with ADHD in both types of working memory measures are evaluated based on the same reference group; a healthy comparison group that performed the verbal and visuospatial working memory test. To establish construct validity of the VSWMT, it would be relevant to correlate verbal and visuospatial working memory performance not only among patients with ADHD, but also among a group of healthy individuals.

As another limitation of the present study, it must be noted that the patients of ADHD1 did not perform any established measures of performance validity. Thus, it cannot be excluded that some of the patients might have exaggerated or even feigned symptoms, which would have distorted the present analysis and results. To control for this problem, we were very careful in the selection of patients with ADHD for the present study and considered collateral information and objective evidence of impairment to support diagnostic status. Additional support for the credibility of the patients of ADHD1 was given by comparing their test performance to the validation sample of patients with ADHD that comprised patients who all 'passed' three established measures of performance validity.

To further explore the utility of the validity indicator of the VSWMT for the identification of noncredible cognitive performance, more research is needed that examines diagnostic accuracies of the VSWMT in comparison to other measures that were previously suggested for the identification of noncredible symptom reporting and performance, including self-report rating scales, personality inventories, or effort tests that were specifically designed to detect feigning (Musso & Gouvier, 2014; L. Tucha et al., 2015). Such research should make use of different methodology, including *known-groups comparisons*, considering that simulation designs are inherently prone to external validity problems. The implications of the present study would be strengthened if the VSWMT also shows high accuracy in identifying individuals who indeed seek clinical evaluation for ADHD, but who are believed to actually feign symptoms.

Finally, to improve diagnostic accuracy of noncredible cognitive performance at clinical evaluation of adult ADHD, it appears promising to explore the utility of the VSWMT in combination with other routine measures of cognition that are commonly performed within a comprehensive neuropsychological assessment of adults with ADHD. Some evidence is given by previous research suggesting measures of attention and processing speed, such as the Continuous Performance Test or the Stroop Test, representing the most promising means among routine measures of cognition as validity indicators of neuropsychological test performance of adults with ADHD (Booksh et al., 2010; Marshall et al., 2010; Quinn, 2003; Sollman, Ranssen, & Berry, 2010; Suhr et al., 2008). A major advantage of embedded measures of performance validity within routine neuropsychological assessment is low face validity, as also well-prepared individuals attempting to feign ADHD are presumably not able to recognize implemented validity indicators

and, thus, cannot adapt their behavior accordingly. VSWMT performance indicating noncredible performance, in combination with other behavioral characteristics that are uncommon for genuine ADHD (e.g., excessive error rates or contradicting performance on cognitive tests measuring similar constructs), might increase sensitivity toward exaggerated or feigned cognitive impairment at clinical evaluation of adults with ADHD.

References

- Alderson, R. M., Kasper, L. J., Hudec, K. L., & Patros, C. H. G. (2013). Attention-deficit/hyperactivity disorder (ADHD) and working memory in adults: A meta-analytic review. *Neuropsychology*, 27, 287–302. <http://dx.doi.org/10.1037/a0032371>
- American Psychiatric Association. (Ed.). (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Press.
- Barkley, R. A., & Murphy, K. R. (1998). *Attention-deficit hyperactivity disorder: A clinical workbook*. New York, NY: Guilford Press.
- Biederman, J., Faraone, S. V., Spencer, T., Wilens, T., Norman, D., Lapey, K. A., . . . Doyle, A. (1993). Patterns of psychiatric comorbidity, cognition, and psychosocial functioning in adults with attention deficit hyperactivity disorder. *The American Journal of Psychiatry*, 150, 1792–1798. <http://dx.doi.org/10.1176/ajp.150.12.1792>
- Booksh, R., 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, 325–338. <http://dx.doi.org/10.1177/1087054708329927>
- Boone, K. (Ed.). (2007). A reconsideration of the Slick et al. (1999) criteria for malingering neurocognitive dysfunction. *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 29–49). New York, NY: Guilford Press.
- Boone, K., Lu, P., & Herzberg, D. (2002a). *The b Test*. Los Angeles, CA: Western Psychological Service.
- Boone, K., Lu, P., & Herzberg, D. (2002b). *The Dot Counting Test*. Los Angeles, CA: Western Psychological Service.
- Boonstra, A. M., Kooij, J. J. S., Oosterlaan, J., Sergeant, J. A., & Buitelaar, J. K. (2010). To act or not to act, that's the problem: Primarily inhibition difficulties in adult ADHD. *Neuropsychology*, 24, 209–221. <http://dx.doi.org/10.1037/a0017670>
- Boonstra, A. M., Oosterlaan, J., Sergeant, J. A., & Buitelaar, J. K. (2005). Executive functioning in adult ADHD: A meta-analytic review. *Psychological Medicine*, 35, 1097–1108. <http://dx.doi.org/10.1017/S003329170500499X>
- Brooks, L. (1968). Spatial and verbal components of act of recall. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 22, 349–368. <http://dx.doi.org/10.1037/h0082775>
- Bush, S. S., Ruff, R. M., Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., . . . Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity NAN policy & planning committee. *Archives of Clinical Neuropsychology*, 20, 419–426. <http://dx.doi.org/10.1016/j.acn.2005.02.002>
- Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Reviews Neuroscience*, 3, 617–628. <http://dx.doi.org/10.1038/nrn896>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Dunn, T. M., Shear, P. K., Howe, S., & Ris, M. D. (2003). Detecting neuropsychological malingering: Effects of coaching and information. *Archives of Clinical Neuropsychology*, 18, 121–134. <http://dx.doi.org/10.1093/arclin/18.2.121>
- Etherton, J. L., Bianchini, K. J., Greve, K. W., & Heinly, M. T. (2005). Sensitivity and specificity of reliable digit span in malingered pain-

- related disability. *Assessment*, 12, 130–136. <http://dx.doi.org/10.1177/1073191105274859>
- Fuermaier, A. B. M., Tucha, L., Koerts, J., Aschenbrenner, S., Kaunzinger, I., Hauser, J., . . . Tucha, O. (2015). Cognitive impairment in adult ADHD—Perspective matters! *Neuropsychology*, 29, 45–58. <http://dx.doi.org/10.1037/neu0000108>
- Fuermaier, A. B. M., Tucha, L., Koerts, J., Weisbrod, M., Grabemann, M., Zimmermann, M., . . . Tucha, O. (2016). Evaluation of the CAARS infrequency index for the detection of noncredible ADHD symptom report in adulthood. *Journal of Psychoeducational Assessment*, 34, 739–750. <http://dx.doi.org/10.1177/0734282915626005>
- Fuermaier, A. B. M., Tucha, O., Koerts, J., Grabski, M., Lange, K. W., Weisbrod, M., . . . Tucha, L. (2016). The development of an embedded figures test for the detection of feigned attention deficit hyperactivity disorder in adulthood. *PLoS ONE*, 11, e0164297. <http://dx.doi.org/10.1371/journal.pone.0164297>
- Greve, K. W., Springer, S., Bianchini, K. J., Black, F. W., Heiny, M. T., Love, J. M., . . . Ciota, M. A. (2007). Malingering in toxic exposure: Classification accuracy of reliable digit span and WAIS–III Digit Span scaled scores. *Assessment*, 14, 12–21. <http://dx.doi.org/10.1177/1073191106295095>
- 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. <http://dx.doi.org/10.1080/09084281003715642>
- Harrison, A. G., Edwards, M. J., & Parker, K. C. H. (2007). Identifying students faking ADHD: Preliminary findings and strategies for detection. *Archives of Clinical Neuropsychology*, 22, 577–588. <http://dx.doi.org/10.1016/j.acn.2007.03.008>
- Harrison, A. G., Rosenblum, Y., & Currie, S. (2010). Examining unusual digit span performance in a population of postsecondary students assessed for academic difficulties. *Assessment*, 17, 283–293. <http://dx.doi.org/10.1177/1073191109348590>
- Iverson, G. L., & Franzen, M. D. (1994). The recognition memory test, digit span, and knox cube test as markers of malingered memory impairment. *Assessment*, 1, 323–334. <http://dx.doi.org/10.1177/107319119400100401>
- Lange, K. W., Hauser, J., Lange, K. M., Makulski-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, 241–248. <http://dx.doi.org/10.1007/s12402-014-0132-3>
- Lehrl, S. (1995). *Mehrfachwahl–Wortschatz–Intelligenztest MWT-B* [multiple choice vocabulary test]. Ballingen, Germany: Spitta-Verlag.
- Lehrl, S., Triebig, G., & Fischer, B. (1995). Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurologica Scandinavica*, 91, 335–345. <http://dx.doi.org/10.1111/j.1600-0404.1995.tb07018.x>
- Lensing, M. B., Zeiner, P., Sandvik, L., & Opjordsmoen, S. (2013). Adults with ADHD: Use and misuse of stimulant medication as reported by patients and their primary care physicians. *ADHD Attention Deficit and Hyperactivity Disorders*, 5, 369–376. <http://dx.doi.org/10.1007/s12402-013-0116-8>
- Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York, NY: Oxford University Press.
- 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. *The Clinical Neuropsychologist*, 24, 1204–1237. <http://dx.doi.org/10.1080/13854046.2010.514290>
- Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44, 377–384. <http://dx.doi.org/10.1097/01.chi.0000153228.72591.73>
- Mathias, C. W., Greve, K. W., Bianchini, K. J., Houston, R. J., & Crouch, J. A. (2002). Detecting malingered neurocognitive dysfunction using the reliable digit span in traumatic brain injury. *Assessment*, 9, 301–308. <http://dx.doi.org/10.1177/1073191102009003009>
- Mostert, J. C., Onnink, A. M. H., Kleins, M., Dammers, J., Harneit, A., Schulten, T., . . . Hoogman, M. (2015). Cognitive heterogeneity in adult attention deficit/hyperactivity disorder: A systematic analysis of neuropsychological measurements. *European Neuropsychopharmacology*, 25, 2062–2074. <http://dx.doi.org/10.1016/j.euroneuro.2015.08.010>
- 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. <http://dx.doi.org/10.1177/1087054712441970>
- Pella, R. D., Hill, B. D., Shelton, J. T., Elliott, E., & Gouvier, W. D. (2012). Evaluation of embedded malingering indices in a non-litigating clinical sample using control, clinical, and derived groups. *Archives of Clinical Neuropsychology*, 27, 45–57. <http://dx.doi.org/10.1093/arclin/acr090>
- Quinn, C. A. (2003). Detection of malingering in assessment of adult ADHD. *Archives of Clinical Neuropsychology*, 18, 379–395. <http://dx.doi.org/10.1093/arclin/18.4.379>
- Rabiner, D. L. (2013). Stimulant prescription cautions: Addressing misuse, diversion and malingering. *Current Psychiatry Reports*, 15, 375. <http://dx.doi.org/10.1007/s11920-013-0375-2>
- Rapport, L. J., Van Voorhis, A., Tzelepis, A., & Friedman, S. R. (2001). Executive functioning in adult attention-deficit hyperactivity disorder. *The Clinical Neuropsychologist*, 15, 479–491. <http://dx.doi.org/10.1076/cln.15.4.479.1878>
- Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, 29, 615–620. <http://dx.doi.org/10.1007/s10979-005-6832-7>
- Rogers, R. (2008). *Clinical assessment of malingering and deception* (3rd ed.). New York, NY: Guilford Press.
- Rösler, M., Retz-Junginger, P., Retz, W., & Stieglitz, R. (2008). *Homburger ADHS-skalen für Erwachsene. Untersuchungsverfahren zur syndromalen und kategorialen Diagnostik der Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) im Erwachsenenalter* [Homburg ADHD scales for adults. Examination procedures for the syndromic and categorical diagnostic of attention-deficit/hyperactivity disorder in adults]. Göttingen, Germany: Hogrefe.
- Schoechlin, C., & Engel, R. R. (2005). Neuropsychological performance in adult attention-deficit hyperactivity disorder: Meta-analysis of empirical data. *Archives of Clinical Neuropsychology*, 20, 727–744. <http://dx.doi.org/10.1016/j.acn.2005.04.005>
- Schuhfried, G. (2013). Vienna test system (VTS) 8 (Version 8.2.00) [Computer software]. Moedling, Austria: Schuhfried.
- Skalská, H., & Freylich, V. (2006). Web-bootstrap estimate of area under ROC curve. *Austrian Journal of Statistics*, 35, 325–330.
- Sollman, M. J., Ranseen, J. D., & Berry, D. T. R. (2010). Detection of feigned ADHD in college students. *Psychological Assessment*, 22, 325–335. <http://dx.doi.org/10.1037/a0018857>
- Steyerberg, E. W., Harrell, F. E., Jr., Borsboom, G. J., Eijkemans, M. J., Vergouw, Y., & Habbema, J. D. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology*, 54, 774–781. [http://dx.doi.org/10.1016/S0895-4356\(01\)00341-9](http://dx.doi.org/10.1016/S0895-4356(01)00341-9)
- Suhr, J., Hammers, D., Dobbins-Buckland, K., Zimak, E., & Hughes, C. (2008). The relationship of malingering test failure to self-reported symptoms and neuropsychological findings in adults referred for ADHD evaluation. *Archives of Clinical Neuropsychology*, 23, 521–530. <http://dx.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. <http://dx.doi.org/10.1080/09084280701509083>
- Thome, J., Ehli, A. C., Fallgatter, A. J., Krauel, K., Lange, K. W., Riederer, P., . . . Gerlach, M. (2012). Biomarkers for attention-deficit/hyperactivity disorder (ADHD). A consensus report of the WFSBP task force on biological markers and the World Federation of ADHD. *The World Journal of Biological Psychiatry*, 13, 379–400. <http://dx.doi.org/10.3109/15622975.2012.690535>
- Tombaugh, T. (1996). *Test of Memory Malingering (TOMM)*. North Tonawanda, NY: Multi-Health Systems.
- Tucha, L., Fuermaier, A. B. M., Koerts, J., Groen, Y., & Thome, J. (2015). Detection of feigned attention deficit hyperactivity disorder. *Journal of Neural Transmission*, 122, 123–134. <http://dx.doi.org/10.1007/s00702-014-1274-3>
- Tucha, L., Sontag, T. A., Walitza, S., & Lange, K. W. (2009). Detection of malingered attention deficit hyperactivity disorder. *Attention Deficit and Hyperactivity Disorders*, 1, 47–53. <http://dx.doi.org/10.1007/s12402-009-0007-1>
- Tucha, L., Tucha, O., Laufkötter, R., Walitza, S., Klein, H. E., & Lange, K. W. (2008). Neuropsychological assessment of attention in adults with different subtypes of attention-deficit/hyperactivity disorder. *Journal of Neural Transmission*, 115, 269–278. <http://dx.doi.org/10.1007/s00702-007-0836-z>
- Tucha, O., Mecklinger, L., Laufkötter, R., Klein, H. E., Walitza, S., & Lange, K. W. (2006). Methylphenidate-induced improvements of various measures of attention in adults with attention deficit hyperactivity disorder. *Journal of Neural Transmission*, 113, 1575–1592. <http://dx.doi.org/10.1007/s00702-005-0437-7>
- van Ewijk, H., Heslenfeld, D. J., Luman, M., Rommelse, N. N., Hartman, C. A., Hoekstra, P., . . . Oosterlaan, J. (2014). Visuospatial working memory in ADHD patients, unaffected siblings, and healthy controls. *Journal of Attention Disorders*, 18, 369–378. <http://dx.doi.org/10.1177/1087054713482582>
- Ward, M. F., Wender, P. H., & Reimherr, F. W. (1993). The Wender Utah Rating Scale: An aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. *The American Journal of Psychiatry*, 150, 885–890.
- Wechsler, D. (1997). *Wechsler adult intelligence scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.
- 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, 1336–1346. <http://dx.doi.org/10.1016/j.biopsych.2005.02.006>

Received February 6, 2017

Revision received August 4, 2017

Accepted August 15, 2017 ■

APA JOURNALS®
Publishing on the Forefront of Psychology

ORDER INFORMATION

Start my 2018 subscription to
Psychological Assessment®
ISSN: 1040-3590

PRICING

APA Member/Affiliate	\$169
Individual Nonmember	\$318
Institution	\$823

Call **800-374-2721** or **202-336-5600**
Fax **202-336-5568** | TDD/TTY **202-336-6123**

Subscription orders must be prepaid. Subscriptions are on a calendar year basis. Please allow 4-6 weeks for delivery of the first issue.

Learn more and order online at:
www.apa.org/pubs/journals/pas

PASA18