

University of Groningen

Utility of the Structured Interview of Reported Symptoms (SIRS-2) in detecting feigned adult attention-deficit/hyperactivity disorder

Becke, Miriam; Fuermaier, Anselm B. M.; Buehren, Jannes; Weisbrod, Matthias; Aschenbrenner, Steffen; Tucha, Oliver; Tucha, Lara

Published in:
Journal of Clinical and Experimental Neuropsychology

DOI:
[10.1080/13803395.2019.1621268](https://doi.org/10.1080/13803395.2019.1621268)

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:
2019

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

Citation for published version (APA):

Becke, M., Fuermaier, A. B. M., Buehren, J., Weisbrod, M., Aschenbrenner, S., Tucha, O., & Tucha, L. (2019). Utility of the Structured Interview of Reported Symptoms (SIRS-2) in detecting feigned adult attention-deficit/hyperactivity disorder. *Journal of Clinical and Experimental Neuropsychology*, 41(8), 786-802. <https://doi.org/10.1080/13803395.2019.1621268>

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.



Utility of the Structured Interview of Reported Symptoms (SIRS-2) in detecting feigned adult attention-deficit/hyperactivity disorder

Miriam Becke, Anselm B. M. Fuermaier, Jannes Buehren, Matthias Weisbrod, Steffen Aschenbrenner, Oliver Tucha & Lara Tucha

To cite this article: Miriam Becke, Anselm B. M. Fuermaier, Jannes Buehren, Matthias Weisbrod, Steffen Aschenbrenner, Oliver Tucha & Lara Tucha (2019) Utility of the Structured Interview of Reported Symptoms (SIRS-2) in detecting feigned adult attention-deficit/hyperactivity disorder, *Journal of Clinical and Experimental Neuropsychology*, 41:8, 786-802, DOI: [10.1080/13803395.2019.1621268](https://doi.org/10.1080/13803395.2019.1621268)

To link to this article: <https://doi.org/10.1080/13803395.2019.1621268>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 03 Jun 2019.



[Submit your article to this journal](#)



Article views: 252



[View Crossmark data](#)

Utility of the Structured Interview of Reported Symptoms (SIRS-2) in detecting feigned adult attention-deficit/hyperactivity disorder

Miriam Becke^{a*}, Anselm B. M. Fuermaier^{a*}, Jannes Buehren^a, Matthias Weisbrod^{b,c}, Steffen Aschenbrenner^d, Oliver Tucha^b and Lara Tucha^a

^aDepartment of Clinical and Developmental Neuropsychology, University of Groningen, Groningen, The Netherlands; ^bDepartment of Psychiatry and Psychotherapy, SRH Clinic Karlsbad-Langensteinbach, Karlsbad, Germany; ^cDepartment of General Psychiatry, Center of Psychosocial Medicine, University of Heidelberg, Heidelberg, Germany; ^dDepartment of Clinical Psychology and Neuropsychology, SRH Clinic Karlsbad-Langensteinbach, Karlsbad, Germany

ABSTRACT

Introduction: The Structured Interview of Reported Symptoms (SIRS-2) utilizes various strategies in the detection of simulated psychiatric disorders. The present study aimed to examine which of these strategies proves most useful in uncovering feigned attention deficit hyperactivity disorder (ADHD) in adulthood.

Method: One-hundred seventy-one individuals instructed to feign ADHD were compared to 46 genuine patients with ADHD as well as 99 neurotypical controls in their reports provided on the SIRS-2.

Results: Responses provided by simulators resembled those of genuine patients with ADHD on all SIRS-2 subscales with the exception of a supplementary scale tapping Overly Specified symptom reports, where a moderate effect emerged ($d = 0.88$). Classification accuracy remained low, with particularly poor sensitivity (sensitivity = 19.30%). Sensitivity was higher when the decision rules postulated in the first edition SIRS were applied instead of its successor's decision model, yet this increase in sensitivity came at the price of unacceptably low specificity.

Conclusion: The present results call for a disorder-specific instrument for the detection of simulated ADHD and offer starting points for the development of such a tool.

ARTICLE HISTORY

Received 21 August 2018
Accepted 10 May 2019

KEYWORDS

Attention deficit hyperactivity disorder; faking; feigning; non-credible performance; symptom validity

Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is among the most common neurodevelopmental conditions, exerting a far-reaching impact on the lives of many. The symptom constellation known as ADHD encompasses hyperactivity, impulsivity, and inattention (American Psychiatric Association, 2013) and first manifests in childhood. While it was long thought that symptom presentation and hence the disorder's impact dissipate as children grow older, it is now well established that symptoms persist into adulthood in a substantial number of cases. Up to 30% of those who experienced symptoms of ADHD in childhood still do so in adult age (Kessler et al., 2005), suggesting that between 1% and 6% of the adult population may be affected (Wender, Wolf, & Wasserstein, 2001).


As awareness of adult manifestations of ADHD increased, so did a consideration of challenges unique to these older age brackets. Alongside questions

regarding symptom severity, comorbidity, and treatment options among adults with the condition (Wender et al., 2001), the possibility of individuals feigning it to access external gains attracted attention. In contrast to children, who have been believed not to be incentivized to or capable of simulating the symptoms of ADHD (Jasinski & Ranseen, 2011; Salekin, Kubak, & Lee, 2008), adults may be motivated to fake the disorder for various reasons.

Treatment options and accommodations offered to those with a diagnosis of ADHD present incentives to malingers. Stimulant medication, a commonly chosen treatment for the condition (Sharma & Couture, 2014), has been found to improve cognitive functioning in people with ADHD across various domains, such as vigilance and mental flexibility (Fuermaier et al., 2016b; Tucha et al., 2006). Amidst heated discussions surrounding “cognitive enhancement”, non-medical use of this class of medication has been presumed to result in similar benefits for neurotypical individuals

CONTACT Anselm B. M. Fuermaier  a.b.m.fuermaier@rug.nl  University of Groningen, Faculty of Behavioural and Social Sciences, Department of Clinical and Developmental Neuropsychology, Grote Kruisstraat 2/1, Groningen 9712 TS, The Netherlands

*These authors contributed equally to this work.

 Supplementary material data can be accessed [here](#).

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

(Ilieva & Farah, 2013; Smith & Farah, 2011). Despite a lack of evidence suggesting beneficial effects of stimulant medication among healthy adults (Hall & Lucke, 2010), these supposed favorable effects on cognition make stimulants desirable to students and professionals alike (Rabiner, 2013; Rabiner et al., 2009; Sansone & Sansone, 2011).

Alongside pharmaceutical interventions, accommodations at school or work may be accessible to those with ADHD (Tucha, Sontag, Walitza, & Lange, 2009). Disability status, as granted following the diagnosis of ADHD, can further lead to tax reductions and additional time off work (Bundesministerium der Justiz und für Verbraucherschutz, 2016; Internal Revenue Service, 2016). In some cases, social security benefits may follow (Pulcini et al., 2015; Steyn, Schneider, & McArdle, 2002). Advantages of an established diagnosis of ADHD may further extend beyond the academic or professional domains to the forensic context, where it may result in diminished criminal liability (for discussion see Eme, 2014). Significant rates of illicit stimulant drug use (Advokat, Guidry, & Martino, 2008) and increasing numbers of self-referrals for evaluations of ADHD (Harrison, Edwards, & Parker, 2008) provide evidence that these incentives may suffice to motivate individuals to feign the condition.

While treatment options and accommodations turn ADHD into an attractive target for individuals attempting to feign the condition, the diagnostic process makes it a feasible diagnosis to simulate. An intricate etiology and vague diagnostic criteria, combined with the lack of objective diagnostic markers (Curatolo, 2005; Fuermaier et al., 2012; Thome et al., 2012; Wankerl et al., 2014), necessitate the reliance on a variety of instruments in the diagnostic process. Ease of use and time constraints encountered in clinical practice contribute to the popularity of self-report measures and clinical interviews as important parts of this process. Insights derived from these measures may be supplemented by collateral information gathered by interviewing examinees' relatives or significant others, or by reviewing school and employment records.

Due to their central role in the diagnostic process of ADHD, self-report measures have garnered significant interest with regard to their utility in detecting feigned presentations of the disorder. Results suggest that existing self-report measures are undependable in distinguishing true patients with ADHD from those aiming to simulate the condition (Bryant et al., 2018; Musso & Gouvier, 2014; Tucha & Fuermaier, 2015), possibly due to the highly subjective nature of the reports (Fisher &

Watkins, 2008; Jachimowicz & Geiselman, 2004; Marshall et al., 2010). Using subsets of items from the Conners' Adults ADHD Rating Scale (CAARS) (Conners, Erhardt, & Sparrow, 1999), a self-report measure commonly used in the diagnostic process, an index has been developed for the sole purpose of differentiating genuine patients from individuals feigning ADHD (Suhr, Buelow, & Riddle, 2011). While initially promising, examination of this Infrequency Index has yielded mixed results (Cook et al., 2017; Cook, Bolinger, & Suhr, 2016; Edmundson et al., 2017; Fuermaier et al., 2016a; Harrison & Armstrong, 2016; Walls, Wallace, Brothers, & Berry, 2017). Classification accuracy was promising, yet further validation is needed to support the index' clinical application.

Alongside self-report measures, clinical interviews have been scrutinized with regard to their ability to detect feigned conditions. The Structured Interview of Reported Symptoms (SIRS), now in its second edition (Rogers, Sewell, & Gillard, 2010), has been established specifically for this purpose. Beyond good sensitivity and specificity rates, the SIRS-2 bears the distinct advantage of utilizing a variety of detection strategies pertinent to the fabrication of psychiatric complaints. Its eight primary scales capitalize on the detection of response patterns uncommon among genuine patients (unlikely detection strategies) as well as excessively impaired presentations (amplified detection strategies including floor effects and symptom validity testing, i. e., identification based on below-chance performance) (Rogers et al., 2010). It is a well-validated measure; its usefulness has been attested in the detection of various feigned psychiatric conditions, such as post-traumatic stress disorder, schizophrenia, and mood disorders (Freeman, Powell, & Kimbrell, 2008; Rogers, Kropp, Bagby, & Dickens, 1992), and across ethnicities (Liu et al., 2013; Rogers et al., 2010).

Recent research on the Structured Interview of Reported Symptoms has, however, suggested significant changes in classification accuracy from the original SIRS publication (Rogers, Bagby, & Dickens, 1992) to the instrument's revised, second edition (Rogers et al., 2010). Whereas the SIRS-2 retains the items and scoring of the first edition SIRS, it features new scales and indices as well as a new classification model. The primary goal of these changes was the reduction of false-positive classifications. Yet the comparison of both SIRS editions strongly suggests that the SIRS-2's improved specificity comes at the cost of meaningfully reduced sensitivity among forensic psychiatric patients and disability claimants, as well as individuals presenting with symptoms of

Dissociative Identity Disorder (Brand, Tursich, Tzall, & Loewenstein, 2014; Green, Rosenfeld, & Belfi, 2013; Tarescavage & Glassmire, 2016; Tylicki et al., 2018). While the newly introduced RS Total Score (please refer to Method Section for descriptions of the scales and indices) appears to serve its intended purpose in reducing false-positive rates (Green et al., 2013), the value of the new MT Index and SS Index has been called into question (Green et al., 2013; Tarescavage & Glassmire, 2016).

Efforts made to detect feigned ADHD have gone beyond self-report measures and clinical interviews to examine the utility of personality inventories (Aita, Sofko, Hill, Musso, & Boettcher, 2017; Butcher, 2009; Morey, 1991; Musso, Hill, Barker, Pella, & Gouvier, 2016; Smith, Cox, Mowle, & Edens, 2017; Young & Gross, 2011), cognitive tests as they are employed in routine neuropsychological examinations (Conners et al., 2000; Fuermaier et al., 2018; Morey, 2016, 2017; Quinn, 2003; Suhr, Sullivan, & Rodriguez, 2011), as well as symptom validity tests developed specifically to uncover feigned cognitive impairment (Edmundson et al., 2017; Fuermaier, Tucha, Koerts, Aschenbrenner, & Tucha, 2017a; Green, 2003, 2008; Leppma, Long, Smith, & Lassiter, 2017) in detecting simulated ADHD. While these efforts have yielded promising results, evidence warranting the use of any single instrument in the detection of aggravated or simulated symptoms is yet lacking.

Taken together, testing the validity of reported complaints amongst individuals presenting for the assessment of possible ADHD has been strongly advised (Marshall, Hoelzle, Heyerdahl, & Nelson, 2016), and the American Academy of Clinical Neuropsychology recommends a multi-method approach in doing so (Heilbronner, Sweet, Morgan, Larrabee, & Millis, 2009), employing multiple instruments based on different detection strategies (Bush et al., 2005).

The present study aims to examine the veracity of the SIRS-2 in detecting feigned presentations of ADHD by employing a simulation design. While the SIRS-2 has been developed to detect malingered psychiatric complaints, rather than simulated cognitive symptoms such as the deficits in attention seen in ADHD, it bears the distinct advantage of enquiring about various strategies used in the fabrication of such complaints. Thus, the SIRS-2 may not be expected to distinguish between genuine and feigned instances of ADHD at levels of accuracy akin to those found in other psychiatric conditions, such as post-traumatic stress disorder or schizophrenia, yet it may offer invaluable information on detection strategies most sensitive to feigned ADHD. In examining such strategies, possible starting points for the development of a disorder-specific symptom validity measure may be revealed.

Method

Participants

Adults with ADHD

Fifty adults with ADHD took part in the study. Local psychiatrists or neurologists referred the patients to the Department of Psychiatry and Psychotherapy at the SHR Clinic Karlsbach-Langensteinbach, Germany, where diagnostic assessments were conducted and diagnoses confirmed by at least two experienced clinicians. Diagnoses of ADHD were secured through a comprehensive clinical assessment, which included a psychiatric interview in accordance with the DSM-IV (American Psychiatric Association, 2000) criteria for ADHD. The interview, devised by Barkley and Murphy (1998), encompasses inquiry into past as well as current symptoms.

Additionally, the assessment involved the identification of objective impairments in line with the diagnosis of ADHD, such as records of failure in school or employment. If possible, inquiries were addressed to multiple informants (e.g., evaluations made by employers, reports made by parents or partners). Participants further completed two standardized self-report rating scales measuring both past and present symptoms of ADHD (WURS-K and ASR) (Adler et al., 2006; Kessler et al., 2005; Ward, Wender, & Reimherr, 1993). As all participants scored above the recommended cut-off scores on these scales, severity of their past and present symptoms was assumed to be clinically relevant.

All patients underwent examination of their reported cognitive complaints to further validate the veracity of their diagnoses and minimize the likelihood of the patient sample including aggravated or dishonest reports. To this end, an independent test was consulted above and beyond the SIRS-2 examined in the study. The Groningen Effort Test (GET) (Fuermaier et al., 2016c, 2017a) is an established instrument developed for the detection of feigned ADHD (see Materials for description). Three patients scored above the recommended cut-off score, suggesting possible non-credible performance. These patients were excluded from further analyses. One additional patient had to be excluded due to recent medication intake: the participant in question had taken stimulant medication less than 24 h prior to their assessment.

The remaining 46 participants in the *ADHD Group* were, on average, 33.85 years of age ($SD = 10.66$) at the time of participation. The group's average Intelligence Quotient (IQ) was estimated to be 102.16 ($SD = 11.92$) (see the "Materials" section for details on the instrument used to estimate intellectual functioning).

Prevalence of ADHD subtypes differed within the sample of the remaining patients with the disorder: 25

(54.35%) presented with adult ADHD of the combined type, while 20 patients (43.48%) met the criteria for the predominantly inattentive type. One patient (2.17%) exhibited symptoms in line with the hyperactive-impulsive type of ADHD.

Twenty-one patients with ADHD (45.65%) reported psychiatric comorbidities, eight of whom presented with more than one comorbid psychiatric disorder. Depression was most commonly reported ($n = 16$), four patients indicated a comorbid anxiety disorder. Five patients had been diagnosed with a personality disorder. Two participants reported a history of substance abuse, with no such abuse in the past six months. Eating disorders ($n = 2$), post-traumatic stress disorder ($n = 1$), and somatoform disorder ($n = 1$), while less common, were also reported. Since psychiatric comorbidities are highly prevalent among adults with ADHD (Biederman et al., 1993), participants with such additional disorders were not excluded from the current study.

Healthy individuals

Matched control group. Recruited through public announcements, researchers’ contacts, and word-of-mouth, 46 neurotypical individuals were matched to the *ADHD Group* according to age, gender, and intellectual functions. These participants formed the *Matched Control Group*, whose descriptive statistics can be found in *Table 1*. This group was included to determine whether diagnostic status is associated with differential responses to SIRS-inquiries: to answer the question, whether adults with ADHD differ from adults without the condition in their responses to SIRS-items, we matched this Control Group with regard to age, gender, and estimated IQ.

Healthy comparison individuals did not differ significantly from adults with ADHD with regard to age ($t(90) = -0.328; p = .744$), gender ($\chi^2(1) = 1.0; p = 1.0$), or intellectual functions ($t(89) = -1.118; p = .267$). However, healthy comparison individuals differed significantly from adults with ADHD with regard to self-reported past ($t(66.297) = 14.186; p < .001$) and present ($t(71.592) = 20.015; p < .001$) symptoms of ADHD. As expected, patients endorsed higher symptomatology on both accounts than their matched neurotypical peers did.

Healthy student group. The original sample of neurotypical volunteers encompassed 293 first-year psychology students at the University of Groningen, The Netherlands, who participated in exchange for course credit. Following the assessment, 69 of these student volunteers had to be excluded from further analyses due to clinically significant levels of self-reported

Table 1. Means and standard deviations for demographics and levels of self-reported ADHD symptoms by group.

Descriptives	ADHD Group	Matched Control Group	ExperimentalControl Group	Combined Simulation Group	Naïve Simulation Group	Symptom-Coached Simulation Group	Test-Coached Simulation Group
<i>n</i>	46	46	53	171	55	60	56
Sex (F/M)	18/28	18/28	44/9	126/45	44/11	46/14	36/20
Age years (SD)	33.85 (10.66)	34.63 (12.19)	22.74 (4.71)	21.76 (3.92)	21.38 (2.71)	22.33 (5.61)	21.50 (2.37)
IQ ^a	102.16 (11.92)	105.02 (12.53)	99.06 (8.36)	100.12 (8.84)	98.67 (8.43)	100.53 (8.46)	101.02 (9.62)
ADHD Symptoms							
Childhood ^b	46.52 (14.35)	12.93 (7.20)	11.38 (6.59)	12.94 (7.09)	12.87 (6.39)	12.60 (7.54)	13.37 (7.37)
Present ^c	35.22 (7.79)	8.74 (4.45)	8.44 (4.06)	8.65 (4.28)	8.75 (4.33)	8.48 (4.59)	8.73 (3.92)

^aMultiple Choice Vocabulary Test; ^bWender Utah Rating Scale; ^cADHD Self-Report Scale.

ADHD symptoms, as measured by self-report on two standardized scales tapping past and present symptoms of ADHD (WURS-K and ASR) (Adler et al., 2006; Kessler et al., 2005; Ward et al., 1993). The remaining healthy student group consisted of 224 participants with an average age of 21.99 years ($SD = 4.13$) and an estimated IQ of 99.92 ($SD = 8.74$). A total of 170 individuals within the student sample identified as female, 54 were male. These participants were randomly allocated to one of four conditions within the healthy student group: an *Experimental Control Group* and three simulation conditions, encompassing a *Naïve Simulation Group*, a *Symptom-Coached Simulation Group*, and a *Test-Coached Simulation Group*. These groups were included to examine whether the instruction to feign symptoms of ADHD altered participants' responses to SIRS inquiries (i.e., manipulation check). Descriptive statistics for all groups, as well as self-reported levels of past and present ADHD symptomatology, may be found in [Table 1](#).

The *Experimental Control Group* encompassed 53 students with a mean age of 22.74 years ($SD = 4.71$) and a mean IQ of 99.06 ($SD = 8.36$). Simulation groups were composed as follows: the *Naïve Simulation Group* included 55 participants, 60 participants belonged to the *Symptom-Coached Simulation Group*, and the *Test-Coached Simulation Group* included 56 participants.

These groups of healthy individuals were comparable with regard to age ($F(3,218) = 23.288, p = .253$), sex ($\chi^2(3) = 6.122, p = .106$), and estimated intelligence ($F(3,188) = 0.804, p = .493$). None of the healthy individuals reported a history of neurological or psychiatric disorders and none had taken or were taking medications known to affect the central nervous system.

Materials

ADHD symptom severity

The short version of the Wender Utah Rating Scale (WURS-K) was used to measure childhood symptoms of ADHD (Ward et al., 1993). The scale encompasses 25 items which are rated on a five-point scale. Participants indicated the current severity of ADHD symptoms by means of the ADHD Self-Report Scale (ASR) (Adler et al., 2006; Kessler et al., 2005), which includes 18 items rated on a four-point scale. Items correspond to the diagnostic criteria for ADHD as defined in the DSM-IV (American Psychiatric Association, 2000). Sum scores were calculated for both rating scales.

Structured interview of reported symptoms (SIRS-2)

The SIRS-2 is a 172-item interview designed to uncover response styles common among those exacerbating or feigning symptoms, each of which is examined in one of eight primary scales: Rare Symptoms (RS), Symptom Combinations (SC), Improbable or Absurd Symptoms (IA), Blatant Symptoms (BL), Subtle Symptoms (SU), Selectivity of Symptoms (SEL), Severity of Symptoms (SEV), and Reported vs. Observed Symptoms (RO). According to Rogers (2008b), two broad categories of detection strategies form the conceptual basis on which each of these response styles is uncovered. Some SIRS-2 scales (i.e. RS, SC, IA, and RO) are based on so-called *unlikely detection strategies*, which aim to identify simulators by the presence of symptoms not commonly reported by genuine patients. *Amplified detection strategies* (i.e. BL, SU, SEL, and SEV scales), on the other hand, encompass plausible items which simulating participants endorse to a larger extent than individuals with a genuine disorder.

If participants answer an inquiry with a "Definite Yes", the answer is scored '2'. Answers of "Qualified Yes/Sometimes" are given a score of '1', whereas "No" is scored as '0'. If no answer is given, "X" is scored for the item in question. Sum scores are calculated and subsequently categorized as falling into the "genuine", "probable feigning" or "definite feigning" range for each of the separate scales. The primary scales are complemented by supplementary scales and index scores aiming to provide further clinical descriptions. Alongside sum scores intended to enable the classification of the examinee's response style, these scales tap honesty towards mental health professionals, defensiveness, and consistency of symptom reports across time. All items, as well as the scoring method, are equivalent to those of the original SIRS (Rogers et al., 1992).

By following a decision model introduced in the interview's manual, the SIRS-2 allows for dichotomous distinctions between genuine responding and probable or definite feigning (Rogers et al., 2010). A three-tiered process is used to reach such a classification. As illustrated on the right-hand side in [Figure 1](#), this process considers the primary scales first, followed by the evaluation of three different sum scores (RS Total, MT Index, and SS Index).

The SIRS-2 has demonstrated good validity and reliability. Internal consistency is high for the primary scales with coefficient alphas ranging from .77 to .92, as is interrater reliability (weighted averages ranging between $\alpha = .95$ to $\alpha = 1.00$, with a mean of $\alpha = .98$) (Rogers et al., 2010). Standard errors of measurement (SEM) further suggest high reliability of individual measurements.

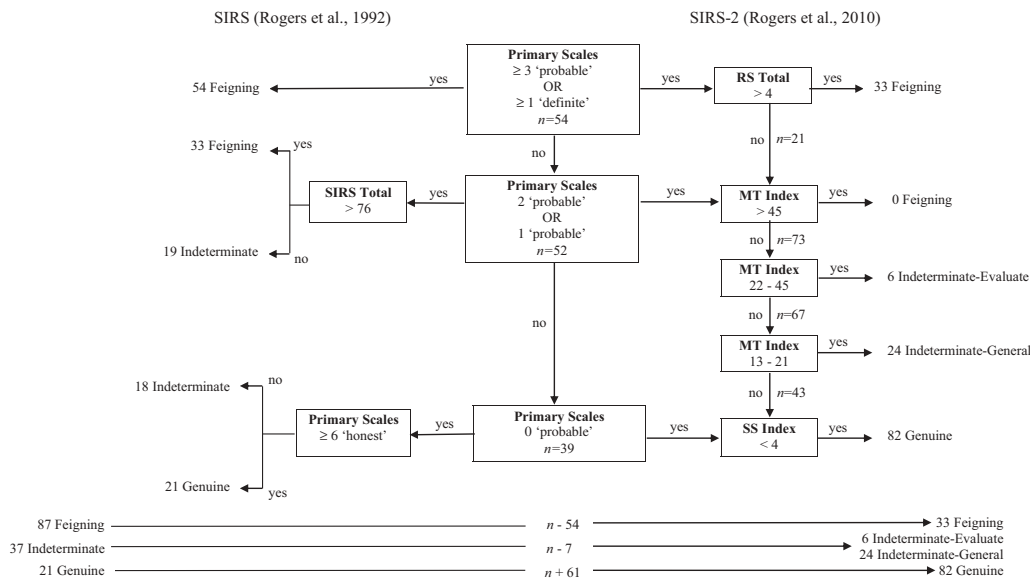


Figure 1. Comparison of classifications yielded by original SIRS decision rules (Rogers, Bagby, et al., 1992) and SIRS-2 decision model (Rogers et al., 2010) for simulating participants ($n = 145$).

Consideration of test–retest reliability ought to be more differentiated: whereas overall test–retest reliability was robust ($r = .82$ for SIRS-2 Total Score) and satisfactory for most scales (*Mean* $r = .77$), significant fluctuations can be observed on a number of scales. In particular, the Symptom Combinations (SC), Improbable or Absurd Symptoms (IA), and Reported vs. Observed Symptoms (RO) Scales may be subject to marked fluctuations across testing sessions. Concordance across repeated testing sessions is, however, high when considering categorization of examinees; there is high agreement in the dichotomous decision of “feigning” or “non-feigning”.

Measures of validity also yield satisfactory results, with convergent validity showing consistent, moderately strong associations with other measures of simulated mental disorders such as the Minnesota Multiphasic Personality Inventory (Butcher, 2009) and the Personality Assessment Inventory (Morey, 1991). Discriminant validity with these measures has been established. Criterion-related validity, as measured by means of effect sizes and accuracy of individual classification, is good (Rogers et al., 2010).

Groningen effort test

Patients’ veracity was assessed by means of the Groningen Effort Test (Fuermaier et al., 2016c, 2017a). The GET was developed to enable the detection of non-credible performance within the diagnostic process of ADHD. The computerized test requires participants to solve a visual

discrimination task designed to appear cognitively taxing, with high demands being exerted on attention and concentration. Most individuals, including those with ADHD, complete the task with ease. Even so, participants are reminded of the test’s supposed significant cognitive demands throughout testing. Mean response times and number of errors (block-wise) are registered as outcome variables. A respective cut-off score allows for the discrimination between credible and non-credible performance with a high degree of accuracy. At the recommended cut-off score, the GET’s sensitivity amounts to 89% and its specificity to 89.5% (Fuermaier et al., 2017b). Given this high classification accuracy, the original cut-off score was retained, rather than choosing a lower cut-off value which would have maximized sensitivity to possible feigning. This approach was chosen as lowering the cut-off score would have resulted in undue loss of specificity, power, and generalizability of the results.

Intellectual functions

Intellectual functions were estimated using Lehl’s Multiple Choice Vocabulary Test (Lehl, Triebig, & Fischer, 1995). The test requires participants to find an authentic target word among three fictitious distractor words. Answers are indicated by underlining the target word in each of the test’s 37 lines. Scores, as derived by summing all correct responses, offer a valid quantification of vocabulary skills transformed to an IQ scale (Lehl et al., 1995).

Design and procedure

Assessment of ADHD group

Patients with adult ADHD were tested individually. They received no reward for participation. Written informed consent was sought from all patients; they were assured that all data collected as part of the research project would be analyzed anonymously, and that these data would not affect clinical assessment or treatment. Patients underwent a comprehensive assessment including standard measures of cognition, a measure of performance validity (i.e., the GET), self-report questionnaires, and the SIRS-2. Completion of the assessment battery took a total of 2 h, divided into two parts to avoid potential effects of fatigue (Lezak, Howieson, Loring, Hannay, & Fischer, 2004). The study complied with the ethical standards of the Helsinki Declaration and was approved by the local institutional ethical committee (Medical Faculty at the University of Heidelberg, Germany).

Assessment of healthy participants

All healthy participants gave written informed consent and were subsequently tested individually in a quiet laboratory. Descriptive and anamnestic information was collected at the beginning of the experiment, prior to the simulation groups being instructed to feign ADHD. Inquiries were made into information such as age, sex, intellectual functions, and self-reported ADHD symptom severity. Additionally, participants were asked about any history of psychiatric or neurological disease, as well as pharmacological treatment. The remainder of the experiment differed between the groups (*Matched Comparison Group* and *Healthy Student Group* allocated either to the *Experimental Control Group*, *Naïve Simulation Group*, *Symptom-Coached Simulation Group*, or *Test-Coached Simulation Group*), with different instructions being given to each. The assessment procedure for healthy participants was approved by the Ethical Committee Psychology (ECP) at the University of Groningen.

Matched control group and experimental control group

Participants in the *Matched Control Group* and the *Experimental Control Group* were instructed to complete all measures, including the SIRS-2, honestly and to the

best of their ability. A day before the assessment, they received an e-mail containing information about the study's clinical significance, but not about its aim. Assessment of a participant took approximately 50 min.

Simulation groups

Participants allocated to either of the three simulation groups (naïve, symptom-coached, or test-coached) were instructed to feign adult ADHD while completing the SIRS-2. To ease participants' adaptation of their roles as adults with ADHD, they were provided with a vignette describing multiple possible incentives (e.g., financial, educational, or vocational accommodations; the prescription of stimulant medication to foster performance at school or work or for recreational use) for someone to feign the condition. The vignette did *not* contain any information on the symptoms or the nature of ADHD. Participants were further explicitly asked to provide believable answers, that is, to feign ADHD in a realistic manner (e.g., avoiding pronounced exaggeration of symptoms). As external incentives have proven effective in simulation designs (Dunn, Shear, Howe, & Ris, 2003; Rogers, Harrell, & Liff, 1993), participants were promised the chance of winning a tablet PC if they feigned the condition most convincingly. Unbeknownst to participants and in line with ethical considerations, the PC was awarded to a participant chosen at random (i.e., irrespective of test performance) from any of the experimental groups.

Information provided next to the vignette differed between groups (an overview can be found in Table 2). The *Naïve Simulation Group* received no further information and no further suggestions on how to feign ADHD convincingly. Symptom-coached participants were provided descriptions of ADHD symptoms as laid out by the DSM-IV (American Psychiatric Association, 2000), equipping participants with enough information to become familiar with the clinical presentation of ADHD (Tucha et al., 2009). Finally, the *Test-Coached Group* of feigning participants was given information on various detection strategies used to uncover feigning. This group received information on general characteristics of neuropsychological assessment, including the role of interviews in the diagnostic process.

Table 2. Information and instructions provided to each group.

	Vignette	Information on Symptoms	Information on Assessment Methods	Instruction
ADHD Group	No	No	No	Honest responding
Matched Control Group	No	No	No	Honest responding
Experimental Control Group	No	No	No	Honest responding
Naïve Simulation Group	Yes	No	No	Feign adult ADHD
Symptom-Coached Simulation Group	Yes	Yes	No	Feign adult ADHD
Test-Coached Simulation Group	Yes	Yes	Yes	Feign adult ADHD

To allow participants sufficient time to prepare strategies for the successful fabrication of symptoms of adult ADHD, information was provided twice; once via mail the day before participation and at the beginning of the assessment. Before the assessment began, participants in both the *Symptom-* and the *Test-Coached Groups* answered questions regarding the information they had been provided with. All participants were able to answer these correctly, so that no participants had to be excluded. Participants were subsequently asked to begin feigning adult ADHD, answering the remainder of the assessment as though they suffered from the disorder. After the assessment, participants were instructed to stop feigning the condition. They were debriefed and asked whether they had followed the given instructions accordingly, including the one asking them to feign ADHD. All participants answered in the affirmative. Assessment of the simulation groups took 70 min. Administrators were aware of the instructions each participant had previously received.

Results

Classification rates based on SIRS and SIRS-2 decision rules

Participants' responses were first classified on the basis of the updated decision model outlined in the SIRS-2 manual (Rogers et al., 2010; see right-hand side of Figure 1). As illustrated in Table 3, this model correctly classified 73.91% of adults with ADHD ($n = 34$) as genuine responders. Misclassification of genuine patients as "Feigning" occurred in 6.52% ($n = 3$) of cases. A total of 33 instructed simulators (19.30%) were correctly identified as such, whereas 59.06% of simulating participants ($n = 101$) were incorrectly classified as honest responders ("Genuine" classification).

As measures of classification accuracy, including estimates of sensitivity and specificity, necessitate a dichotomous classification of either "genuine" or "feigning" (see Green et al., 2013), participants classified as "Indeterminate-General" and "Indeterminate-Evaluate" were combined with the "Genuine" responding group to determine accuracy of classification using the SIRS-2 decision model. This yielded a sensitivity of .19 and a specificity of .93. To account for the increased chance of feigning among

individuals classified as "Indeterminate-Evaluate" (Rogers et al., 2010), accuracy rates were also calculated when counting participants with an "Indeterminate-Evaluate" classification towards the "Feigning" cases. Using this approach, sensitivity amounted to .24, specificity equaled .89.

Estimates of sensitivity and specificity were further calculated in the manner described by Rogers (2010), whereby all "Indeterminate" cases are excluded before carrying out the calculations. Doing so resulted in the exclusion of nine patients and 37 simulators, a sensitivity of .25, and specificity of .92.

Responses provided by 289 participants were re-evaluated according to the first edition of the SIRS' professional manual (Rogers et al., 1992; left-hand side of Figure 1). Classifications based on the original manual were missing for 26 simulating participants and one honest responding neurotypical control, as SIRS Total Scores could not be retrieved for these cases. Applying the decision rules described in the original professional manual, 23.91% ($n = 11$) of patients with ADHD were correctly identified as honest responders (see Table 3). Fifteen adults with ADHD (32.61%) were wrongfully classified as "Feigning" based on the SIRS. Among members of the *Combined Simulation Group*, 60% of cases were correctly identified to be "Feigning" ($n = 87$). Misclassification as honest responding ("Genuine" classification) occurred in 14.48% of cases ($n = 21$).

Sensitivity and specificity were calculated for classifications produced by the original decision rules in the same manner as those yielded by the second edition's decision model. Considering "Indeterminate" cases not to be feigning resulted in a sensitivity of .60 and a specificity of .67. If "Indeterminate" cases were instead considered to be feigning, sensitivity amounted to .86, specificity to .24. Excluding "Indeterminate" cases, as illustrated by Rogers (2010), before calculating accuracy estimates resulted in a sensitivity of .81 and a specificity of .42. In doing so, data of 20 patients and 37 simulators were excluded.

Comparison of SIRS and SIRS-2 classifications

Classifications derived from the SIRS and the SIRS-2 coincided for 56.75% of the 289 participants (see Supplementary Table 2) when "Indeterminate" SIRS-2 classifications (i.e., "Indeterminate-General" and

Table 3. Agreement of SIRS classification and group membership.

Group	SIRS Classification			SIRS-2 Classification			
	Genuine	Indeterminate	Feigning	Genuine	Indeterminate-General	Indeterminate-Evaluate	Feigning
ADHD Group	11 (23.91%)	20 (43.48%)	15 (32.61%)	34 (73.91%)	7 (15.22%)	2 (4.35%)	3 (6.52%)
Combined Simulation Group	21 (14.48%)	37 (25.52%)	87 (60%)	101 (56.06%)	29 (16.96%)	8 (4.68%)	33 (19.30%)
Combined Control Group	93 (94.90%)	4 (4.08%)	1 (1.02%)	98 (99%)	1 (1%)	0 (0%)	0 (0%)

“Indeterminate-Evaluate”) were collapsed into one category. Agreement was perfect for participants whose responses yielded a “Genuine” classification based on the original SIRS decision rules, as all of them were sorted into the same category on the SIRS-2 ($n = 125$, 100%). The number of “Feigning” classifications, however, decreased from 103 to 36 when scoring answers according to the SIRS-2 rather than the SIRS. Thirty participants (29.13%), whose answers resulted in a SIRS-based “Feigning” classification, were sorted into the “Genuine” category by the SIRS-2 decision rules. The remaining SIRS-identified feigning participants were moved into “Indeterminate-General” ($n = 29$, 28.16%) or “Indeterminate-Evaluate” ($n = 8$, 7.77%) classifications based on the SIRS-2.

Agreement between the two SIRS editions was highest amongst honest-responding neurotypical controls ($n = 98$), with classifications coinciding in 94.90% of cases ($n = 93$). Classifications yielded by the SIRS and SIRS-2 agreed in 39.31% of cases in the *Combined Simulation Group* ($n = 145$). Discordance occurred primarily for participants who were classified as “Feigning” based on the original SIRS decision rules ($n = 87$). Twenty-seven (31.03%) of these initially correctly identified simulators were considered to be honest responders based on the SIRS-2 decision model. Additional 31% ($n = 27$) were categorized as “Indeterminate” on the SIRS’ second edition. Accordingly, 37.93% ($n = 33$) of simulators who had been correctly identified by the SIRS were also regarded as “Feigning” on the SIRS-2. Considering the group of adults with ADHD ($n = 46$), classifications based on the SIRS and SIRS-2 converged in 30.43% of

cases. Discrepancies largely arose as all “Indeterminate” cases brought forward by the SIRS decision rules fell into the “Genuine” category when the SIRS-2 decision model was applied ($n = 20$, 100%). Additionally, 12 patients with ADHD considered to be “Feigning” by SIRS standards were sorted into the “Genuine” ($n = 3$, 20%) or “Indeterminate” ($n = 9$, 60%) categories on the SIRS-2.

All agreement rates, as well as further information on the distribution of the SIRS Total, RS Total, and MT Index may be found in the Supplementary Material.

Experimental analyses of SIRS-2 subscales across groups

Means and standard deviations of SIRS-2 scores by group can be found in Table 4. Supplementary Table 1 further lists SIRS-2 scores for each of the three stimulation conditions, which did not differ significantly from one another ($p > .011$ for all SIRS-2 scales). A Bonferroni-corrected significance level of $p = .05/16 = .003$ was considered in the interpretation of all null hypothesis significance tests to account for multiple comparisons. Given the previously described sample sizes for patients with ADHD ($n = 46$) and simulators ($n = 171$), this adjusted significance level was associated with a power of 0.96 or larger for effect sizes of $d \geq 0.75$. In accordance with Roger’s classification of effect sizes in malingering research (Rogers, 2009), an effect of this size (i.e., ≥ 0.75) was considered a moderate one. Values ≥ 1.25 were considered large, whereas those ≥ 1.5 were deemed very large.

The comparison of item endorsement (Kruskal–Wallis Test) revealed significant differences between groups on

Table 4. SIRS scores by group (means and standard deviations) with results yielded by Kruskal–Wallis tests.

	ADHD Group ($n = 46$)	Matched Control Group ($n = 46$)	Experimental Control Group ($n = 53$)	Combined Simulation Group ($n = 171$)	H (df)	p
RS	1.91 (2.49)	0.63 (0.99)	0.42 (0.82) ^b	3.05 (2.95)	76.67 (5)	<.001
SC	1.72 (1.22)	0.39 (0.77) ^a	0.42 (0.84) ^b	3.73 (3.15) ^a	119.14 (5)	<.001
IA	0.41 (0.75)	0.59 (2.96)	0.04 (0.28) ^b	1.05 (1.88)	34.69 (5)	<.001
BL	5.37 (4.29)	0.48 (1.21) ^a	0.55 (0.85) ^b	4.99 (3.84)	137.77 (5)	<.001
MT	9.41 (6.62)	2.09 (3.78) ^a	1.38 (1.80) ^b	12.82 (10.11)	142.87 (5)	<.001
SU	14.00 (5.57)	2.98 (3.05) ^a	3.00 (2.38) ^b	16.32 (5.16) ^a	193.23 (5)	<.001
SEL	14.02 (4.54)	3.07 (2.95) ^a	3.32 (2.51) ^b	14.32 (4.34)	178.34 (5)	<.001
SEV	5.33 (5.53)	0.39 (1.11) ^a	0.21 (0.60) ^b	6.51 (4.58)	141.36 (5)	<.001
RO	3.00 (1.89)	0.72 (1.31) ^a	1.36 (2.08) ^b	4.87 (5.46)	63.69 (5)	<.001
RS-T	3.70 (2.40)	1.15 (1.09) ^a	1.15 (1.05) ^b	4.51 (3.63)	104.77 (5)	<.001
DA	3.85 (2.62)	2.70 (1.74) ^c	2.32 (1.45) ^c	3.58 (3.06) ^c	10.93 (5) ^c	.053
DS	25.50 (5.91)	17.48 (6.29) ^a	16.96 (6.09) ^b	26.07 (6.78)	91.40 (5)	<.001
IF	2.37 (1.55)	1.00 (0.92) ^a	1.06 (0.95) ^b	2.39 (2.22)	49.02 (5)	<.001
OS	0.02 (0.15)	0.24 (0.64)	0.15 (0.46) ^b	1.51 (1.90) ^a	86.69 (5)	<.001
SS	31.74 (7.49)	21.41 (6.93) ^a	20.49 (6.66) ^b	33.56 (8.64)	144.89 (5)	<.001
INC	1.48 (1.67)	0.59 (1.20) ^a	0.58 (1.46) ^b	2.69 (2.56) ^a	81.96 (5)	<.001

SIRS, Structured Interview of Reported Symptoms; RS, Rare Symptoms; SC, Symptom Combinations; IA, Improbable or Absurd Symptoms; BL, Blatant Symptoms; MT (Modified Total Index), sum score of RS, SC, IA, and BL items; SU, Subtle Symptoms; SEL, Selectivity of Symptoms; SEV, Severity of Symptoms; RO, Reported vs. Observed Symptoms; RS-T (RS-Total), Rare Symptoms sum score; DA, Direct Appraisal of Honesty; DS, Defensive Symptoms; IF, Improbable Failure; OS, Overly Specified Symptoms; SS (SS Index), sum score encompassing DA, DS, IF, and OS items; INC, Inconsistency of Symptoms.

^aStatistically significantly different from ADHD Group; ^bStatistically significantly different from Combined Simulation Group; ^cNo pairwise comparisons calculated, as Kruskal–Wallis Test did not reach significance.

all scales, but the index tapping Direct Appraisals of Honesty (DA) ($H(5) = 10.93, p = .053$) (see Table 4). Follow-up analyses (Mann–Whitney U Tests) were conducted to investigate pairwise comparisons.

Comparison of ADHD group and combined simulation group

The *Combined Simulation Group* differed significantly from adults with ADHD on four SIRS-2 scales. On scales assessing Symptom Combinations (SC) ($U = 2421.50; p < .001; d = 0.70$), Subtle Symptoms (SU) ($U = 2806.00; p = .003, d = 0.44$), Overly Specified Symptoms (OS) ($U = 1706.50; p < .001, d = 0.88$), or Inconsistency of Symptoms (INC) ($U = 2701.50; p < .001; d = 0.51$), simulators’ responses yielded significantly higher scores than those of adults with ADHD. No statistically significant differences were detected between the groups on any of the remaining scales ($p \geq .005; 0.01 \leq d \leq 0.40$). The magnitude of differences between patients and simulators was negligible: most comparisons between patients and simulators yielded trivial effect sizes (see Table 5 and Figure 2 for 95%-confidence intervals around the effect sizes). The supplementary scale examining the report of Overly Specified Symptoms (OS) presented the sole exception with a moderate effect based on Roger’s classification ($d = 0.88; 95\%CI[0.55, 1.22]$). Endorsement of individual items composing this scale was further compared between the *ADHD Group* and the

Combined Simulation Group. The sum score for the OS scale is attained by summing up responses provided to seven items (items number 24, 33, 41, 110, 114, 119, and 135). As illustrated in Figure 3, adults with ADHD did not endorse any item belonging to this scale. Among simulating participants, endorsement of item number 135 was most common (“Do you spend much time worrying about your physical health? Would this average between 30 and 40 min each day?”). Supplementary Table 3 presents results of a ROC analysis, which provided limited support for the diagnostic accuracy of the OS scale.

Table 5 illustrates significant differences found between patients with ADHD and their neurotypical counterparts on most SIRS-2 scales, which arose because adults with ADHD endorsed – on average – higher levels of symptomatology than their matched peers. It further shows significant differences between the *Combined Simulation Group* and *Experimental Control Group*, which confirm that the instruction to feign ADHD successfully altered the participants’ responding behavior.

Discussion

Its consideration of various detection strategies constitutes one of the SIRS’ strengths, and has allowed the present study to examine the utility of these very strategies in the detection of feigned adult ADHD. Herein, the focus lay on determining which detection strategies

Table 5. Results of Mann–Whitney U tests and effect sizes (Cohen’s d) for pairwise group comparisons.

Scale	Group Comparison								
	ADHD Group (n = 46) vs. Matched Control Group (n = 46)			Experimental Control Group (n = 53) vs. Combined Simulation Group (n = 171)			ADHD Group (n = 46) vs. Combined Simulation Group (n = 171)		
	U	p	d	U	p	d	U	p	d
RS	749.00	.009	0.67	1618.50	<.001	1.01	2886.00	.005	0.40
SC	421.00	<.001	1.30	1221.50	<.001	1.19	2421.50	<.001*	0.70
IA	920.50	.120	0.08	2892.00	<.001	0.61	3339.00	.066	0.37
BL	213.00	<.001	1.55	885.00	<.001	1.31	3744.50	.617	0.10
MT	213.00	<.001	1.36	569.00	<.001	1.29	3232.00	.063	0.36
SU	71.50	<.001	2.45	118.50	<.001	2.86	2806.00	.003*	0.44
SEL	63.00	<.001	2.86	278.00	<.001	2.76	3771.50	.668	0.07
SEV	394.00	<.001	1.24	668.00	<.001	1.57	3193.50	.050	0.25
RO	339.50	<.001	1.40	2395.50	<.001	0.72	3600.00	.375	0.38
RS-T	330.00	<.001	1.37	1436.00	<.001	1.05	3623.00	.409	0.24
DA	–	–	–	–	–	–	–	–	–
DS	380.00	<.001	1.32	1440.00	<.001	1.38	3668.50	.484	0.09
IF	427.50	<.001	1.07	2805.00	<.001	0.67	3536.50	.285	0.01
OS	940.50	.046	0.48	2275.00	<.001	0.81	1706.50	<.001*	0.88
SS	333.50	<.001	1.43	1053.00	<.001	1.58	3424.00	.178	0.22
INC	656.50	.001	0.61	1719.50	<.001	0.91	2701.50	.001*	0.51

RS, Rare Symptoms; SC, Symptom Combinations; IA, Improbable or Absurd Symptoms; BL, Blatant Symptoms; MT (Modified Total Index), sum score of RS, SC, IA, and BL items; SU, Subtle Symptoms; SEL, Selectivity of Symptoms; SEV, Severity of Symptoms; RO, Reported vs. Observed Symptoms; RS-T (RS-Total), Rare Symptoms sum score; DA, Direct Appraisal of Honesty; DS, Defensive Symptoms; IF, Improbable Failure; OS, Overly Specified Symptoms; SS (SS Index), sum score encompassing DA, DS, IF, and OS items; INC, Inconsistency of Symptoms.

*Statistically significant differences between genuine patients with ADHD and instructed simulators at a significance level of $p \leq .003$; no pairwise comparisons were conducted for the DA scale, as the Kruskal–Wallis Test did not reach significance for this scale.

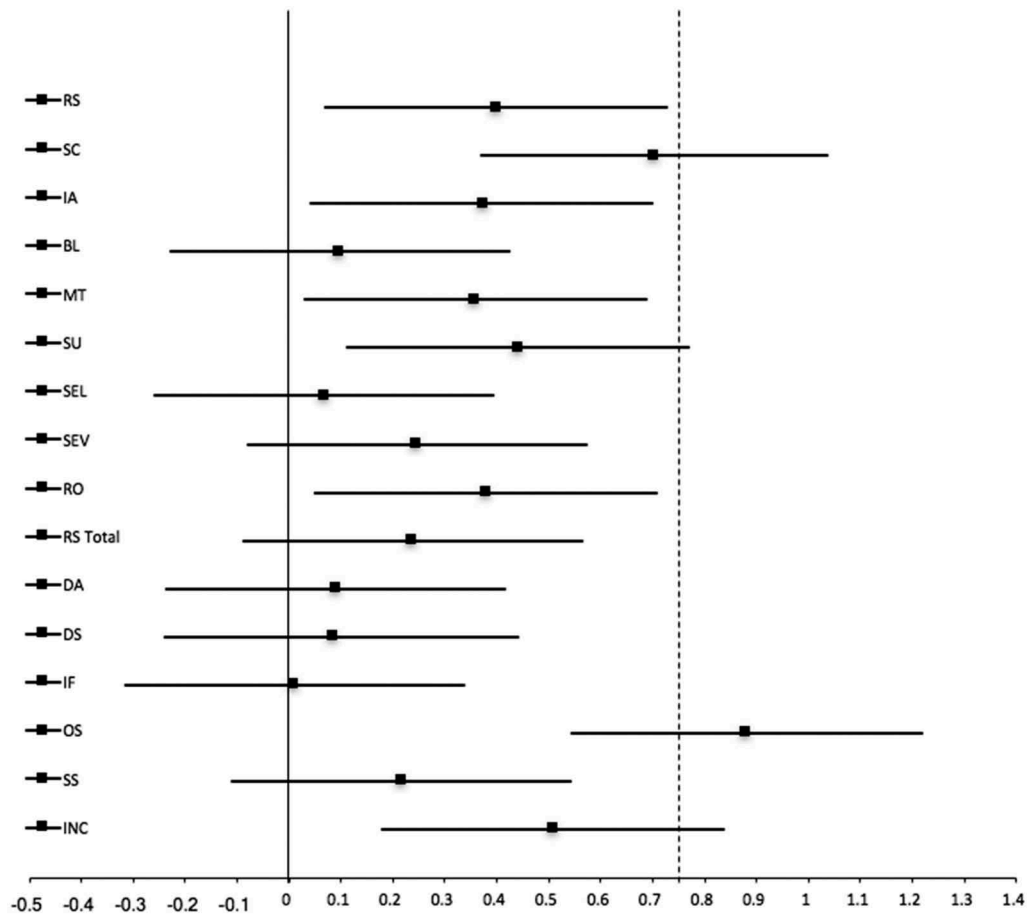


Figure 2. Ninety-five percent confidence intervals for effect sizes (Cohen's d) yielded by the comparison between adults with ADHD and the combined simulation group.

Notes: y-Axis denotes SIRS scales; x-Axis displays effect sizes (Cohen's d); dashed vertical line indicates Roger's standard for medium effect size ($d = 0.75$); RS – Rare Symptoms; SC – Symptom Combinations; IA – Improbable or Absurd Symptoms; BL – Blatant Symptoms; MT (Modified Total Index) – sum score of RS, SC, IA, and BL items; SU – Subtle Symptoms; SEL – Selectivity of Symptoms; SEV – Severity of Symptoms; RO – Reported vs. Observed Symptoms; RS-T (RS-Total) – Rare Symptoms sum score; DA – Direct Appraisal of Honesty; DS – Defensive Symptoms; IF – Improbable Failure; OS – Overly Specified Symptoms; SS (SS Index) – sum score encompassing DA, DS, IF, and OS items; INC – Inconsistency of Symptoms.

included in the SIRS-2 are most suitable for distinguishing genuine from simulated adult manifestations of ADHD.

Application of the decision model described in the SIRS-2 manual (Rogers et al., 2010) resulted in isolated misclassification of genuine cases of ADHD and poor classification rates among simulators. Whereas the majority of genuine patients was classified as such, a considerable minority was categorized as “Indeterminate”. Two adults with ADHD were falsely assumed to be feigning their complaints. Further, mere 19.3% of instructed simulators were detected. The majority of simulators (59.1%) was incorrectly classified as honest responders. Thus, sensitivity was low, with estimates ranging from .19 if “Indeterminate” cases were counted towards the “Genuine” classifications and .25 if Rogers' approach, whereby “Indeterminate” cases are excluded before sensitivity is calculated, was applied. Clearly, either estimate deviates from the sensitivity of .80 listed in the instrument's manual (Rogers et al., 2010).

Specificity, on the other hand, was high, with estimates ranging from .89 to .93.

A different pattern emerged when responses were categorized according to the decision rules outlined in the first edition's manual (Rogers et al., 1992) rather than the SIRS-2's revised decision model. Sensitivity was generally higher than that observed for the SIRS' second edition, ranging from .60 if “Indeterminate” cases were considered not to be feigning and .86 if those with an “Indeterminate” classification were counted towards the “Feigning” cases. This is in line with previous research reporting significant decrements in sensitivity when the SIRS-2 is compared to its predecessor (Brand et al., 2014; Green et al., 2013; Tarescavage & Glassmire, 2016; Tylicki et al., 2018). As approximately 30% of adults with ADHD were wrongfully accused of feigning; however, specificity of the original SIRS was unacceptably low (.24 to .67).

Whereas the SIRS and its successor were in perfect agreement in cases of SIRS-identified honest responders,

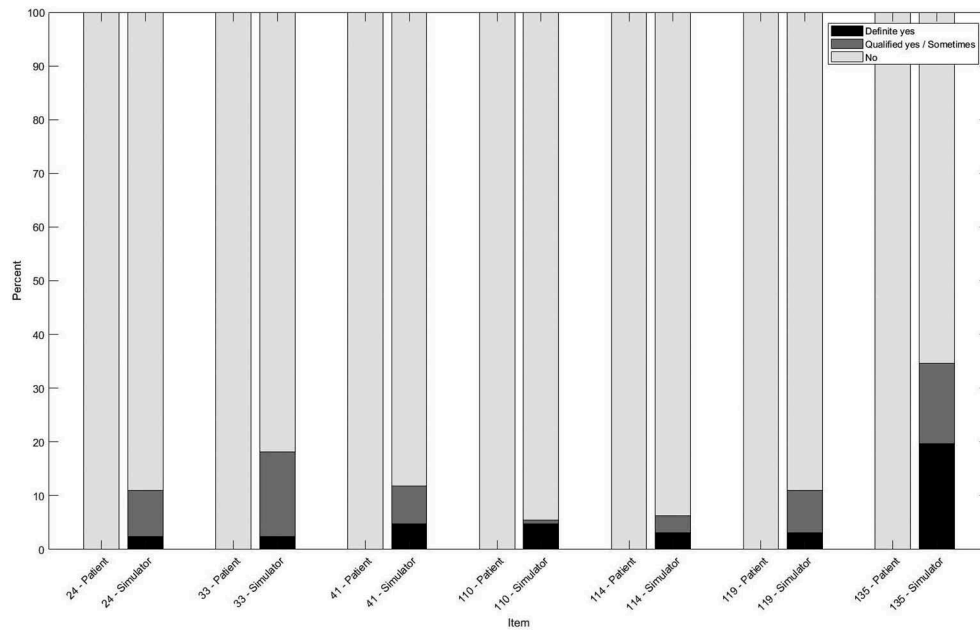


Figure 3. Percentage of participants endorsing items composing the Overly Specified Symptoms (OS) Scale by ADHD group and combined simulation group.

drastically reduced numbers of “Feigning” classifications upon application of the SIRS-2 resulted in an overall convergence of 56.75%. The majority of SIRS-identified feigners (62.07%) was detected based on relevant elevations on the primary scales, while the remaining 37.93% were classified as “Feigning” due to SIRS Total scores above the recommended cut-off value. The additional classifications steps introduced with the advent of the SIRS-2 served their intended purpose of reducing the false-positive rate in that five adults with ADHD avoided a wrongful “Feigning” classification due to an unremarkable RS Total score. Twenty-one simulators, on the other hand, went undetected despite relevant elevations on the primary scales since their RS Total scores did not exceed the recommended cut-off value. In line with previous research (Tarescavage & Glassmire, 2016), the value of the newly introduced MT Index appeared limited. Only one simulator scored above the cut-off score suggested by Rogers and colleagues to be indicative of feigning. The overall distribution of MT Index scores provides further evidence suggesting that the recommended cut-off scores may be inadequate (see Supplementary Material): a score of 13 corresponded to the 60th percentile in the present sample, whereas a score of 22 was equivalent to the 85th percentile, and an MT Index score of 45 equaled the 99th percentile. The SS Index was not further examined as all participants produced unremarkable scores. In sum, these findings underscore the urgency with which the instrument’s authors caution against the SIRS’ application to disorders other than those described in its manual

(Rogers et al., 2010). This may hold particularly true for disorders manifesting primarily in cognitive dysfunction, such as the attention deficits seen in ADHD.

Exploratory comparisons of genuine patients with ADHD and simulators suggested that responses provided by simulators closely resembled those of genuine patients with ADHD on most scales. Considering individual SIRS-2 scales, comparisons between patients with ADHD and a combined group of simulators who received various types and levels of instructions yielded mostly non-significant differences between the groups, with most effects failing to meet an effect size that is considered moderate in the field of malingering research (i.e., $d > 0.75$ according to Roger’s standard, 2008b). A supplementary scale assessing report of Overly Specified Symptoms (OS) emerged as a sole exception, resulting in not only statistically significant differences between the groups but also a moderate effect by the aforementioned standards. The primary scale aiming to uncover simulated symptom reports based on the endorsement of unusual symptom combinations (SC), while also yielding significant differences between genuine and simulated cases, narrowly failed to meet Roger’s requirements for a satisfactory effect. The largest effect found in the present study thus concerned the OS scale and arose, as patients with ADHD endorsed its items to a similar degree as their matched neurotypical peers, whereas individuals feigning the condition tended to endorse items composing this scale to a larger extent. As exemplified by other

SIRS-2 scales (e.g., BL, MT, SU, SEL, SEV and RO scales), patients provided answers which resulted in higher scores than those given by healthy controls. This was not the case on the OS scale, where adults with ADHD did not differ from matched neurotypical controls. Simulators, on the other hand, exaggerated their accounts, thus resulting in a moderate effect when compared to genuine patients.

Upon examination of each individual item contributing to the OS scale, it became apparent that simulators more frequently deemed the enquiries into overly specified symptoms as applicable to ADHD than genuine patients with ADHD report such symptoms. Approximately one-third of simulating participants endorsed Item 135 (“Do you spend much time worrying about your physical health? Would this average between 30 and 40 min each day?”), whereas none of the adults with ADHD did. Adults with a genuine presentation of ADHD did not endorse any of the remaining items on this scale either, while approximately 5–18% of simulators did. Classification accuracy of the OS scale, as determined through an ROC analysis, was nonetheless unsatisfactory. At an adequate specificity of 97.8%, the sensitivity of this scale amounted to 57.9%.

Neither the OS scale, nor the SIRS-2 in its entirety, are suitable as sole means of detecting feigned instances of adult ADHD. They may, however, present invaluable starting points for the development of a disorder-specific instrument. Based on the present findings, items of such an instrument could tap highly specified symptoms which appear relevant to ADHD to those who aim to simulate the disorder. Examining simulators’ endorsement of individual SIRS-2 items suggests that they deem item number 135 more applicable to ADHD than other items based on the same detection strategy. Few simulators, for instance, endorsed items 24 (“Do you have exactly two nightmares every evening?”) or 41 (“Do you hear unnatural sounds or voices? Do they come from closets and nowhere else?”), which apparently bear little relevance to ADHD. Clearly, the item content of such a disorder-specific instrument should be coherent with peoples’ understanding of the condition.

Alongside items based on highly specific symptoms, a disorder-specific instrument for the detection of simulated ADHD could further profit from a second detection strategy employed in the SIRS-2. The primary scale aiming to identify simulators based on their endorsement of uncommon symptom combinations (SC) narrowly failed to meet Roger’s standard of moderate effects between genuine and simulated instances of ADHD in this study. Inquiries such as, “Do you have a need to wash your hands frequently? Is this

related to any unique or special powers that you possess?” (Item 50) or “Do you have strong religious thoughts and periods of giggling? Do these happen together?” (Item 67) may not have struck participants as immediately relevant to ADHD. If adapted to ADHD-specific symptoms, such as instances of inattention, hyperactivity, and impulsivity, this unlikely detection strategy may nonetheless be of great value.

The use of unlikely detection strategies in uncovering feigned ADHD has previously been advocated based on promising results found across numerous studies (see also Harrison, 2006). Suhr and colleagues (Suhr, Buelow, et al., 2011), for example, developed an infrequency index for the Conner’s Adult ADHD Rating Scales (Conners et al., 1999), aiming to detect simulated ADHD by means of a self-report questionnaire commonly used in clinical practice. To this end, Suhr and colleagues determined which CAARS items were rarely ever endorsed by genuine patients with ADHD as well as non-treatment seeking adults. This approach is arguably very akin to Roger’s concept of unlikely detection strategies and has proven promising in later studies (Cook et al., 2016). Even so, this index was derived solely through statistical means, and its sensitivity may be limited (Fuermaier et al., 2016a).

Limitations and future directions

This study employed a simulation design, which has been criticized for its questionable ecological validity (Rogers, 2008a). The motivation of and incentives provided to participants as part of this study are hardly comparable to those faced by individuals who intend to simulate ADHD in real life (Rogers et al., 1993). The current findings should thus be followed up by studies using known-groups comparisons (Rogers, 2008c). However, participants who were asked to simulate adult ADHD as part of this study were verifiably motivated to do so convincingly. Each of them was able to correctly answer questions regarding the simulation instructions and further assured that they had followed these instructions throughout the experiment upon its completion. Lastly, responses provided by simulating participants differed significantly from those of honest responding control participants, suggesting that the instruction to feign ADHD altered their response style ($p < .001$ for all comparisons but the DA scale for which $p = .018$; $0.46 < d < 2.86$). Furthermore, participants’ performance on the GET, a performance validity test, was considered before including them in further analyses. The SIRS-2, on the other hand, is a measure of symptom validity rather than performance validity. As tests of performance validity and symptom validity have been shown to yield diverging

results (Hirsch & Christiansen, 2018), participants in future studies should complete both a performance validity and a symptom validity test prior to the inclusion of their data in further analyses.

Conclusions

The present results strongly advise against reliance on the SIRS or SIRS-2 in the detection of feigned ADHD. Whereas application of the original SIRS' decision rules resulted in acceptable sensitivity and unacceptably low specificity rates, its successor's decision model fared poorly at the detection of feigned ADHD. Specificity of the SIRS-2, on the other hand, was high. These findings extend upon recent studies reporting significantly reduced sensitivity when the SIRS and SIRS-2 are compared (Brand et al., 2014; Green et al., 2013; Tarescavage & Glassmire, 2016; Tylicki et al., 2018). Additionally, they are in line with the developer's cautions (Rogers et al., 2010) which draw attention to the distinction between the detection of feigned mental disorders versus feigned cognitive complaints. The developers underscore the SIRS' utility in uncovering feigned psychiatric symptoms, while the instrument was not designed to detect feigning of the predominantly cognitive complaints manifesting in ADHD. Whereas neither the SIRS-2's classification scheme nor individual scales result in satisfactory classification accuracy, the detection strategies employed in the interview may present useful starting points for the development of a disease-specific symptom validity measure. Items of such a measure may present examinees with symptoms commonly occurring as part of ADHD's clinical manifestation, yet specify these to a degree not usually reported by genuine patients. Symptoms deemed representative of ADHD by laypeople might further be combined to construct disorder-specific items resembling those of the SIRS-2's SC scale.

Acknowledgments

We thank LB for assisting in the realization of the figures presented as part of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Data availability statement

The data that support the findings of this study are available from the corresponding author, ABMF, upon reasonable request.

ORCID

Oliver Tucha  <http://orcid.org/0000-0001-8427-5279>

References

- Adler, L. A., Spencer, T., Faraone, S. V., Kessler, R. C., Howes, M. J., Biederman, J., & Secnik, K. (2006). Validity of pilot adult ADHD Self-Report Scale (ASRS) to rate adult ADHD symptoms. *Annals of Clinical Psychiatry, 18*(3), 145–148.
- Advokat, C. D., Guidry, D., & Martino, L. (2008). Licit and illicit use of medications for attention-deficit hyperactivity disorder in undergraduate college students. *Journal of American College Health, 56*(6), 601–606.
- Aita, S. L., Sofko, C. A., Hill, B. D., Musso, M. W., & Boettcher, A. C. (2017). Utility of the personality assessment inventory in detecting feigned Attention-Deficit/Hyperactivity Disorder (ADHD): The feigned adult ADHD index. *Archives of Clinical Neuropsychology*, (November), 1–13. doi:10.1093/arclin/acx113
- American Psychiatric Association. (2013). DSM 5. *American Journal of Psychiatry* doi:10.1176/appi.books.9780890425596.744053
- Association, A. P. (2000). *Diagnostic and statistical manual of mental disorders*. Washington, DC: IV TR. doi:10.1176/appi.books.9780890423349
- Barkley, R. A., & Murphy, K. R. (1998). *Attention-deficit hyperactivity disorder: a clinical workbook*. New York, NY: Guilford Publications Inc.
- 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*(12), 1792–1798.
- Brand, B. L., Tursich, M., Tzall, D., & Loewenstein, R. J. (2014). Utility of the SIRS-2 in distinguishing genuine from simulated dissociative identity disorder. *Psychological Trauma: Theory, Research, Practice, and Policy, 6*(4), 308–317.
- Bryant, A. M., Lee, E., Howell, A., Morgan, B., Cook, C. M., Patel, K., ... Suhr, J. A. (2018). The vulnerability of self-reported disability measures to malingering: A simulated ADHD study. *Clinical Neuropsychologist, 32*(1), 109–118.
- Bundesministerium der Justiz und für Verbraucherschutz. (2016). §33 Einkommensteuergesetz (EStG). Retrieved from <https://www.gesetze-im-internet.de/estg/BJNR010050934.html>
- 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*(4), 419–426.
- Butcher, J. N. (2009). Minnesota multiphasic personality inventory. *Corsini Encyclopedia of Psychology, 21*(1–3), 2010.
- Conners, C. K., Erhardt, D., & Sparrow, M. A. (1999). *Conners' Adult ADHD Rating Scales (CAARS)*. New York, NY: Multihealth Systems.
- Conners, C. K., Staff, M. H. S., Connelly, V., Campbell, S., MacLean, M., & Barnes, J. (2000). Conners' continuous

- performance test II (CPT II V. 5). *Multi-Health Systems Inc*, 29(1), 175–196.
- Cook, C., Buelow, M. T., Lee, E., Howell, A., Morgan, B., Patel, K., ... Suhr, J. (2017). Malingered Attention Deficit/Hyperactivity Disorder on the Conners' Adult ADHD rating scales. *Journal of Psychoeducational Assessment*, 073428291769693. doi:10.1177/0734282917696934
- Cook, C. M., Bolinger, E., & Suhr, J. (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(4), 358–364.
- Curatolo, P. (2005). The neurology of attention deficit/hyperactivity disorder. *Brain and Development*, 27(8), 541–543.
- Dunn, T. M., Shear, P. K., Howe, S., & Ris, M. D. (2003). Detecting neuropsychological malingering: Effects of coaching and information. *Archives of Clinical Neuropsychology : the Official Journal of the National Academy of Neuropsychologists*, 18(2), 121–134.
- Edmundson, M., Berry, D. T. R., Combs, H. L., Brothers, S. L., Harp, J. P., Williams, A., ... Scott, A. B. (2017). The effects of symptom information coaching on the feigning of adult ADHD. *Psychological Assessment*, 29(12), 1429–1436.
- Eme, R. (2014). Attention deficit hyperactivity disorder: A mitigation of criminal responsibility? *Journal of Forensic Psychology Practice*, 14(3), 221–236.
- Fisher, A. B., & Watkins, M. W. (2008). ADHD rating scales' susceptibility to faking in a college student sample. *Journal of Postsecondary Education and Disability*, 20(2), 81–92. Retrieved from: <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ825775&site=ehost-live%5Cnhttp://www.ahead.org/publications/jped>
- Freeman, T., Powell, M., & Kimbrell, T. (2008). Measuring symptom exaggeration in veterans with chronic post-traumatic stress disorder. *Psychiatry Research*, 158(3), 374–380.
- Fuermaier, A. B. M., Tucha, L., Koerts, J., Mueller, A. K., Lange, K. W., & Tucha, O. (2012). Measurement of stigmatization towards adults with attention deficit hyperactivity disorder. *PLoS ONE*, 7, 12.
- Fuermaier, A. B. M., Tucha, L., Koerts, J., Weisbrod, M., Grabemann, M., Zimmermann, M., ... Tucha, O. (2016a). Evaluation of the CAARS infrequency index for the detection of noncredible ADHD symptom report in adulthood. *Journal of Psychoeducational Assessment*, 34(8), 739–750.
- Fuermaier, A. B. M., Tucha, L., Koerts, J., Weisbrod, M., Lange, K. W., Aschenbrenner, S., & Tucha, O. (2016b). Effects of methylphenidate on memory functions of adults with ADHD. *Applied Neuropsychology: Adult*, 9095(June), 1–13.
- Fuermaier, A. B. M., Tucha, O., Koerts, J., Grabski, M., Lange, K. W., Weisbrod, M., ... Tucha, L. (2016c). The development of an embedded figures test for the detection of feigned attention deficit hyperactivity disorder in adulthood. *Plos One*, 11(10), e0164297.
- Fuermaier, A. B. M., Tucha, O., Koerts, J., Send, T., Weisbrod, M., Aschenbrenner, S., & Tucha, L. (2018). Is motor activity during cognitive assessment an indicator for feigned ADHD in adults? *Journal of Clinical and Experimental Neuropsychology*.
- Fuermaier, A. B. M., Tucha, L., Koerts, J., Aschenbrenner, S., & Tucha, O. (2017a). *Vienna Test System (VTS): The Groningen Effort Test (GET)*. Vienna, Austria: Schuhfried.
- Fuermaier, A. B. M., Tucha, O., Stenger, H., Koerts, J., Aschenbrenner, S., & Tucha, L. (2017b). *Test manual: The Groningen Effort Test (GET)*. Vienna, Austria: Schuhfried.
- Green, D., Rosenfeld, B., & Belfi, B. (2013). New and improved? A comparison of the original and revised versions of the structured interview of reported symptoms. *Assessment*, 20(2), 210–218.
- Green, P. (2003). Manual for the word memory test for windows: User's manual and program. In *Edmonton*. Alberta, Canada: Green's Publishing.
- Green, P. (2008). Green's Nonverbal Medical Symptom Validity Test (NV-MSVT) for microsoft windows: User's manual. In *Edmonton*. Alberta, Canada: Green's Publishing.
- Hall, W. D., & Lucke, J. C. (2010). The enhancement use of neuropharmaceuticals: More scepticism and caution needed. *Addiction*, 105(12), 2041–2043.
- Harrison, A. G. (2006). Adults faking ADHD: You must be kidding! *The ADHD Report*, 14(4), 1–7.
- Harrison, A. G., & Armstrong, I. T. (2016). Development of a symptom validity index to assist in identifying ADHD symptom exaggeration or feigning. *Clinical Neuropsychologist*, 30(2), 265–283.
- Harrison, A. G., Edwards, M. J., & Parker, K. C. H. (2008). Identifying students feigning dyslexia: Preliminary findings and strategies for detection. *Dyslexia*, 14, 228–246.
- 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.
- Hirsch, O., & Christiansen, H. (2018). Faking ADHD? Symptom validity testing and its relation to self-reported, observer-reported symptoms, and neuropsychological measures of attention in adults with ADHD. *Journal of Attention Disorders*, 22(3), 269–280.
- Ilieva, I. P., & Farah, M. J. (2013). Enhancement stimulants: Perceived motivational and cognitive advantages. *Frontiers in Neuroscience*, 7(October), 198.
- Internal Revenue Service. (2016). Publication 502. Retrieved from <https://www.irs.gov/publications/p502/ar02.html>
- Jachimowicz, G., & Geiselman, R. E. (2004). Comparison of ease of falsification of attention deficit hyperactivity disorder diagnosis using standard behavioral rating scales. *Cognitive Science*, 2, 6–20. Retrieved from <http://cogsci-online.ucsd.edu/2/2-1.pdf>
- Jasinski, L. J., & Ranseen, J. D. (2011). Malingered ADHD evaluations: A further complication for accommodations reviews. *NCBE Annual Bar Admissions Conference*, San Francisco, California.
- Kessler, R. C., Adler, L., Ames, M., Demler, O., Faraone, S., Hiripi, E., ... Walters, E. E. (2005). The World Health Organization adult ADHD self-report scale (ASRS): A short screening scale for use in the general population. *Psychological Medicine*, 35(2), 245–256.
- Kessler, R. C., Adler, L. A., Barkley, R., Biederman, J., Conners, C. K., Faraone, S. V., ... Zaslavsky, A. M. (2005). Patterns and predictors of attention-deficit/hyperactivity disorder persistence into adulthood: Results from

- the national comorbidity survey replication. *Biological Psychiatry*, 57(11), 1442–1451.
- 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(5), 335–345.
- Leppma, M., Long, D., Smith, M., & Lassiter, C. (2017). Detecting symptom exaggeration in college students seeking ADHD treatment: Performance validity assessment using the NV-MSVT and IVA-plus. *Applied Neuropsychology: Adult*, 1–9. doi:10.1080/23279095.2016.1277723
- Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2004). *Neuropsychological assessment* (4th ed.). New York, NY: Oxford University Press.
- Liu, C., Liu, Z., Chiu, H. F. K., Carl, T. W. C., Zhang, H., Wang, P., ... Xue, Z. (2013). Detection of malingering: Psychometric evaluation of the Chinese version of the structured interview of reported symptoms-2. *BMC Psychiatry*, 13. doi:10.1186/1471-244X-13-254
- 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(7), 1204–1237.
- 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(10), 1290–1302.
- Morey, L. C. (1991). *Personality assessment inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Morey, L. C. (2016). *Tests of attentional distraction*. College Station, TX: Morey Consulting.
- Morey, L. C. (2017). Examining a novel performance validity task for the detection of feigned attentional problems. *Applied Neuropsychology: Adult*, 1–13. doi:10.1080/23279095.2017.1409749
- Musso, M. W., & Gouvier, W. D. (2014). "Why is this so hard?": A review of detection of malingered ADHD in college students. *Journal of Attention Disorders*, 18(3), 186–201.
- Musso, M. W., Hill, B. D., Barker, A. A., Pella, R. D., & Gouvier, W. D. (2016). Utility of the personality assessment inventory for detecting malingered ADHD in college students. *Journal of Attention Disorders*, 20(9), 763–774.
- Pulcini, C. D., Perrin, J. M., Houtrow, A. J., Sargent, J., Shui, A., & Kuhlthau, K. (2015). Examining trends and coexisting conditions among children qualifying for SSI under ADHD, ASD, and ID. *Academic Pediatrics*, 15(4), 439–443.
- Quinn, C. A. (2003). Detection of malingering in assessment of adult ADHD. *Archives of Clinical Neuropsychology*, 18(4), 379–395.
- Rabiner, D. L. (2013). Stimulant prescription cautions: Addressing misuse, diversion and malingering topical collection on attention-deficit disorder. *Current Psychiatry Reports*, 15(7). doi:10.1007/s11920-013-0375-2
- Rabiner, D. L., Anastopoulos, A. D., Costello, E. J., Hoyle, R. H., McCabe, S. E., & Swartzwelder, H. S. (2009). The misuse and diversion of prescribed ADHD medications by college students. *Journal of Attention Disorders*, 13(2), 144–153.
- Rogers, (2009). Clinical assessment of malingering and deception, 3rd edition. *British Journal of Psychiatry*, 195 doi:10.1192/bjp.bp.108.060855
- Rogers, R. (2008a). An introduction to response styles. In *Clinical assessment of malingering and deception* (3rd ed., pp. 3–13). New York, NY: Guilford Publications Inc.
- Rogers, R. (2008b). Detection strategies for malingering and defensiveness. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 14–35). New York, NY: Guilford Press.
- Rogers, R. (2008c). Researching response styles. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 411–434). New York, NY: Guilford Press.
- Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *Structured interview of reported symptoms: Professional manual* (1st ed.). Lutz, FL: Psychological Assessment Resources.
- Rogers, R., Harrell, E. H., & Liff, C. D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review*, 13, 255–274.
- Rogers, R., Kropp, R. P., Bagby, M. R., & Dickens, S. E. (1992). Faking specific disorders: A study of the Structured Interview of Reported Symptoms (SIRS). *Journal of Clinical Psychology*, 48(5), 643–648.
- Rogers, R., Sewell, K. W., & Gillard, N. D. (2010). *Structured interview of reported symptoms 2nd Edition: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Salekin, R. T., Kubak, F. A., & Lee, Z. (2008). Deception in children and adolescents. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 343). New York, NY: The Guilford Press.
- Sansone, R. A., & Sansone, L. A. (2011). Faking attention deficit hyperactivity disorder. *Innovations in Clinical Neuroscience*, 8(8), 10–13. Retrieved from: <http://www.ncbi.nlm.nih.gov/pubmed/21922064>
- Sharma, A., & Couture, J. (2014). A review of the pathophysiology, etiology, and treatment of Attention-Deficit Hyperactivity Disorder (ADHD). *Annals of Pharmacotherapy*, 48, 209–225.
- Smith, M. E., & Farah, M. J. (2011). Are prescription stimulants "smart pills"? The epidemiology and cognitive neuroscience of prescription stimulant use by normal healthy individuals. *Psychological Bulletin*, 137(5), 717–741.
- Smith, S. T., Cox, J., Mowle, E. N., & Edens, J. F. (2017). Intentional inattention: Detecting feigned attention-deficit/hyperactivity disorder on the personality assessment inventory. *Psychological Assessment*, 29(12), 1447–1457.
- Steyn, B. J., Schneider, J., & McArdle, P. (2002). The role of disability living allowance in the management of attention-deficit hyperactivity disorder. *Child: Care, Health And Development*, 28(2), 523–527.
- Suhr, J. A., Buelow, M., & Riddle, T. (2011). Development of an infrequency index for the CAARS. *Journal of Psychoeducational Assessment*, 29(2), 160–170.
- Suhr, J. A., Sullivan, B. K., & Rodriguez, J. L. (2011). The relationship of noncredible performance to continuous performance test scores in adults referred for attention-deficit/hyperactivity disorder evaluation. *Archives of Clinical Neuropsychology*, 26(1), 1–7.

- Tarescavage, A. M., & Glassmire, D. M. (2016). Differences between Structured Interview of Reported Symptoms (SIRS) and SIRS-2 sensitivity estimates among forensic inpatients: A criterion groups comparison. *Law and Human Behavior, 40*(5), 488–502.
- Thome, J., Ehlis, 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. *World Journal of Biological Psychiatry, 13*(5), 379–400.
- Tucha, L., & Fuermaier, A. B. M. (2015). Detection of feigned attention deficit hyperactivity disorder. *Journal of Neural Transmission, 122*(Suppl 1), S123–S134.
- 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.
- 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*(10), 1575–1592.
- Tylicki, J. L., Wygant, D. B., Tarescavage, A. M., Frederick, R. I., Tyner, E. A., Granacher, R. P., & Sellbom, M. (2018). Comparability of Structured Interview of Reported Symptoms (SIRS) and Structured Interview of Reported Symptoms-Second Edition (SIRS-2) classifications with external response bias criteria. *Psychological Assessment, 30*(9), 1144–1159.
- Walls, B. D., Wallace, E. R., Brothers, S. L., & Berry, D. T. R. (2017). Utility of the conners' adult ADHD rating scale validity scales in identifying simulated attention-deficit hyperactivity disorder and random responding. *Psychological Assessment, 29*(12), 1437–1446.
- Wankerl, B., Hauser, J., Makulska-Gertruda, E., Reißmann, A., Sontag, T. A., Tucha, O., & Lange, K. W. (2014). Neurobiology of attention deficit hyperactivity disorder. *Fortschritte Der Neurologie-Psychiatrie, 82*(1), 9–29.
- Ward, M. F., Wender, P. H., & Reimherr, F. W. (1993). The wender utah rating-sclae - An aid in the retrospective diagnosis of childhood attention-deficit hyperactivity disorder. *American Journal of Psychiatry, 150*(6), 885–890.
- Wender, P. H., Wolf, L. E., & Wasserstein, J. (2001). Adults with ADHD. An overview. *Annals of the New York Academy of Sciences, 931*, 1–16.
- Young, J. C., & Gross, A. M. (2011). Detection of response bias and noncredible performance in adult attention-deficit/hyperactivity disorder. *Archives of Clinical Neuropsychology, 26*(3), 165–175.