



How Plausible Is the Implausible? Students' Plausibility and Prevalence Ratings of the Self-Report Symptom Inventory

Irena Boskovic^{1,2} · Thomas Merten³ · Harald Merckelbach²

Received: 20 January 2021 / Accepted: 16 April 2021 / Published online: 23 April 2021
© The Author(s) 2021

Abstract

Some self-report symptom validity tests, such as the Self-Report Symptom Inventory (SRSI), rely on a detection strategy that uses bizarre, extreme, or very rare symptoms. Thus, items are constructed to *invite* respondents with an invalid response style to affirm pseudosymptoms that are usually not experienced by genuine patients. However, these pseudosymptoms should not be easily recognizable, because otherwise sophisticated over-reporters could strategically avoid them and go undetected. Therefore, we tested how well future psychology professionals were able to differentiate between genuine complaints and pseudosymptoms in terms of their plausibility and prevalence.

Psychology students ($N=87$) received the items of the SRSI online and were given the task to rate each item as to its plausibility and prevalence in the community.

Students evaluated genuine symptoms as significantly more plausible and more prevalent than pseudosymptoms. However, 56% of students rated pseudosymptoms as moderately plausible, whereas 17% rated them as moderately prevalent in the general public.

Overall, it appears that psychology students are successful in distinguishing bizarre, unusual, or rare symptoms from genuine complaints. Yet, the majority of students still attributed relatively high *prima facie* plausibility to pseudosymptoms. We contend that if such a trusting attitude is true for psychology students, it may also be the case for young psychology practitioners, which, consequently, may diminish the probability of employing self-report validity measures in psychological assessments.

Keywords Self-Report Symptom Inventory · Feigning · Pseudosymptoms · Overreporting · Malingering · Plausibility

Introduction

Some people malingering symptoms, that is, they engage in intentional amplification or complete fabrication of health complaints so as to secure certain benefits such as disability pensions or financial compensation. In medicolegal contexts, malingering occurs on a non-trivial level, ranging from 20 to 40% according to some authors (Bass & Halligan, 2014; Greve et al., 2009; Mittenberg et al., 2002), while even higher prevalence rates were reported by other authors for selected referral backgrounds

or certain types of claimed symptomatology (e.g., whiplash injury, Schmand et al., 1998; social security claimants, Chafetz et al., 2007; see also Dandachi-FitzGerald et al., 2020).

The recommended approach to screen for intentional over-reporting of symptoms is to administer symptom validity tests (SVTs, Bianchini et al., 2005; Bush et al., 2014; Chafetz et al., 2015; Dandachi-FitzGerald et al., 2013). One prominent type of SVT is the self-report validity tests. Self-report-based SVTs have attained an ever-increasing importance in the detection of non-valid symptom claims. The majority of them rely on either of two approaches: The first approach is to identify a level of symptom reporting that is so extreme that its authenticity and believability are questionable. The best known example of an SVT that operates on the basis of this principle is the Symptom Validity Scale, formerly known as Fake Bad Scale (FBS), originally developed by Lees-Haley et al. (1991), of the Minnesota Multiple Personality Inventory-2 (MMPI-2; Butcher et al., 2001). The second approach is to gauge the tendency of malingerers to endorse unlikely

✉ Irena Boskovic
boskovic@essb.eur.nl;
irena.boskovic@maastrichtuniversity.nl

¹ Erasmus University Rotterdam, Rotterdam, South Holland, Netherlands

² Maastricht University, Maastricht, Limburg, Netherlands

³ Department of Neurology, Vivantes Klinikum Im Friedrichshain, Berlin, Friedrichshain, Germany

symptoms, which is bizarre, rare or extreme symptoms. Such symptoms may be occasionally experienced or reported by a genuine patient, but the probability that genuine patients endorse a fair number of them is low. The best known example of a freestanding SVT predominately using the second approach is the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997). The items of four of its five subscales describe atypical, extreme, or bizarre complaints, related to amnesic disorders, psychosis, low intelligence, and neurological impairment (see Martin et al., 2015; Van Impelen et al., 2014).

With the aim to overcome the drawbacks of SIMS, such as high face validity and limited relevance for non-criminal forensic contexts, Merten et al. (2016) developed the Self-Report Symptom Inventory (SRSI), with a comprehensive professional manual published in German language (Merten et al., 2019). It comprises 107 true-false items, with 50 items describing potentially genuine symptoms and 50 items referring to pseudosymptoms. The remaining seven items check for participants' compliance with instructions (2 items) and their consistency in reporting health complaints (5 items). The genuine and the pseudosymptom scales each consist of five subscales pertaining to frequently reported complaints in the medico-legal context, such as cognitive complaints, pain, nonspecific somatic symptoms, and anxiety/PTSD. The genuine symptoms and the pseudosymptoms are randomly mixed within the SRSI. With multiple language versions of the instrument available (e.g., German, Dutch, French, English, Serbian, Norwegian, Russian), research concerning the SRSI has addressed a variety of psychometric issues (e.g., Boskovic et al., 2020; Geurten et al., 2018; Stevens et al., 2018).

A good SVT should effectively distinguish between honest respondents and invalid responders. In terms of test theory, this would correspond to optimal sensitivity and specificity. In order to fulfill this requirement, one important condition is that the test should “invite” overreporting respondents to affirm symptoms that are usually rejected by honest responders who fully understand the items and truthfully respond to them. Ideally, this means that pseudosymptom items should describe health complaints that appear to occur with some frequency in true patient populations (bona-fide patients) when, in fact, they rarely (or, in the extreme, never) occur in such populations. As said before, when pseudosymptom items are too obvious, malingerers will avoid them resulting in low detection rates, i.e., sensitivity. With this in mind, a crucial, and so far unexplored, question with regard to the SRSI is whether its pseudosymptoms possess enough *prima facie* plausibility.

Previous cross-cultural research showed that psychologists in practice agree that items of SIMS are odd and rare, yet, their ratings also revealed moderate, rather than low, plausibility of such claims (Boskovic et al., 2017). Such

findings indicate a tendency to take patients' claims at their face value or so-called truth bias (Beach et al., 2017), hence, a lack of skepticism (Lilienfeld et al., 2016) among practitioners. If this was the case for the items quite bizarre in their quality, one can only assume that the issue might be more severe for the items less obvious in their implausibility. Hence, the current study examined whether psychology students are able to distinguish between the genuine and pseudosymptom items of the SRSI, as well as the extent to which they find this task difficult.

Method

Participants

In total, 87 bachelor level psychology students participated in our study. The majority of the sample were women (89.7%), and the average age was 20.2 years ($SD = 2.07$). We asked participants to rate their English proficiency on a 5-point Likert scale (1 = low and 5 = extremely good), and their ratings indicated overall high English proficiency ($M = 4.28$, $SD = 0.62$).

Before administering the list of symptoms, we asked students whether they experienced any symptoms at the time of their participation. The majority of our sample (72.4%) reported not having any health complaints, whereas 8.0% confirmed having some complaints, and 6.9% reported having a chronic health condition. We also offered an option “No, but people close to me do have health issues,” which 12.6% of participants selected.

Measures

Participants were presented with the genuine and pseudosymptom items of the Self-Report Symptom Inventory (SRSI; Merten et al., 2016). For the current study, the two warming-up items and five consistency items of the instrument were discarded and only plausibility and prevalence ratings of the 100 symptom and pseudosymptom items were obtained. Genuine symptom items encompass five subscales that address health concerns patients often report such as (1) cognitive complaints, (2) depression, (3) pain, (4) nonspecific somatic complaints, and (5) PTSD/anxiety. The pseudosymptom scale includes five subscales tapping into unlikely complaints in the following domains: (1) cognitive/memory, (2) neurological (motor) complaints, (3) neurological (sensory) complaints, (4) pain, and (5) anxiety/depression. The pseudosymptoms of the SRSI were generated in a two-step procedure. In the first step, a group of experts had listed potential pseudosymptom items, and in the second step, pseudosymptoms underwent an empirical item selection procedure (Merten et al., 2016). Several studies showed

that genuine patient groups seldom endorsed the pseudo-symptoms. For example, van Helvoort et al. (2019) found in their study including 39 forensic patients that they endorsed on average only 1.63 items ($SD = 2.31$) of the 50 pseudo-symptoms, but, when instructed to malingering symptoms, they endorsed on average 24.54 pseudo-symptoms ($SD = 13.39$).

Procedure

This study was conducted online, using Qualtrics. The study was posted on the student research participation portal (SONA), and the only eligibility criterion was sufficient proficiency in English as English-language SRSI items were presented for judgment. After following the link, students were briefly explained the purpose of the study and asked to provide informed consent. After the demographic questions, we administered the SRSI items. In the original SRSI, items describe potential health problems and can be answered by “true” if a symptom description is affirmed as present and “false” if not. However, in the current study, after reading each item, students were asked to grade the plausibility and prevalence of the symptoms on two 5-point scales (1 = low and 5 = high). After finishing the questionnaire, participants were asked to rate their motivation and difficulty of the task to rate symptoms on plausibility and prevalence using 5-point scales (1 = low and 5 = high). Finally, all participants received a debriefing form and were compensated with 0.5 research credit points. This study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience, Maastricht University, the Netherlands.

Data Analyses

Below, we present mean plausibility and estimated prevalence ratings along with the 95% confidence intervals (CIs). We employed paired t tests for comparing genuine symptom and pseudo-symptom scales and subscales. Cohen's d s are provided for effect sizes and were corrected for the correlation between the variables. Further, we looked at the proportion of participants who found, on average, the pseudo-symptoms highly plausible (≥ 3) and prevalent (≥ 3), as well as at the proportion of students who rated them as implausible (≤ 2) and relatively rare (≤ 2).

Results

Integrity Check

Overall, participants were moderately motivated to participate in our study ($M = 3.24$, $SD = 0.68$; range 2–5). They found it moderately difficult to grade plausibility and prevalence, $M = 3.20$ ($SD = 0.92$) and $M = 3.31$ ($SD = 0.92$), respectively. The full dataset and outputs can be found at Open Science platform, <https://osf.io/vth9k/>.¹

Plausibility

At the level of two main scales, plausibility and prevalence ratings of genuine symptoms ($r = 0.456$, $p < 0.001$), and of pseudo-symptoms ($r = 0.289$, $p = 0.007$) correlated significantly and positively (for subscales see Supplemental Table 1). Mean plausibility and prevalence ratings for each domain can be found in Table 1. At the level of main scales, students rated pseudo-symptoms as less plausible than genuine symptoms: paired $t(86) = 19.96$, $p < .001$, 95% CI for mean difference [.85, 1.04], Cohen's $d = 2.13$. We also looked into ratings subscales that are content-wise comparable—cognitive symptoms, pain, anxiety/depression/PTSD—and evaluated students' scores with paired t tests: cognitive symptoms, $t(86) = 13.20$, $p < .001$; 95% CI [.59, .73], Cohen's $d = 1.13$; pain, $t(86) = 15.30$, $p < .001$; 95% CI [.75, .98], Cohen's $d = 1.63$, and anxiety/depression/PTSD, $t(86) = 10.81$, $p < .001$; 95% CI [.46, .67], Cohen's $d = 1.14$. Altogether, the 95% CI's for mean differences in plausibility scores indicate that the distance between genuine and pseudo-symptoms rarely exceeded one scale point (1.0). More importantly, the 95% CIs for means (see Table 1) show that pseudo-symptoms were generally rated at or above the scale midpoint.

Prevalence

Genuine symptoms were rated as more prevalent than were pseudo-symptoms: $t(86) = 27.97$, $p < .001$, 95% CI [.89, 1.02], Cohen's $d = 3.03$. The 95% CIs indicated that the prevalence of pseudo-symptoms is generally rated lower than their plausibility (see Table 1). The paired t tests for cognitive symptoms, pain, and anxiety/depression/PTSD are as follows: cognitive symptoms, $t(86) = 18.37$, $p < .001$; 95% CI [.58, .73], Cohen's $d = 1.97$; Pain, $t(86) = 18.36$, $p < .001$; 95% CI [.85, 1.05], Cohen's $d = 1.98$, and anxiety/depression/PTSD, $t(86) = 13.79$, $p < .001$; 95% CI [.85, .1.05], Cohen's $d = 1.48$.

Frequency of Plausibility and Prevalence Ratings

We inspected how many students obtained total plausibility scores on the pseudo-symptoms scale that were above the midpoint (≥ 3), and 56.3% ($n = 48$) met that condition. In contrast, 17.2% ($n = 15$) gave pseudo-symptoms an overall plausibility rating of ≤ 2 . The proportion of students who rated the prevalence of pseudo-symptoms scale above the midpoint was lower than for the plausibility ratings, but still non-trivial: 17.2%

¹ All the analyses were also performed without students whose motivation ratings were low (< 2 ; $n = 9$), but the results were basically similar. Output is provided at OSF platform.

Table 1 Means, standard deviations, and 95% confidence intervals for all Self-Report Symptom Inventory scales and subscales on plausibility and prevalence ratings

		Plausibility ratings			Prevalence ratings		
		<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI
Genuine	SRSI scale						
	Symptoms total	3.98	0.80	3.80, 4.15	3.57	0.45	3.48, 3.67
	Cognitive symptoms	3.79	0.81	3.61, 3.96	3.36	0.46	3.26, 3.46
	Depression symptoms	4.01	0.86	3.82, 4.19	3.63	0.47	3.53, 3.73
	Pain symptoms	3.89	1.01	3.68, 4.12	3.53	0.62	3.40, 3.66
	Non-specific somatic symptoms	4.25	0.71	4.10, 4.40	3.89	0.47	3.78, 3.99
Pseudo	PTSD/anxiety symptoms	3.94	0.89	3.74, 4.12	3.47	0.52	3.35, 3.58
	Pseudosymptoms total	3.03	0.89	2.84, 3.22	2.61	0.48	2.51, 2.72
	Cognitive pseudosymptoms	3.10	0.93	2.90, 3.29	2.70	0.50	2.60, 2.81
	Motor pseudosymptoms	2.67	0.89	2.47, 2.85	2.36	0.51	2.24, 2.47
	Sensory pseudosymptoms	3.00	0.95	2.80, 3.20	2.54	0.54	2.42, 2.65
	Pain pseudosymptoms	3.03	0.97	2.82, 3.23	2.57	0.56	2.45, 2.69
	Anxiety/depression/PTSD pseudosymptoms	3.37	0.91	3.17, 3.56	2.91	0.53	2.80, 3.02

SRSI Self-Report Symptom Inventory

(*n* = 15). Most importantly, 6.9% (*n* = 6) gave an overall prevalence rating of ≤ 2 to pseudosymptoms.² For the frequency of ratings on the level of subscales, see Table 2.

Discussion

Were the pseudosymptoms of the SRSI judged as reasonably plausible and prevalent by undergraduate psychology students? Or were they immediately recognizable as bogus

symptoms, thereby suppressing the potential sensitivity of the SRSI? Our results show that students rated genuine symptoms of the SRSI as significantly more plausible than its pseudosymptoms. For this kind of self-report validity assessment to work, pseudosymptoms should maintain their character to “invite” persons with invalid symptom claims to endorse them. Thus, they should appear somewhat plausible. If this were not the case, questionnaires like the SIMS or the SRSI would not work in a clinical or forensic practice. Hence, a subtle trade-off between face validity and item dif-

Table 2 Frequency of rating ≤ 2, between 2 and 3, and ≥ 3 of plausibility and prevalence of SRSI subscales and main scales

	SRSI subscales	Plausibility ratings (%)			Prevalence ratings (%)		
		≤ 2	2–3	≥ 3	≤ 2	2–3	≥ 3
Genuine	Cognitive	3.4	12.6	83.9	–	17.2	82.8
	Depression	2.3	8.0	89.7	–	6.9	93.1
	Pain	8.0	9.2	82.8	2.3	12.6	85.1
	Nonspecific somatic	2.3	4.6	93.1	–	4.6	95.4
	Anxiety/PTSD	4.6	8.0	87.4	–	14.9	85.1
	Total	2.3	10.3	87.4	–	8.0	92.0
Pseudo	Cognitive	14.9	23.0	62.1	5.7	64.4	29.9
	Motor	27.6	32.3	40.2	28.7	58.6	12.6
	Sensory	18.4	27.6	54.0	19.5	58.6	21.8
	Pain	17.2	26.4	56.3	19.5	57.5	23.0
	Anxiety/PTSD/depression	10.3	17.2	72.4	3.4	49.4	47.1
	Total	17.2	26.4	56.3	6.9	75.9	17.2

² We also checked whether participants without symptoms (*n* = 63) and participants with symptoms or with close people with symptoms (*n* = 24) differed in their plausibility and prevalence ratings on genuine symptoms and pseudosymptom scales. There results indicated no significant differences (*Welch's ts* < 1.39; *ps* > .167; see Supplemental Table 2).

ficulty/item validity must be obtained. However, from multiple data including different referral contexts and different cultural backgrounds (as summarized in the test manual, Merten et al., 2019), it is well known that pseudosymptom endorsement is much lower than genuine symptom endorsement, even in patients or research participants with invalid

and/or exaggerated symptom reports. It lies in the (partly bizarre, partly extreme) nature of pseudosymptoms that, on average, they present with an a priori lower plausibility and a low expectation to occur in patients than do common genuine symptoms.

Looking at the mean ratings of genuine symptoms and pseudosymptoms in the current study, both types of items were attributed a moderate level of plausibility, which was higher than expected for the pseudosymptoms, but, at the same time, lower than expected for the genuine symptom items. This suggests that, at least for students, the SRSI does possess this subtle trade-off.

Our findings are in line with those of a previous study that found practitioners across different cultures to rate bizarre items taken from the SIMS as moderately plausible (Boskovic et al., 2017). The general tendency to give moderate plausibility rating even to the bizarre items might simply signal the idiographic approach practitioners and psychologists in training often take, which lowers their skepticism and encourages a “believe bias” or “truth bias” (see Beach & Taylor, 2017; Lilienfeld et al., 2016). Thus, our findings justify concerns that genuine symptoms and bogus symptom are not differentiated sufficiently in the mental maps of future psychologists; consequently, pseudosymptoms may be mistaken for genuine complaints. Outside the context of SVTs, this would create room for incorrect diagnostic decisions with potentially harmful consequences.

Inspecting all genuine symptoms subscales, it is noticeable that the lowest plausibility scores ($M=3.79$, $SD=0.81$, 95% CI [3.61, 3.97]) were given to cognitive complaints, whereas the highest ($M=4.25$, $SD=0.71$, 95% CI [4.10, 4.40]) were attributed to non-specific somatic symptoms. Among the pseudosymptoms subscales, the highest mean ratings were for anxiety/depression/PTSD subscale ($M=3.37$, $SD=0.91$, 95% CI [3.17, 3.56]) and the lowest for motor pseudosymptoms ($M=2.67$, $SD=0.89$, 95% CI [2.47, 2.85]). This pattern is interesting because prior studies showed that the SRSI has high sensitivity when it comes to detecting fabricated anxiety-related complaints (around 80%; Boskovic et al., 2020; Merten et al., 2019), whereas one would expect detection of feigned motor complaints (e.g., tics) to be low. However, it has to be noted that the purpose of the pseudosymptom scale is to detect a (usually generalized) tendency of respondents to over-report symptoms across different symptom domains. The SRSI subscales were not developed and not validated for identifying highly specific and circumscribed invalid symptom claims in isolation. Considering that the cut scores of the SRSI were based on the sum of endorsed pseudosymptoms, the plausibility of one subscale does not per se determine the detection accuracy.

Inspecting the prevalence estimates of genuine symptoms and pseudosymptoms in the SRSI led to more obvious

differences. Genuine symptoms were rated as significantly more frequent in the general public than pseudosymptoms. Looking at the subscales, we can see that this trend was consistent, meaning that scores on the five genuine scales included less variability than plausibility ratings. The lowest prevalence score for genuine symptoms attributed to cognitive complaints ($M=3.36$, $SD=0.46$, 95% CI [3.26, 3.46]), and the highest was given to the non-specific somatic issues ($M=3.89$, $SD=0.47$, 95% CI [3.78, 3.99]). These findings fit nicely with the results of Petrie et al. (2014), who observed that somatic complaints, such as headaches, stomach, and back pain, were the most frequently reported symptoms in the general population. For the pseudosymptom subscales, the lowest average rate was given to motor pseudosymptoms ($M=2.36$, $SD=0.51$, 95% CI [2.24, 2.47]), and the highest prevalence score was attributed to anxiety/depression/PTSD pseudosymptoms ($M=2.91$, $SD=0.53$, 95% CI [2.80, 3.02]). Higher ratings given to the psychological rather than physical complaints might again reflect the “belief bias” among psychology students (Lilienfeld et al., 2016). Specifically, psychologists are often trained to accept as a priori trustworthy even the most implausible reported experiences. This “truth bias” even for non-believable symptom reports is usually justified by referring to the subjective quality of psychological complaints (e.g., Noeker & Petermann, 2011).

However, if the large effect sizes of differences in both plausibility and prevalence ratings between genuine symptom scale and pseudosymptom scale are considered, one might think that the above-mentioned issue is trivial. Yet, inspecting the frequency of participants whose plausibility rating scores were at or above the mid-point, it is noticeable that the majority of students indeed perceived SRSI pseudosymptom subscales as semi-plausible. The only exception was the motor pseudosymptoms (40.2%). This trend disappeared in the prevalence ratings, indicating that students, despite perceiving symptoms as semi-plausible, did not expect such complaints to occur frequently in the general public. The one exception to this was the items of the anxiety/depression/PTSD-related pseudosymptom subscale (47.1%). This appears to confirm the presence of a belief bias (Lilienfeld et al., 2016) to be found when psychology undergraduates perceive a would-be-symptom to belong to the spectrum of common mental-health complaints.

Some limitations deserve mentioning. Our sample consisted of bachelor level students. Thus, it might be that some of the students were not yet familiar with clinical psychology, and the findings might differ were the study to be replicated with students of master level or even professionals in the clinical field. Second, the questionnaire was provided in English, which was not the mother tongue of most of the participating students. However, as the bachelor psychology program at our institution is international and delivered in English, resorting to the English-language

questionnaire version was considered appropriate. We did not check for the nationality of our sample, which is something that future investigation should include. Finally, the study was conducted online, which might have provided an opportunity for students to divide their attention and only partially focus on the task.

Overall, our results showed that psychology students rated genuine symptoms of the SRSI to be more plausible and more prevalent than pseudosymptoms. Yet, scores for the pseudosymptoms were, on average, graded as moderately plausible, that is much more plausible (and probable) than they are in truth. One could argue that this truth bias of junior psychology students is due to their incomplete training. However, some published results indicate that even professionals do not always value the exceptional nature of extreme or bizarre symptom claims (e.g., Cernovsky et al., 2019). The lack of skepticism among psychologists toward unusual or bizarre complaints may diminish the probability of them employing symptom validity tests and properly interpreting their outcomes. Hence, teaching junior psychologists about self-report-based SVTs, and training them to remain open-minded with respect to merely subjective symptom claims can increase the quality and validity of psychological assessments.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12207-021-09409-x>.

Declarations

Ethical Consideration This study was conducted in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

Conflict of Interest Two of the authors of this paper are also the developers of the instrument used in the study, the Self-Report Symptom Inventory (SRSI), which is commercially distributed by Hogrefe publishers, Göttingen, Germany.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bass, C., & Halligan, P. (2014). Factitious disorders and malingering: Challenges for clinical assessment and management. *The Lancet*, 383, 1422–1432. [https://doi.org/10.1016/S0140-6736\(13\)62186-8](https://doi.org/10.1016/S0140-6736(13)62186-8)
- Beach, S. R., Taylor, J. B., & Kontos, N. (2017). Teaching psychiatric trainees to “think dirty”: Uncovering hidden motivations and deception. *Psychosomatics*, 58, 474–482. <https://doi.org/10.1016/j.psych.2017.04.005>
- Bianchini, K. J., Greve, K. W., & Glynn, G. (2005). On the diagnosis of malingered pain-related disability: Lessons from cognitive malingering research. *The Spine Journal*, 5, 404–417. <https://doi.org/10.1016/j.spinee.2004.11.016>
- Boskovic, I., Merckelbach, H., Merten, T., Hope, L., & Jelicic, M. (2020). The Self-Report Symptom Inventory as an instrument for detecting symptom over-reporting. *European Journal of Psychological Assessment*, 36, 730–739. <https://doi.org/10.1027/1015-5759/a000547>
- Boskovic, I., Van Der Heide, D., Hope, L., Merckelbach, H., & Jelicic, M. (2017). Plausibility judgments of atypical symptoms across cultures: An explorative study among Western and non-Western experts. *Psychological Injury and Law*, 10, 274–281. <https://doi.org/10.1007/s12207-017-9294-6>
- Bush, S. S., Heilbronner, R. L., & Ruff, R. M. (2014). Psychological assessment of symptom and performance validity, response bias, and malingering: Official position of the Association for Scientific Advancement in Psychological Injury and Law. *Psychological Injury and Law*, 7, 197–205. <https://doi.org/10.1007/s12207-014-9198-7>
- Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2). Manual for administration, scoring, and interpretation.* (rev). University of Minnesota Press.
- Cernovsky, Z., Mendonça, J. D., Oyewumi, L. K., Ferrari, J. R., Sidhu, G., & Campbell, R. (2019). Content validity of the psychosis subscale of the Structured Inventory of Malingered Symptomatology (SIMS). *International Journal of Psychology and Cognitive Science*, 5, 121–127.
- Chafetz, M. D., Abrahams, J. P., & Kohlmaier, J. (2007). Malingering on the social security disability consultative exam: A new rating scale. *Archives of Clinical Neuropsychology*, 22, 1–14. <https://doi.org/10.1016/j.acn.2006.10.003>
- Chafetz, M. D., Williams, M. A., Ben-Porath, Y. S., Bianchini, K. J., Boone, K. B., Kirkwood, M. W., & Ord, J. S. (2015). Official position of the American Academy of Clinical Neuropsychology Social Security Administration policy on validity testing: Guidance and recommendations for change. *The Clinical Neuropsychologist*, 29, 723–740. <https://doi.org/10.1080/13854046.2015.1099738>
- Dandachi-FitzGerald, B., Merckelbach, H., Bošković, I., & Jelicic, M. (2020). Do you know people who feign? Proxy respondents about feigned symptoms. *Psychological Injury and Law*, 13, 225–234. <https://doi.org/10.1007/s12207-020-09387-6>
- Dandachi-FitzGerald, B., Ponds, R. W., & Merten, T. (2013). Symptom validity and neuropsychological assessment: A survey of practices and beliefs of neuropsychologists in six European countries. *Archives of Clinical Neuropsychology*, 28, 771–783. <https://doi.org/10.1093/arclin/act073>
- Geurten, M., Meulemans, T., & Seron, X. (2018). Detecting over-reporting of symptoms: The French version of the Self-Report Symptom Inventory. *The Clinical Neuropsychologist*, 32(Suppl. 1), 164–181. <https://doi.org/10.1080/13854046.2018.1524027>
- Greve, K. W., Ord, J. S., Bianchini, K. J., & Curtis, K. L. (2009). Prevalence of malingering in patients with chronic pain referred for psychologic evaluation in a medico-legal context. *Archives of*

- Physical Medicine and Rehabilitation*, 90, 1117–1126. <https://doi.org/10.1016/j.apmr.2009.01.018>
- Lees-Haley, P. R., English, L. T., & Glenn, W. J. (1991). A Fake Bad Scale on the MMPI-2 for personal injury claimants. *Psychological Reports*, 68, 208–210. <https://doi.org/10.2466/pr0.1991.68.1.203>
- Lilienfeld, S. O., Watts, A. L., Robinson, B. A., & Smith, S. F. (2016). Scientific research in forensic samples. In Cima, M. (ED.), *The Handbook of Forensic Psychopathology and Treatment* (pp. 25–51). New York, NY: Routledge.
- Martin, P. K., Schroeder, R. W., & Odland, A. P. (2015). Neuropsychologists' validity testing beliefs and practices: A survey of North American professionals. *The Clinical Neuropsychologist*, 29, 741–776. <https://doi.org/10.1080/13854046.2015.1087597>
- Merten, T., Giger, P., Merckelbach, H., & Stevens, A. (2019). Self-Report Symptom Inventory (SRSI) - deutsche Version. Manual [Manual of the German version of the Self-Report Symptom Inventory]. Göttingen, Germany: Hogrefe.
- Merten, T., Merckelbach, H., Giger, P., & Stevens, A. (2016). The Self-Report Symptom Inventory (SRSI): A new instrument for the assessment of distorted symptom endorsement. *Psychological Injury and Law*, 9, 102–111. <https://doi.org/10.1007/s12207-016-9257-3>
- Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, 24, 1094–1102. <https://doi.org/10.1076/jcen.24.8.1094.8379>
- Noeker, M., & Petermann, F. (2011). Simulation neurologischer versus psychischer Beschwerden. Notwendigkeit unterschiedlicher Validierungsstrategien [Malingering of neurological vs. mental complaints. Necessity of different validation strategies]. *Psychotherapeut*, 56, 449–454. <https://doi.org/10.1007/s00278-011-0851-2>
- Petrie, K. J., Faasse, K., Crichton, F., & Grey, A. (2014). How common are symptoms? Evidence from a New Zealand national telephone survey. *British Medical Journal Open*, 4, e005374. <https://doi.org/10.1136/bmjopen-2014-005374>
- Schmand, B., Lindeboom, J., Schagen, S., Heijt, R., Koene, T., & Hamburger, H. L. (1998). Cognitive complaints in patients after whiplash injury: The impact of malingering. *Journal of Neurology, Neurosurgery, and Psychiatry*, 64, 339–343. <https://doi.org/10.1136/jnnp.64.3.339>
- Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law*, 25, 183–189. <https://doi.org/10.1037/t04573-000>
- Stevens, A., Schmidt, D., & Hautzinger, M. (2018). Major depression - a study on the validity of clinicians' diagnoses in medicolegal assessment. *The Journal of Forensic Psychiatry & Psychology*, 29, 794–809. <https://doi.org/10.1080/14789949.2018.1477974>
- van Helvoort, D., Merckelbach, H., & Merten, T. (2019). The Self-Report Symptom Inventory (SRSI) is sensitive to instructed feigning, but not to genuine psychopathology in male forensic inpatients: An initial study. *The Clinical Neuropsychologist*, 33, 1069–1082. <https://doi.org/10.1080/13854046.2018.1559359>
- van Impelen, A., Merckelbach, H., Jelicic, M., & Merten, T. (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. *The Clinical Neuropsychologist*, 28, 1336–1365. <https://doi.org/10.1080/13854046.2014.984763>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.