The Self-Report Symptom Inventory (SRSI) is sensitive to instructed feigning, but not to genuine psychopathology in male forensic inpatients: an initial study

Daniël van Helvoort, Harald Merckelbach & Thomas Merten

Published online: 30 Jan 2019.

Submit your article to this journal ⬀

Article views: 36

View Crossmark data ⬀

Routledge
Taylor & Francis Group

The Self-Report Symptom Inventory (SRSI) is sensitive to instructed feigning, but not to genuine psychopathology in male forensic inpatients: an initial study

Daniël van Helvoort[a] (iD), Harald Merckelbach[a] and Thomas Merten[b]

[a]Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands; [b]Department of Neurology, Vivantes Klinikum im Friedrichshain, Berlin, Germany

ABSTRACT

Objective: The *Self-Report Symptom Inventory* (SRSI) is a new symptom validity test that, unlike other symptom over-reporting measures, contains both genuine symptom and pseudosymptom scales. We tested whether its pseudosymptom scale is sensitive to genuine psychopathology and evaluated its discriminant validity in an instructed feigning experiment that relied on carefully selected forensic inpatients ($n = 40$).
Method: We administered the SRSI twice: we instructed patients to respond honestly to the SRSI (T1) and then to exaggerate their symptoms in a convincing way (T2).
Results: On T1, the pseudosymptom scale was insensitive to patients' actual psychopathology. Two patients (5%) had scores exceeding the liberal cut point (specificity = 0.95) and no patient scored above the more stringent cut point (specificity = 1.0). Also, the SRSI cut scores and ratio index discriminated well between honest (T1) and exaggerated (T2) responses (AUCs were 0.98 and 0.95, respectively).
Conclusions: Given the relatively few false positives, our data suggest that the pseudosymptom scale of the SRSI is a useful measure of symptom over-reporting in forensic treatment settings.

Introduction

Poor symptom validity refers to patients' exaggerated self-reports of impairments or symptoms. Because unstructured clinical evaluations of symptom validity have a poor track record, with error rates often exceeding 20% (e.g. Dandachi-FitzGerald, Merckelbach, & Ponds, 2017; Resnick & Harris, 2002; Rosen & Phillips, 2004), neuropsychologists developed tasks and tests that may help to detect poor symptom validity (Sweet & Guidotti Breting, 2013). These tools have now become an essential part of neuropsychological evaluations (Bush et al., 2005) and routine administration

of validity tests is recommended by various professional organizations (e.g. Chafetz et al., 2015; Heilbronner et al., 2009; Institute of Medicine, 2015). Validity tests aim to measure negative response bias during diagnostic assessment, which may occur for a variety of reasons (e.g. feigning, careless responding; e.g. Merckelbach, Boskovic, Pesy, Dalsklev, & Lynn, 2017). Negative response bias may take the form of underperformance on cognitive tasks and/or over-endorsement of items on symptom scales (i.e. over-reporting). These phenomena are gauged by embedded or stand-alone performance validity tests (PVTs) and symptom validity tests (SVTs), respectively. In clinical practice, both types of validity tests should preferably be used, as some patients may fail one but not the other, irrespective of the diagnostic referral question (Dandachi-FitzGerald, Ponds, Peters, & Merckelbach, 2011). Poor symptom validity as indexed by underperformance and/or over-reporting occurs on a non-trivial scale in various settings. Prevalence rates may be as high as 10% in regular medical settings and 30% in forensic evaluation settings (Ardolf, Denney, & Houston, 2007; Dandachi-FitzGerald et al., 2011; Mittenberg, Patton, Canyock, & Condit, 2002).

Surveys among experts have shown that the *Structured Inventory of Malingered Symptomatology* (SIMS; Widows & Smith, 2005) is the most widely used stand-alone SVT (Dandachi-FitzGerald, Ponds, & Merten, 2013; Martin, Schroeder, & Odland, 2015). The SIMS consists of 75 true-false items that span five symptom domains (i.e. psychosis, neurology, amnesia, mental disability, and affective disorder). The SIMS has reasonable psychometric properties (e.g. high sensitivity), but is also subject to a number of limitations. One important limitation is the suboptimal specificity of the SIMS. In patients with serious psychopathology (e.g. schizophrenia; van Impelen, Merckelbach, Jelicic, & Merten, 2014), the SIMS may produce false positives of poor symptom validity. In addition, the SIMS may be easily recognizable as an SVT because it exclusively relies on implausible (e.g. extreme, bizarre, or rare) symptoms. Furthermore, several symptom domains that may be relevant in litigation cases are not covered by the SIMS (e.g. fatigue, pain, anxiety, trauma). For example, one recent study found that the SIMS possesses only low to moderate sensitivity to detect over-reporting of posttraumatic stress disorder symptoms (PTSD; Parks, Gfeller, Emmert, & Lammert, 2016).

The *Self-Report Symptom Inventory* (SRSI; Merten, Giger, Merckelbach, & Stevens, in press; Merten, Merckelbach, Giger & Stevens, 2016) was recently developed as an alternative to the SIMS. The 100 symptoms listed by the SRSI are organized into two main scales of 50 items each. The genuine symptom main scale addresses credible manifestations of (1) cognitive, (2) depressive, (3) pain, (4) non-specific somatic, and (5) PTSD/anxiety symptoms. The pseudosymptom main scale lists non-credible symptoms in these domains. Apart from the 100 symptom items, the SRSI contains seven additional items that measure test taking attitude. They comprise two warming-up items that inquire whether the respondent is willing to cooperate and five items that are evenly distributed over the instrument and that check for careless responding. The rationale underlying the SRSI is that patients who over-report symptoms tend to endorse genuine symptoms and pseudosymptoms in an indiscriminate way. Besides the number of endorsed pseudosymptoms, one additional parameter that can be

derived from the SRSI is the ratio of endorsed pseudosymptoms to endorsed genuine symptoms. A high ratio would be indicative of indiscriminate symptom reporting.

In their psychometric evaluation of the SRSI in a pooled sample of healthy controls, patients involved in litigation, and young prison inmates ($n = 387$), Merten et al. (2016) found that the number of endorsed pseudosymptoms on the SRSI correlated positively with SIMS scores ($r = 0.82$), but negatively ($r = -0.45$) with performance on the Word Memory Test (WMT; Green, 2003). Internal consistency estimates for the main SRSI scales were high (α's > 0.90) and their test-retest stability was good (>0.85). A receiver operating characteristic (ROC) analysis revealed an area under the curve (AUC) of 0.93 (95% Confidence Interval = [0.90, 0.96]) when SIMS was used as the external criterion. Merten et al. (2016) recommended the following cut scores: >6 endorsed pseudosymptoms for screening purposes (sensitivity = 83%, false positives < 10%) and >9 endorsed pseudosymptoms for diagnostic purposes (sensitivity = 62%, false positives < 5%).

Recent studies found further support for the psychometric qualities of the SRSI. Giger and Merten (in press) instructed non-clinical participants ($n = 40$) to complete the SRSI under honest and instructed feigning conditions and replicated the good classificatory accuracy for both cut scores (sensitivity > 90%; specificity = 100%). Geurten, Meulemans, and Seron (2018) conducted exploratory and confirmatory factor analyses on the SRSI in healthy adults and found a two-factor structure corresponding to genuine and pseudosymptoms. The main scales had good internal reliabilities (α's > 0.85). Discriminant validity was tested by administering the SRSI to healthy controls ($n = 575$), patients with genuine cognitive impairments ($n = 13$), and healthy participants who were instructed to feign cognitive impairments ($n = 68$). The SRSI pseudosymptom scale differentiated reasonably well between patients and instructed feigners, with AUCs being >0.75.

Merckelbach, Merten, Dandachi-FitzGerald, and Boskovic (2018) instructed undergraduates to complete the SRSI in an honest way ($n = 51$) or to feign pain symptoms ($n = 54$) or anxiety symptoms ($n = 53$) on the SRSI. Main and subscales had high internal consistencies (all α's > 0.80). Overall, the AUC was good: 0.90 (95% Confidence Interval = [0.85, 0.95]). For the cut scores of >6 and >9, sensitivity rates were ≥48% and false-positive rates were ≤10%. Moreover, support was found for the provisional criterion that index ratios of >0.288 (i.e. more than about 1 pseudosymptom per 3 genuine symptoms) are suspect. In the Merckelbach et al. (2018) sample, the mean ratio was 0.43 (95% Confidence Interval = [0.33, 0.52]) for feigners and 0.20 (95% Confidence Interval = [0.14, 0.27]) for honest controls. A ROC analysis of the ratio index data yielded an AUC of 0.80 (95% Confidence Interval = [0.73, 0.86]).

Stevens, Schmidt, and Hautzinger (2018) administered the SRSI and a structured psychiatric interview to claimants ($n = 127$) who had been diagnosed with a major depression. None of them fulfilled the major depression criteria addressed during a structured clinical interview and 40% endorsed a heightened number of SRSI pseudosymptoms. Most importantly, their SRSI scores were not dependent on educational background. Also, Lehrner (in Merten et al., in press) observed in a sample of memory clinic patients ($n = 106$) that cognitive impairments in this group were not associated with elevated endorsement of pseudosymptoms on the SRSI.

So far, research has not examined to what extent patients with genuine psychopathology endorse pseudosymptoms on the SRSI. If such patients would endorse SRSI pseudosymptoms to a degree that exceeds the cut points, false-positive classifications might ensue (i.e. bona fide patients misclassified as over-reporters). Finding such false positives on a fairly large scale would casts serious doubts on the diagnostic safety of the SRSI (Hartman, 2002; Morel & Marshman, 2008).

With this in mind, we wanted to test the false-positive potential of the SRSI. We carefully selected a sample of forensic inpatients from a high security, post-trial treatment setting and administered the SRSI twice to them. Our study was predicated on the assumptions that relative to non-forensic samples and claimants in pre-trial evaluation settings, forensic inpatients more often would suffer from serious psychopathology and would have less incentive to exaggerate their symptoms. We conducted a test-retest study. On the first test, patients were instructed to respond honestly to the SRSI (*T1*; honest condition). We hypothesized that patients would endorse many SRSI genuine symptoms but only few pseudosymptoms. Such a pattern would suggest that the SRSI pseudosymptoms are immune to genuine psychopathology. On the second test, to test the discriminant power of the SRSI, we instructed patients to exaggerate symptoms on the SRSI at *T2* (instructed feigning condition). We expected that a majority would exceed the cut scores for the SRSI pseudosymptoms (>6 and >9) as well as the ratio index (>0.288). Also, we anticipated that ROC analyses involving *T1* and *T2* data would yield high sensitivity and specificity for the detection of symptom over-reporting in the present sample.

Methods

Participants

The study was conducted at de Rooyse Wissel, a maximum security forensic psychiatric hospital in Oostrum, The Netherlands. The hospital provides mandatory treatment to mentally disordered offenders. Every 1–2 years, the criminal court re-evaluates whether or not to continue a patient's treatment, based on reports by an independent expert committee that looks into the patient's recidivism risk. Thus, for patients in the current sample, mandatory inpatient treatment is contingent on regular psychiatric evaluations and therefore, they generally have little reason to exaggerate their symptoms. Doing so might prolong their mandatory stay in high security facilities. A priori analysis indicated a minimum required sample size of 34 at a power of 0.80, one-tailed (computed by G*Power; Faul, Erdfelder, Lang, & Buchner, 2007). Taking potential drop-outs into account, we recruited 41 patients. Participation was compensated with a 7.50 euro (approximately 9 dollars) gift voucher. The study was approved by the Ethical Committees of the Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands (Master_184_06_10_2017) and de Rooyse Wissel, Oostrum, The Netherlands.

We excluded participants based on criteria that we a priori thought to be indicative of invalid responding. Thus, patients were excluded based on the following criteria: (1) IQ lower than 70; (2) suspected substance abuse at time of testing and/or suspected drug trading while in the facility; (3) failure on ≥ 2 SVTs and/or PVTs in any previous

assessment. Moreover, we excluded patients who displayed a pattern of chronic unco-operativeness, or had a preference to stay within the hospital. Specifically, we excluded: (4) patients who had been in the current forensic hospital for longer than ten years; and (5) patients who were indicated for "longstay" (i.e. long-term non-treatment hospitalization for non-responsive patients). Furthermore, based on the most recent available bi-annual treatment report, we excluded: (6) patients who displayed a prolonged pattern of treatment frustration; (7) patients who were recently isolated in an extra-security setting due to a violent incident; and (8) patients with a "time-out" or crisis status (e.g. due to acute psychosis). Patients were also screened for indicators of dissimulation (e.g. social desirability; faking good). Particularly, patients were excluded when (9) dissimulative behavior was repeatedly mentioned in their patient files or noted by their therapist. Applying these selection criteria led to the exclusion of 60% of potential participants.

During testing, one patient appeared under the influence of drugs and was excluded from the analysis. Also, one patient failed to complete the SRSI at *T2* (the patient reported not to feel comfortable with the instructed feigning role), but we included his *T1* data. This explains the fluctuating degrees of freedom in the statistical tests to be reported below. The remaining forensic inpatients ($n = 40$) consisted of male adults with a mean age of 41.6 years ($SD = 9.24$, range $= 20–62$). Their average stay in the clinic was 38.7 months ($SD = 35.0$, range $= 1–119$, capped at <120, see exclusion criteria). As to their index crimes, 72% of patients had committed a non-sexual violent crime (murder, manslaughter, and/or assault), 40% a sexual crime (adult and/or minor victim), and 35% property theft, damage or arson. As to their psychopathological status, 87.5% had been diagnosed (see below) with a personality disorder, 82.5% with a substance problem (abuse and/or dependence), and 35% with a psychotic disorder. Also, 27.5% had borderline IQ levels between 70 and 79. Forty-five percent of patients were prescribed antipsychotics. Table 1 gives a summary of the background data.

Measures and materials

Information about patients' index crime(s), DSM-IV diagnoses, IQ scores, medication, and prior performance on validity measures was retrieved from patient files after patients had given explicit and written permission for this. Diagnoses reported within patients' files were made by trained clinicians who employed standard measures with adequate psychometric properties (e.g. the Structured Clinical Interview for DSM-IV Axis I Disorders; the Structured Clinical Interview for DSM-IV Axis II Disorders; the Minnesota Multiphasic Personality Inventory; the Structured Interview for DSM-IV Personality; the Wechsler Adult Intelligence Scale, 3rd/4th edition).

Self-Report Symptom Inventory (SRSI)
The SRSI is a 107-item self-report instrument to measure symptom over-reporting (Merten et al., 2016). Both main scales – genuine symptoms and pseudosymptoms – consist of five subscales that each contain 10 items. Examples of genuine symptoms are "I can barely remember names'; 'My life is strongly affected by pain"; "I am often

Table 1. Demographic characteristics of the forensic inpatient participants ($n = 40$).

Characteristic	n	%
Crime type[a]		
Murder or manslaughter	15	37.5
Assault	20	50
Sexual violence (minor)	9	22.5
Sexual violence (adult)	7	17.5
Property	8	20
Arson or vandalism	7	17.5
Primary Diagnoses[b]		
Substance dependence	24	60
Substance abuse	15	37.5
Psychotic disorder	14	35
Borderline intelligence (IQ 70–79)	11	27.5
Paraphilia	6	15
ADHD[c] impulse regulation disorder	5	12.5
Autism spectrum disorder	3	7.5
Anxiety disorder/PTSD[d]	3	7.5
Dysthymia	1	2.5
None	9	22.5
Personality Disorder		
Personality disorder unspecified	23	57.5
Antisocial	6	15
Schizotypal	3	7.5
Narcissistic	2	5
None	7	17.5
Medication		
Antipsychotics	18	45
Somatic	8	20
Antidepressants	7	17.5
Stimulants	5	12.5
GABAergic/Sedatives	4	10
Antiadrenergic	1	2.5
None	10	25

Notes. Totals do not add up to $n = 40$ due to the presence of multiple characteristics within some patients.
[a]Crime type refers only to the index crimes for which mandatory treatment was given.
[b]Here we refer to the broader disorder category rather than specific disorders.
[c]Attention deficit/hyperactivity disorder.
[d]Posttraumatic stress disorder.

exhausted"; and "I have nightmares about things that happened to me". Examples of pseudosymptoms are "I have completely lost my sense of taste"; "I can hardly raise my shoulders anymore"; "Even when I am touched lightly I flinch with pain"; and "I can't remember what happened to me, but I constantly dream about it".

Exit questionnaire

After the experiment, patients completed an exit questionnaire that contained open-ended questions about the SRSI (*T1*) and the comprehensibility and strangeness of its items. Patients were also asked to rate their feigning performance at *T2*. Specifically, using 5-point scales, they rated how convincing they thought their feigning was at *T2*, how difficult it was to play the feigning role, and how skillful they generally were in exaggerating complaints.

Additional measures

We also administered the Dissociative Experiences Scale Taxon (DES-T; Waller, Putnam, & Carlson, 1996) and the 20-item Toronto Alexithymia Scale (TAS-20; Bagby, Parker, &

Taylor, 1994) at baseline, which measure trait dissociation and alexithymia, respectively. The data obtained with these scales will not be further considered in the current paper.[1]

Procedure

To establish rapport, patients were approached and tested individually. Patients were given extensive verbal and written information that emphasized the voluntary and anonymous nature of participation. They were told that the study was about the association between common symptoms and traits. Patients were also informed that they would fill out questionnaires under various instructions, that this would take 30 min, and that they would be compensated with a 7.50 euro gift voucher. Patients were given a week to think about their willingness to participate and then were revisited. They were given the opportunity to ask questions and were reminded that participation was anonymous and voluntary. Informed consent was obtained from all participants.

In scheduled individual meetings, patients completed the TAS-20 and DES-T, which were presented as questionnaires on personality traits. Next, the patients were instructed to fill out the SRSI (*T1*) in an honest way. The SRSI was presented as a questionnaire about cognitive, depressive, somatic, pain, and anxiety problems. After completion of the SRSI (*T1*), patients were instructed to fill out the SRSI once more but now while playing the role of a person who exaggerates his problems. Specifically, patients were told the following: "Imagine that I am a psychologist hired by an insurance company. You may gain a large sum of money from the insurance company when you convince the insurer that you are unable to work. You can do this by exaggerating your complaints on the five topics that you saw in the previous questionnaire and that addressed cognitive complaints, depression complaints, physical complaints, pain complaints, and anxiety complaints. Thus, we ask you to fill out the previous questionnaire once more, but now while assuming this role".

If necessary, further clarifications were given by the experimenter (the first author). Patients were instructed to calibrate their simulation appropriately, in line with instructions by Merckelbach et al. (2018): "Present your symptoms in a convincing way. Do not select complaints that are perhaps not very plausible in the eyes of medical experts. Be clever in which symptoms you select by keeping in mind the goal you want to achieve by exaggerating your problems. If you are convincing in feigning complaints, you will participate in a lottery, in which you can win an extra 20 euros (in fact, all participants took part in the lottery). The patients were then given the SRSI a second time (*T2*). Finally, the patients were asked in a brief exit questionnaire to give their opinion about the SRSI at *T1* and the instructed feigning at *T2*. We did not counterbalance *T1* and *T2* instructions because feigning at *T1* could have distorted genuine responding at *T2* (Merckelbach, Jelicic & Pieters, 2011). After completion of the study, all participants were debriefed and fully informed about the research questions."

Statistical analyses

We performed a series of paired-sample *t* tests (Time: honest responding [*T1*] vs. experimental simulation [*T2*]) on the following variables: SRSI genuine symptom scale total, SRSI pseudosymptom scale total, and the ratio index (Σ pseudosymptoms/ Σ genuine symptoms). Using the recommended cut scores, we examined the SRSI's sensitivity and specificity for the detection of instructed feigning. For the combined *T1* and *T2* SRSI data, ROC analyses were performed. We examined the AUC for the total number of pseudosymptoms and for the ratio index. All data associated with this article can be found in DataverseNL at https://hdl.handle.net/10411/QQE3VA.

Results

Manipulation check

With regard to the warming-up items of the SRSI at *T1*, all participants confirmed their willingness to fill out the scale in an honest way. As to the open question to the patients about whether they had noted anything remarkable about the SRSI items, only 7.5% responded that some symptoms made a strange impression. Patients rated several questions on 5-point scales (anchors: 1 = very poor/very ineffective; 5 = very good/very effective). A majority evaluated the comprehensibility of SRSI items favorably (*M* = 3.95; *SD* = 0.90; only three respondents gave a rating lower than 3; two patients did not provide a rating). Overall, patients thought that they had been effective in playing their role at *T2* (*M* = 3.46; *SD* = 1.05; none of the respondents gave a rating lower than 2; one patient did not provide a rating). Patients evaluated their 'general skill in credible exaggeration' as modest (*M* = 3.00; *SD* = 0.95).

Discriminant validity

Table 2 summarizes SRSI scores obtained during honest responding (*T1*) and instructed symptom exaggeration (*T2*). Cronbach's alpha's for the genuine and pseudosymptom

Table 2. Symptom endorsement for SRSI main and subscales at *T1* (honest responding) and *T2* (instructed feigning) in forensic inpatients (*n* = 40).

Variables	T1 M (SD)	T1 R.ange	T2 M (SD)	T2 Range	Cohen's d	t
Genuine symptom scale total	12.70 (7.31)	2–30	36.15 (7.51)	17–47	2.48	15.49*
Cognitive	1.85 (2.26)	0–8	7.64 (2.19)	2–10	1.97	12.28*
Depressive	2.35 (1.31)	1–6	6.54 (1.67)	3–10	2.11	13.19*
Pain	2.10 (2.11)	0–8	7.26 (2.75)	0–10	1.47	9.16*
Non-specific somatic	2.55 (2.45)	0–9	7.54 (2.64)	0–10	1.43	8.92*
PTSD/Anxiety	3.85 (2.18)	0–10	7.18 (2.09)	2–10	1.14	7.15*
Pseudosymptom scale total	1.63 (2.31)	0–8	24.54 (13.39)	4–49	1.79	11.16*
Cognitive	0.38 (0.67)	0–2	4.97 (2.87)	0–10	1.69	10.53*
Neurological: motor	0.30 (0.69)	0–3	4.10 (3.02)	0–10	1.27	7.94*
Neurological: sensory	0.38 (0.95)	0–4	4.21 (3.35)	0–10	1.15	7.18*
Pain	0.28 (0.51)	0–2	5.51 (3.42)	0–10	1.58	9.86*
Anxiety/Depression	0.30 (0.61)	0–3	5.74 (2.87)	1–9	1.89	11.82*
Ratio index	0.12 (0.12)	0.01–0.44	0.67 (0.35)	0.14–1.41	1.64	10.25*

Notes. T1 data are based on *n* = 40, T2 data on *n* = 39 patients.
*p < .001.

scales at *T1* were 0.86 and 0.85, respectively. At *T1*, patients endorsed few pseudo-symptoms ($M = 1.63$, $SD = 2.31$, range from 0 to 8), on average far below the cut scores of >6 (screening cutoff) and >9 (routine cutoff). Only 2 of the 40 patients scored above the cut score of >6 (specificity $= 0.95$), neither of whom attained a score >9 (specificity $= 1.0$). In contrast, at *T2*, patients endorsed many pseudosymptoms ($M = 24.54$, $SD = 13.39$, range from 4 to 49). In total, 36 of the 39 patients had a pseudosymptom score that exceeded the cut score of 6 (sensitivity $= 92.3\%$), whereas 31 patients had a score above the cut score of 9 (sensitivity $= 79.5\%$). The increase in pseudosymptom endorsement from *T1* to *T2* ($M = 22.90$; $SD = 12.81$; 95% Confidence Interval $= [18.75, 27.05]$) was statistically significant and the corresponding effect size was large [$t(38)=11.16$, $p<.001$, Cohen's $d = 1.79$]. Patients endorsed a moderate number of genuine symptoms at *T1* ($M = 12.70$, $SD = 7.31$, range from 2 to 30), but endorsed many more genuine symptoms at *T2* ($M = 36.15$, $SD = 7.51$, range from 17 to 47). Again, the increase ($M = 23.26$; $SD = 9.38$; 95% Confidence Interval $= [20.22, 26.30]$) was statistically significant and was associated with a large effect size [$t(38)=15.49$, $p<.001$, $d = 2.48$]. An ROC analysis performed on the number of endorsed pseudosymptoms at *T1* and *T2* yielded an AUC of 0.983 (95% Confidence Interval $= [0.96, 1.00]$), which is, according to widely accepted standards, excellent.

Table 2 also shows the ratio scores (Σ endorsed pseudosymptoms/Σ endorsed genuine symptoms) at *T1* and *T2*. The ratio was significantly higher at *T2* ($M = 0.67$, $SD = 0.35$, range 0.14–1.41) than at *T1* ($M = 0.12$, $SD = 0.12$, range 0.01–0.44), [$t(38)=10.25$, $p<.001$, $d = 1.64$]. The mean increase in ratio ($M = 0.56$; $SD = 0.34$; 95% Confidence Interval $= [0.45, 0.67]$) was statistically significant and the corresponding effect size was large [$t(38)=10.25$, $p<.001$, $d = 1.65$]. At *T1*, only 2 of the 40 honest patients were incorrectly classified (specificity $= 95\%$), using the ratio cut score of >0.288. At *T2*, 31 of the 39 feigners were correctly classified (sensitivity $= 79.5\%$). An ROC analysis performed on ratio data yielded an AUC of 0.952 (95% Confidence Interval $= [0.91, 0.99]$).

Discussion

The SRSI is a new SVT that was designed to detect one particular manifestation of poor symptom validity, namely symptom over-reporting (Merten et al., 2016, in press). The instrument was developed against the background of a remarkable lack of stand-alone SVTs in most European languages. The only instrument that is currently widely available is the SIMS, but this SVT focuses on blatant forms of symptom exaggeration (e.g. total memory loss) typically seen in criminal forensic settings. Also, the SIMS has been widely publicized in the last few years and may be vulnerable to preparation by patients on the basis of previous forensic reports or Internet searching (e.g. Ruiz, Drake, Glass, Marcotte, & van Gorp, 2002; Wetter & Corrigan, 1995). Furthermore, the development of additional SVTs is of crucial importance because ideally, the diagnostic expert administers multiple validity instruments (e.g. Slick, Sherman, & Iverson, 1999).

Recent studies have provided supportive evidence for the sensitivity of the SRSI, i.e. its diagnostic power to detect over-reporting (Geurten et al., 2018; Giger & Merten, in press; Merckelbach et al., 2018). Its specificity has been less well-researched. Lehrner

(in Merten et al., in press) observed in a sample of memory clinic patients that cognitive impairments per se were not associated with elevated endorsement of pseudo-symptoms on the SRSI. Furthermore, Stevens et al. (2018) found no correlation between pseudosymptoms and educational background in their sample of claimants. Still, so far, no study looked at whether patients who suffer from serious psychopathology and who have no incentive to exaggerate, score below the thresholds on the pseudosymptom scale of the SRSI. Finding such a pattern would support the safety of the SRSI, i.e. its ability to avoid false-positive classifications. The current study addressed this issue by relying on a forensic inpatient sample from a high security treatment setting. The sample was carefully recruited in the sense that patients suspected of over-reporting tendencies, under-reporting tendencies, and/or lack of motivation were excluded (for the importance of such pre-selection see: van Egmond & Kummeling, 2002; van Egmond, Kummeling, & Balkom, 2005). Thus, we had good reasons to believe that our carefully selected sample of forensic inpatients were willing and able to report honestly about their genuine psychopathology.

The main findings of our study can be summarized as follows. First, and most importantly, the pseudosymptom scale of the SRSI was largely immune to patients' psychopathology. That is, when instructed to fill out the SRSI in an honest way (*T1*), forensic inpatients endorsed a moderate number of genuine symptoms, but only few pseudosymptoms. Two patients scored above the liberal screening cut score of 6, which corresponds to a false-positive rate of 5%; none of the patients scored above the more conservative cut score of 9 (false-positive rate = 0%).

Second, we found that the cut scores of >6 (for screening) and >9 (for diagnosis) differentiated excellently between patients' SRSI pseudosymptom scores at *T1* (i.e. when instructed to respond honestly) and their pseudosymptom scores at *T2* (i.e. when instructed to exaggerate symptoms in a convincing way). Furthermore, the ratio index indicated that patients increased their pseudosymptom endorsement relatively more than their genuine symptom endorsement from *T1* to *T2*. Our data confirm that ratio indices >0.288 flag over-reporting tendencies (i.e. poor symptom validity). Indeed, in the current study, the ratio index was equally diagnostic as the >6 cut score. These results are in line with those of previous studies that included a wide variety of samples such as population-based healthy controls, clinical groups, and litigating claimants (Geurten et al., 2018; Merckelbach et al., 2018; Merten et al., in press). They converge upon the conclusion that the SRSI is reasonably effective in detecting poor symptom validity. What the current study adds to this is that the SRSI appears to be a safe instrument, i.e. does not produce false-positive cases of poor symptom validity in patients with genuine mental health issues.

Several limitations of our study deserve comment. To begin with, our study was based on a selective group of patients. Although the careful recruitment of our sample guaranteed that participants suffered from psychopathology and had no incentive to exaggerate their problems, it remains to be seen whether our results can be generalized to other clinical groups in other types of settings. For example, there were no women in our sample. Also, certain types of psychopathology were over-represented (e.g. personality disorders). Furthermore, the mean IQ of our sample was in the low average to average range. Therefore, we opted to keep our self-report battery as brief

as possible to minimize inattentive or fatigued responding. For this reason, we could not include a measure of faking good such as the Supernormality Scale – Revised (SS-R; Cima, van Bergen, & Kremer, 2008). However, without such a measure, exclusion of patients who engaged in faking good may have been suboptimal in our study. Future SRSI validation studies may want to include more diverse clinical groups and the SS-R or another measure of faking good.

A second limitation of our study was that we instructed patients to feign symptoms and did not include a separate "known" group that exhibits poor symptom validity. Instructed feigners may behave differently than such known groups (Niesten et al., 2017). Specifically, instructed feigning may inflate sensitivity and specificity rates of an SVT. Therefore, the diagnostic accuracy parameters of the SRSI provided in the current study should be interpreted with caution.

For neuropsychologists who want to include an SVT in their test battery, the SRSI may hold several advantages over the SIMS. First, as our data suggest, unlike the SIMS (Parks et al., 2016; van Impelen et al., 2014), the SRSI pseudosymptom scale appears to be rather immune to actual psychopathology (e.g. cognitive deficits; anxiety disorders), thereby lowering the risks of false-positive classifications. Second, the SRSI contains multiple indices of symptom over-reporting (number of endorsed pseudo-symptoms; ratio). Third, unlike the SIMS, the SRSI contains symptom scales that are relevant to symptomatology typically prominent in civil litigation cases, such as pain, fatigue, and PTSD. With the exception of the Inventory of Problems-29 (IOP-29; Viglione, Giromini, & Landis, 2016) and the Modified Somatic Perception Questionnaire and the Pain Disability Index (MSPQ and PDI; Crighton, Wygant, Applegate, Umlauf, & Granacher, 2014), no other SVTs address these types of problems. It would be inform-ative if future studies would compare psychometric performance of the SRSI, IOP-29, MSPQ, and PDI across various clinical samples.

In sum, we found in a selective sample of forensic inpatients recruited from a high security treatment setting, the SRSI pseudosymptom scales to be immune to actual psychopathology. Also, the SRSI cut scores and ratio index discriminated effectively between SRSI scores obtained in an honest responding condition and SRSI scores obtained when patients were instructed to feign their symptoms. Our data suggest that in a forensic treatment setting, the SRSI is a promising measure of symptom over-reporting.

Note

1. Of the 40 patients who completed the DES-T and TAS-20 at *T1*, eight patients (20%) had DES-T scores that exceeded the cut point of 13, whereas six patients (15%) had TAS-20 scores that exceeded the cut point of 60.

Acknowledgments

Disclosure statement

The second and the third author developed the Self-Report Symptom Inventory and authored the manual for this test, which is currently in print for commercial distribution.

ORCID

Daniël van Helvoort iD http://orcid.org/0000-0002-1354-2402

References

Ardolf, B. R., Denney, R. L., & Houston, C. M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist*, 21(6), 899–916. doi:10.1080/13825580600966391

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, 38(1), 23–32. doi:10.1016/0022-3999(94)90005-1

Bush, S., Ruff, R., Troster, A., Barth, J., Koffler, S., Pliskin, N., … Silver, C. (2005). Symptom validity assessment: Practice issues and medical necessity NAN Policy & Planning Committee. *Archives of Clinical Neuropsychology*, 20, 419–426. doi:10.1016/j.acn.2005.02.002

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(6), 723–740. doi:10.1080/13854046.2015.1099738

Cima, M., van Bergen, S., & Kremer, K. (2008). Development of the supernormality scale-revised and its relationship with psychopathy. *Journal of Forensic Sciences*, 53(4), 975–981. doi:10.1111/j.1556-4029.2008.00740.x

Crighton, A. H., Wygant, D. B., Applegate, K. C., Umlauf, R. L., & Granacher, R. P. (2014). Can brief measures effectively screen for pain and somatic malingering? Examination of the Modified Somatic Perception Questionnaire and Pain Disability Index. *The Spine Journal*, 14(9), 2042–2050. doi:10.1016/j.spinee.2014.04.012

Dandachi-FitzGerald, B., Merckelbach, H., & Ponds, R. W. H. M. (2017). Neuropsychologists' ability to predict distorted symptom presentation. *Journal of Clinical and Experimental Neuropsychology*, 39(3), 257–264. doi:10.1080/13803395.2016.1223278

Dandachi-FitzGerald, B., Ponds, R. W. H. M., & 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(8), 771–783. doi:10.1093/arclin/act073

Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J., & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist*, 25(5), 812–828. doi:10.1080/13854046.2011.583280

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. doi:10.3758/bf03193146

Geurten, M., Meulemans, T., & Seron, X. (2018). Detecting over-reporting of symptoms: The French version of the Self-Report Symptom Inventory (SRSI). *The Clinical Neuropsychologist*, 32, 164–181. doi:10.1080/13854046.2018.1524027

Giger, P., & Merten, T. (in press). Equivalence of the German and the French versions of the Self-Report Symptom Inventory. *Swiss Journal of Psychology*.

Green, P. (2003). *Green's word memory test for windows: User's manual*. Seattle, WA: Green's Publishing Inc.

Hartman, D. E. (2002). The unexamined lie is a lie worth fibbing: Neuropsychological malingering and the word memory test. *Archives of Clinical Neuropsychology*, *17*(7), 709–714. doi:10.1093/arclin/17.7.709

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference Participants 1. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, *23*, 1093–1129. doi:10.1080/13854040903155063[Mismatch]

Institute of Medicine. (2015). *Psychological testing in the service of disability determination*. Washington, DC: The National Academies Press. doi:10.17226/21704

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*(6), 741–776. doi:10.1080/13854046.2015.1087597

Merckelbach, H., Boskovic, I., Pesy, D., Dalsklev, M., & Lynn, S. J. (2017). Symptom over-reporting and dissociative experiences: A qualitative review. *Consciousness and Cognition*, *49*, 132–144. doi:10.1016/j.concog.2017.01.007

Merckelbach, H., Jelicic, M., & Pieters, M. (2011). The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology*, *33*(1), 131–139. doi:10.1080/13803395.2010.495055

Merckelbach, H., Merten, T., Dandachi-FitzGerald, B., & Boskovic, I. (2018). De Self-Report Symptom Inventory (SRSI): Een instrument voor klachtenoverdrijving [The Self-Report Symptom Inventory (SRSI): An instrument to measure symptom overreporting]. *De Psycholoog*, *53*(3), 32–40.

Merten, T., Giger, P., Merckelbach, H., & Stevens, A. (in press). *Self-Report Symptom Inventory (SRSI). - deutsche Version (SRSI). Manual* [Professional 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*(2), 102–111. doi: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*(8), 1094–1102. doi:10.1076/jcen.24.8.1094.8379

Morel, K. R., & Marshman, K. C. (2008). Critiquing symptom validity tests for posttraumatic stress disorder: A modification of Hartman's criteria. *Journal of Anxiety Disorders*, *22*(8), 1542–1550. doi:10.1016/j.janxdis.2008.03.008

Niesten, I. J., Merckelbach, H., Van Impelen, A., Jelicic, M., Manderson, A., & Cheng, M. (2017). A lab model for symptom exaggeration: What do we need? *Journal of Experimental Psychopathology*, *8*, 55–75. doi:10.5127/jep.051815

Parks, A. C., Gfeller, J., Emmert, N., & Lammert, H. (2016). Detecting feigned postconcussional and posttraumatic stress symptoms with the Structured Inventory of Malingered Symptomatology (SIMS). *Applied Neuropsychology: Adult*, *24*(5), 429–438. doi:10.1080/23279095.2016.1189426

Resnick, P. J., & Harris, M. R. (2002). Retrospective assessment of malingering in insanity defense cases. In: R. I. Simon & D. W. Shuman (Eds.), *Retrospective assessment of mental states in litigation: Predicting the past* (pp. 101–134). Washington, DC: American Psychiatric Publishing.

Rosen, G. M., & Phillips, W. R. (2004). A cautionary lesson from simulated patients. *Journal of the American Academy of Psychiatry and the Law Online*, *32*, 132–133.

Ruiz, M. A., Drake, E. B., Glass, A., Marcotte, D., & van Gorp, W. G. (2002). Trying to beat the system: Misuse of the Internet to assist in avoiding the deception of psychological symptom dissimulation. *Psychological Psychology: Research and Practice*, *33*, 294–299. doi:10.1037/0735-7028.33.3.294

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, *13*(4), 545–561. doi:10.1076/1385-4046(199911)13:04;1-y;ft545

Stevens, A., Schmidt, A., & Hautzinger, M. (2018). Major depression – A study on the validity of clinician's diagnoses in medicolegal assessment. *The Journal of Forensic Psychiatry & Psychology*, 29, 794–809. doi:10.1080/14789949.2018.1477974

Sweet, J. J., & Guidotti Breting, L. M. (2013). Symptom validity test research: Status and clinical implications. *Journal of Experimental Psychopathology*, 4, 1–14. doi:10.5127/jep.022311

van Egmond, J. J., & Kummeling, I. (2002). A blind spot for secondary gain affecting treatment outcome. *European Psychiatry*, 17(1), 46–54. doi:10.1016/S0924-9338(02)00622-3

van Egmond, J., Kummeling, I., & Balkom, T. (2005). Secondary gain as hidden motive for getting psychiatric treatment. *European Psychiatry*, 20, 416–421. doi:10.1016/j.eurpsy.2004.11.012

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(8), 1336–1365. doi:10.1080/13854046.2014.984763

Viglione, D. J., Giromini, L., & Landis, P. (2016). The development of the Inventory of Problems-29: A brief self-administered measure for discriminating bona fide from feigned psychiatric and cognitive complaints. *Journal of Personality Assessment*, 99(5), 534–544. doi:10.1080/00223891.2016.1233882

Waller, N. G., Putnam, F. W., & Carlson, E. B. (1996). Types of dissociation and dissociative types: A taxometric analysis of dissociative experiences. *Psychological Methods*, 1(3), 300–321. doi:10.1037/1082-989X.1.3.300

Wetter, M. W., & Corrigan, S. K. (1995). Providing information to clients about psychological tests: A survey of attorneys' and law students' attitudes. *Professional Psychology: Research and Practice*, 26(5), 474–477. doi:10.1037/0735-7028.26.5.474

Widows, M. R., & Smith, G. P. (2005). *SIMS – Structured inventory of malingered symptomatology. Professional manual*. Lutz, FL: Psychological Assessment Resources.