Psychology & Neuroscience

Detecting Symptom Exaggeration in Compensation-Seeking Individuals, Psychotherapy Clients, and Individuals Referred for Job Assessments: Psychometric Features of the French and Dutch Versions of the Self-Report Symptom Inventory

Brechje Dandachi-FitzGerald, Louis De Page, and Harald Merckelbach

Online First Publication, June 29, 2023. https://dx.doi.org/10.1037/pne0000318

Psychology & Neuroscience

ISSN: 1983-3288
https://doi.org/10.1037/pne0000318

Detecting Symptom Exaggeration in Compensation-Seeking Individuals, Psychotherapy Clients, and Individuals Referred for Job Assessments: Psychometric Features of the French and Dutch Versions of the Self-Report Symptom Inventory

Brechje Dandachi-FitzGerald[1, 2], Louis De Page[3, 4], and Harald Merckelbach[1]

[1] Department of Clinical Psychological Science, Faculty of Psychology and Neuroscience, Maastricht University

[2] Department of Clinical Psychology, Faculty of Psychology, Open University

[3] Mediter Psychotherapy Center, Halle, Belgium

[4] Centre Hospitalier Jean Titeca (CHJT), Brussels, Belgium

Objective: The Self-Report Symptom Inventory (SRSI) is a relatively new instrument to detect symptom exaggeration. It contains a mix of plausible and pseudosymptoms, the rationale being that people who intend to exaggerate symptoms will overendorse both types of symptoms, whereas individuals responding truthfully will selectively endorse primarily plausible symptoms. The present study examined whether there are any differences in Dutch and French versions of the SRSI as a first step in determining their psychometric equivalence. ***Method:*** Relying on a differential prevalence design, we compared the Dutch and French SRSI in a mixed sample of compensation-seeking individuals (n = 226), psychotherapy clients (n = 95), and job selection candidates (n = 130). Participants were tested at the same bilingual facility and either had a Dutch language background (n = 263) or a French language background (n = 188). ***Results:*** Internal reliability estimates of SRSI subscales were highly comparable across the two language groups. Both language groups exhibited a pattern in which compensation-seeking individuals reported the highest level of pseudosymptoms, psychotherapy clients an intermediate level, and the job selection subgroup the lowest level. There was no difference in this regard between the two language backgrounds. ***Conclusions:*** Given the increasing diversity in the cultural backgrounds of clients, it is important for experts to have different language versions of the same instrument in their toolbox that possess a comparable quality. Our data show that the Dutch and French SRSI possess similar psychometric and conceptual features.

Brechje Dandachi-FitzGerald https://orcid.org/0000-0002-8984-8192

Harald Merckelbach https://orcid.org/0000-0002-5116-7826

There was no third party funding for the article. The authors have no conflicts of interest to declare. Harald Merckelbach is coauthor of the German-language Self-Report Symptom Inventory; Brechje Dandachi-FitzGerald and Harald Merckelbach are coauthors of its Dutch-language version. The test is commercially distributed by Hogrefe Publishers.

The data associated with this article are available via the Open Science Framework and can be accessed at https://osf.io/kv2x8/.

Correspondence concerning this article should be addressed to Brechje Dandachi-FitzGerald, Department of Clinical Psychological Science, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 616, 6200 MD, Maastricht, The Netherlands. Email: b.fitzgerald@maastrichtuniversity.nl

Public Significance Statement
To avoid wrong diagnoses and treatments, it is important to first establish the validity of self-reported symptoms, mainly by excluding tainted presentations due to, for example, symptom exaggeration. The Self-Report Symptom Inventory (SRSI) is a test that helps clinicians to identify patients who engage in symptom exaggeration. The SRSI has been translated in many languages and we found that its Dutch and French versions did not differ significantly in terms of internal reliability and scores across different referral groups, which is a first indication that both versions are equivalent. Establishing equivalence of SRSI language versions is important considering the highly diverse backgrounds of patients.

Keywords: symptom exaggeration, Self-Report Symptom Inventory, Dutch French equivalence, differential prevalence

Supplemental materials: https://doi.org/10.1037/pne0000318.supp

Sometimes, clinicians suspect that clients exaggerate their symptoms and this might be particularly true in a forensic context (e.g., compensation-seeking individuals; Bianchini et al., 2006). To screen for symptom exaggeration, researchers developed several types of dedicated instruments, often referred to as symptom validity tests. Widely used examples are the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997) and the Miller-Forensic Assessment of Symptoms Test (M-FAST; Miller, 2001). Typically, such instruments list unlikely symptoms and rest on the assumption that people who exaggerate their symptoms will overendorse these items. Although tools such as the SIMS and the M-FAST have their merits as part of the evaluation process to identify feigning (see, for reviews, Detullio et al., 2019; Shura et al., 2022; Van Impelen et al., 2014), they suffer from two obvious limitations. First, the majority of their items allude to extreme forms of psychopathology (e.g., psychosis, amnesia) and cover more prevalent symptom domains (e.g., anxiety and depression) less well. Second, in these screening tools, rare, bizarre, or nonexistent symptoms are overrepresented, making them easily recognizable as bogus items for well-informed or coached clients, thereby potentially reducing sensitivity in detecting symptom exaggeration.

With these limitations in mind, Merten et al. (2016, 2022) developed the 107-item Self-Report Symptom Inventory (SRSI). Basically, the SRSI consists of a mix of pseudosymptoms and genuine symptoms, which makes its measurement intention less obvious. Furthermore, both types of items cover a broad spectrum of symptoms (e.g., anxiety, depression, functional neurological symptoms). Thus, the SRSI is a potentially useful screening tool for forensic assessments in which more restricted forms of mental impairment rather than extreme manifestations of psychopathology are to be evaluated, as is the case in, for example, litigation procedures where compensatory damages are sought.

The psychometric qualities of the German SRSI have been well researched (Merten et al., 2019; see also Giromini et al., 2022). Merten et al. (2016) reported for their mixed sample of forensic patients and normal population participants ($N = 367$) internal reliabilities that were satisfactory, with Cronbach αs > .90 for both the pseudosymptoms and genuine symptoms scales. The test–retest stability over a period of 2 weeks in a subsample ($n = 30$) was $r = 0.81$ and $r = 0.91$ for pseudosymptoms and genuine symptoms, respectively. Furthermore, the authors found a correlation of $r = .82$ between SIMS scores and the number of endorsed SRSI pseudosymptoms, which supports the convergent validity of the SRSI. More recent research, summarized in Merten et al. (2022), replicates the substantial correlation between the number of endorsed SRSI pseudosymptoms and scores on concurrent symptom validity tests or indices (e.g., SIMS, the *F-r* index of Minnesota Multiphasic Personality Inventory–2–Restructured Form; Ben-Porath & Tellegen, 2008), with *r*s typically being in the .70 −.85 range.

Based on an aggregated data set, Merten et al. (2016, 2022) recommended two different cutoff values to detect symptom exaggeration, depending on the circumstances. For screening purposes,

they recommended a cutoff of >six pseudosymptoms (sensitivity = .83; specificity = .91; likelihood ratio [LR] = 9.31). For psychodiagnostic standard purposes, they proposed to use the higher threshold of >nine pseudosymptoms (sensitivity = .62; specificity = .96; LR = 13.73).

Making psychological tests available in multiple languages is important to ensure that psychologists in various countries have access to accurate tools for their psychological assessment. Moreover, as the numbers of expats and immigrants continue to rise, forensic and clinical experts are increasingly faced with clients who have different cultural and language backgrounds (e.g., Franzen et al., 2022; Merten et al., in press; Weiss & Rosenfeld, 2012, 2017). To accurately assess these clients, it is crucial to use carefully translated measures, including symptom validity tests such as the SRSI. The SRSI has been translated in 10 different languages, among which Dutch (e.g., Van Helvoort et al., 2019) and French (e.g., Geurten et al., 2018), with the translated versions of the German SRSI being the result of fine-tuning and adjustments through systematic comparison between original and back-translated versions.

Previous research that focused exclusively on either the Dutch or French SRSI has yielded promising results. For example, Van Helvoort et al. (2019) administered the Dutch SRSI to forensic patients and students instructed to feign symptoms. These researchers concluded that the pseudosymptoms scale of the SRSI is insensitive to actual psychopathology. They also concluded that with the cutoff of >nine pseudosymptoms, this scale discriminates fairly good between honest and feigning participants (area under the curve [AUC] = .98). Geurten et al. (2018) conducted factor analyses on the French SRSI in healthy adults and found a two-factor solution corresponding to genuine and pseudosymptoms. They found the genuine and pseudosymptoms scales to possess good internal reliabilities (Cronbach αs > 0.85). These researchers tested discriminatory power by administering the SRSI to several groups of healthy controls and patients, as well as healthy participants instructed to feign cognitive impairments. The SRSI pseudosymptom scale differentiated reasonably well between controls and instructed feigners, with AUCs being in the range of 0.81–0.93.

Although studies focusing on translated versions are a first step in establishing the equivalence of original and translated instruments (e.g., Weiss & Rosenfeld, 2012), they are not sufficient. A next and necessary step would be to collect normative data across groups with different language backgrounds and systematically compare these data. Two instructed feigning studies were conducted with bilingual participants to directly compare the test performance of different language versions of the SRSI. Thus, Giger and Merten (2019) showed that the original German and French versions of the SRSI produced a highly comparable pattern in Swiss bilingual participants. Similarly, in their study with bilingual participants, Dandachi-FitzGerald et al. (in press) used a Bayesian inference approach and found evidence in favor of equivalence between the Dutch and German versions of the SRSI.

In the present study, we compared the Dutch and French versions in a mixed sample of Belgian participants who either spoke Dutch or French and who were referred to the same bilingual assessment facility either to undergo psychodiagnostic evaluation in the context of compensatory damages procedures, psychotherapeutic treatment, or job recruitment. We expected the two language versions to possess similar internal reliabilities for the pseudosymptoms and genuine symptom scales. Most importantly, we expected the two language groups to exhibit a similar pattern of differential prevalence of symptom exaggeration. The differential prevalence approach (e.g., Nijdam-Jones & Rosenfeld, 2017) is based on the notion that symptom exaggeration will be more prevalent in forensic groups (e.g., litigating participants) than in nonforensic participants (e.g., psychotherapy clients), whereas in other samples (e.g., job evaluees), it will be nearly absent because they have an incentive for symptom denial rather than symptom exaggeration. Finding such a differential prevalence pattern would support the rationale on which the SRSI rests and finding a similar pattern across two language versions of the test would support the conceptual equivalence of its translated versions. Practically, it would mean that the corpus of psychometric data obtained with one language version of the SRSI bears also relevance to the other language version.

Method

Participants

Participants (N = 477) were recruited at Mediter Center, a facility specialized in (neuro)-psychological assessment and outpatient psychotherapy, located in Halle, Belgium. Given the bilingual nature of Belgian society, both Dutch- and French-speaking individuals are referred to the center. They are referred for various reasons, which mainly relate to assessments of neuropsychological impairments in the context of insurance claims (i.e., compensation-seeking group), assessments prior to psychotherapeutic treatment (i.e., psychotherapy group), or assessments in the context of job selection and recruitment (i.e., job selection group). In the present study, a consecutive series of 477 participants were tested with the SRSI (see below) as part of the psychodiagnostic assessment. Depending on the precise referral question, this assessment consisted of various other psychometric tools (e.g., clinical scales, the Minnesota Multiphasic Personality Inventory–2–Restructured Form). The data obtained with these other tools will not be considered here. Records with more than five missing SRSI items were excluded from the analysis (n = 20, all in the compensation-seeking group). Six participants manifested a lack of cooperation on the SRSI control items (see below): five in the compensation-seeking group and one in the job selection group. These participants were also excluded from the analysis. There were no obvious cases of random/careless responding on the SRSI control items, leaving 451 participants in the final sample. The study was approved by a standing ethical committee of the Vrije Universiteit Brussel (VUB-Ref 2020-527). Table 1 summarizes demographic background data of the sample broken down by language and referral question.

The men–women ratio was not similar across groups, $\chi^2(2) = 57.21$, $p < .001$, $\eta = .32$. The job selection group had a higher proportion of men than women compared with either the compensation-seeking, $\chi^2(1) = 52.81$, $p < .001$, $\eta = .39$, or the psychotherapy group, $\chi^2(1) = 39.39$, $p < .001$, $\eta = .42$. Men–women distribution was, however, similar in the compensation-seeking and psychotherapy groups, $\chi^2(1) = .003$, $p = .95$, $\eta < .01$. Also, a one-way analysis of variance (ANOVA) revealed a statistically significant difference in age across groups, $F(2, 447) = 175.06$, $p < .001$, $\eta^2 = .44$. Tukey's honestly significant difference test for multiple comparisons indicated that as group, compensation-seeking individuals were significantly older than psychotherapy clients, $p < .001$, 95% CI [4.19, 10.41] who, in turn, were older than job selection candidates, $p < .01$, 95% CI [11.12, 18.00].

In the compensation-seeking group, the most frequent referral questions had to do with work-related mental health concerns (e.g., burnout), psychosomatic complaints, compensation for mild

Table 1
Background Information of Sample Broken Down By Language and Referral Group

	Dutch (n = 263)		French (n = 188)		Full sample (N = 451)	
Background characteristic[a]	*n/M*	*%/SD*	*n/M*	*%/SD*	*n/M*	*%/SD*
Compensation seeking	145	55.1	81	43.1	226	50.1
Gender						
Women	72	49.7	43	53.1	115	50.9
Men	73	50.3	38	46.9	111	49.1
Age	46.26	11.15	45.26	10.59	45.78	10.96
Psychotherapy	55	20.9	40	21.3	95	21.1
Gender						
Women	28	50.9	20	50.0	48	50.5
Men	27	49.1	20	50.0	47	49.5
Age	23.48	9.07	24.33	8.16	23.92	8.59
Job application	63	23.9	67	35.6	130	28.8
Gender						
Women	8	12.7	8	11.9	16	12.3
Men	55	87.3	59	88.1	114	87.7
Age	39.05	14.18	36.48[b]	13.92[b]	37.98[b]	14.11[b]

[a] For referral group and gender, the n and % are given; for age, the M and SD are given. [b] Based on n = 187 and n = 450, respectively, because one participant did not indicate age.

head injury due to an accident, or the presence of trauma-related symptoms due to victimization. The most frequent referral reasons for the psychotherapy group were depression, problematic personality features, occupational stress, psychotic symptoms, anxiety disorders, and marital problems. The job selection group mainly consisted of individuals who had applied for high skills jobs (e.g., pilots).

Self-Report Symptom Inventory

The SRSI (Merten et al., 2016, 2019) contains 107 items with a true-or-false response format. Items describe symptoms- and health-related issues. The SRSI was constructed with a focus on item formulation that utilizes a simple syntactic and semantic structure, avoiding complexities such as double negatives, conditional clauses, ambiguities, and similar issues. This ensures that individuals with a reading proficiency equivalent to a secondary school diploma level and without intellectual disabilities can easily understand and answer the items (Merten et al., 2019). It takes approximately 15 min to complete the instrument. The SRSI includes two main scales. One taps into plausible complaints (i.e., genuine symptoms; e.g., items of the following type "Feeling no interest in things").[1] The other main scale lists unlikely symptoms (i.e., pseudosymptoms; e.g., items of the following type "On a scale from 0 [*no headache*] to 10 [*maximum headache*], my headaches are at "10" almost all the time"). The two main scales each include five subscales, with every subscale consisting of 10 items that cover either genuine or pseudosymptoms in specific domains. The subscales of the genuine symptoms main scale are: cognitive problems, depression, pain, nonspecific somatic complaints, and posttraumatic stress disorder/anxiety. Those of the pseudosymptoms main scale are: cognitive/memory problems, neurological motoric complaints, neurological sensory complaints, pain, and anxiety/depression. The range of possible subscale scores varies between 0 and 10 for each of the 10 subscales, and between 0 and 50 for the two main scales. The SRSI contains an additional seven items, of which two check cooperativeness and five serve as a check on random/careless responding. In the present study, Flemish participants were administered the Dutch version of the SRSI, whereas French-speaking participants from Bruxelles and Wallonia received the French version.

Data Analysis

We sought to answer two questions. First, do the Dutch and French SRSI versions produce similar internal reliabilities across the three referral groups (i.e., compensation seeking, psychotherapy, job selection)? Accordingly, we computed Cronbach αs and average interitem correlations for main scales across groups and compared language versions with each other. Second, do the Dutch and French versions exhibit a similar pattern of genuine and pseudosymptom scores across the three groups? To address this issue, we conducted a 2 (Language: Dutch vs. French) × 3 (Referral Groups) × 2 (Symptoms: Genuine Symptoms vs. Pseudosymptoms) ANOVA on the SRSI data, with repeated measures on the last factor. The absence of a main effect of language or the absence of an interaction effect with language would provide initial support for the equivalence of Dutch and French versions. The presence of a significant interaction of groups and symptoms could reflect a differential prevalence pattern and in case it would occur, we planned to conduct follow-up *t* tests to determine whether compensation-seeking individuals, indeed, had the highest level of pseudosymptoms, psychotherapy clients an intermediate level, and job selection candidates the lowest level. As a further step, we looked at the proportion of participants in each subgroup who scored above the standard cutpoint of nine pseudosymptoms, expecting that irrespective of the language of administration, this proportion would be the highest in the compensation-seeking group, intermediate in psychotherapy clients, and lowest in the job selection candidates. We evaluated group differences with Fisher's exact probability tests.

Results

The anonymized data file can be found at Open Science Framework at https://osf.io/kv2x8/.

Demographic Background and SRSI Symptom Endorsement

For the full sample, age correlated significantly with the number of endorsed genuine symptoms

[1] To ensure test security, the items presented here are for illustrative purposes only and do not reflect actual items from the SRSI.

($r = .56$, $p < .001$, $r^2 = .32$) and with the number of endorsed pseudosymptoms ($r = .34$, $p < .001$, $r^2 = .12$). For the full sample, men and women differed with regard to genuine symptoms ($M_{men} = 16.17$, $M_{women} = 24.37$, $p < .001$, $\eta^2 = .08$) and pseudosymptoms ($M_{men} = 4.25$, $M_{women} = 6.67$, $p < .001$, $\eta^2 = .03$). Because younger men were overrepresented in the job selection group, these correlations might moderate the outcomes of the $2 \times 3 \times 2$ ANOVA.

Internal Reliability

Table 2 shows internal reliabilities for genuine and pseudosymptom scales broken down by language and referral groups. As can be seen, Cronbach αs were mostly satisfactory (>.70), the exceptions being αs for the pseudosymptom scales in job selection participants, possibly because these αs were suppressed by the excess of no answers in the context of a dichotomous scale (e.g., Lissitz & Green, 1975). Most importantly, Cronbach αs were highly comparable across language groups. Much the same is true for the mean interitem *r*s. For both language versions, average interitem *r*s were in the range of 0.15–0.35, which one would expect when higher order concepts (e.g., "symptom exaggeration") are targeted (e.g., Clark & Watson, 1995).

Differential Prevalence Pattern

Table 3 gives mean scores of the three groups and shows proportions of participants scoring above the cutoff. Scores at the level of the subscales are provided in the Supplemental Material File. A 2 (Language) × 3 (Referral Groups) × 2 (Symptoms: Genuine vs. Pseudosymptoms) ANOVA, with repeated measures on the last factor, yielded main effects of symptoms, $F(1, 445) = 1{,}738.91$, $p < .001$, $\eta^2 = .80$, and referral groups, $F(2, 445) = 225.19$, $p < .001$, $\eta^2 = .50$. As expected, there was a significant interaction between symptoms and referral groups, $F(2, 445) = 334.88$, $p < .001$, $\eta^2 = .601$. Post hoc Bonferonni analysis revealed that symptom scores were significantly higher in the compensation-seeking group than in the psychotherapy group, which scored higher than the job selection group (all *p*s < .001). Importantly, the main effect of language fell short of significance, $F(1, 445) = .01$, $p = .927$, $\eta^2 < .01$. The same held for interaction effects with language. Specifically, there were no significant two-way interactions of language and symptoms, $F(2, 445) = 3.82$, $p = .051$, $\eta^2 < .01$, or of language and referral group, $F(2, 445) = .13$, $p = .88$, $\eta^2 < .01$. Neither did the three-way interaction of language, referral groups, and symptoms reach significance, $F(2, 445) = .71$, $p = .49$, $\eta^2 < .01$. Thus, the language version of the SRSI did not impact the differential pattern of symptom reporting across the three groups.

Fisher's exact tests performed on proportions of participants failing the SRSI (i.e., scoring above the cutpoint of 9 on the pseudosymptoms scale) revealed no significant differences between the Dutch and French SRSI within the separate referral groups (compensation seeking: $p = .47$; psychotherapy: $p = .38$). On the other hand and in accordance with the differential prevalence prediction, the proportion of participants who failed the SRSI was higher among compensation-seeking individuals than psychotherapy clients ($p < .01$) and higher among psychotherapy clients than job selection candidates ($p < .01$). None of the job selection candidates failed the SRSI.

Table 2
Cronbach Alphas (α) *and Mean Interitem r (SD) for Genuine and Pseudosymptom Scales Broken Down By Language and Referral Group*

	Genuine symptoms				Pseudosymptoms			
	Dutch		French		Dutch		French	
Referral group	α	Mean item *r* (*SD*)	α	Mean item *r* (*SD*)	α	Mean item *r* (*SD*)	α	Mean item *r* (*SD*)
Compensation seeking	.92	.19 (.15)	.92	.20 (.16)	.92	.19 (.12)	.90	.15 (.16)
Psychotherapy	.93	.20 (.17)	.85	.10 (.22)	.88	.14 (.20)	.94	.31 (.28)
Job selection	.73	.08 (.21)	.88	.22 (.28)	.70	.13 (.33)	.67	.18 (.33)
Total	.96	.33 (.17)	.96	.34 (.21)	.93	.21 (.11)	.93	.21 (.13)

Table 3
Mean (SD) Scores on the Genuine Symptoms and Pseudosymptoms Scale of the SRSI and % SRSI Fail Broken Down by Language and Referral Group

	Genuine symptoms						Pseudosymptoms						SRSI fail (pseudosymptoms > 9)			
	Dutch			French			Dutch			French			Dutch		French	
Referral group	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	%	*n*	%
Compensation	145	28.11	10.43	81	28.75	10.36	145	8.73	8.11	81	7.33	6.82	54	37.2	26	32.1
Psychotherapy	55	20.20	10.77	40	21.05	8.27	55	4.84	5.48	40	4.50	7.04	10	18.2	4	10.0
Job selection	63	2.81	2.75	67	3.34	4.14	63	0.27	.88	67	0.37	0.95	0	0	0	0

Note. SRSI = Self-Report Symptom Inventory.

Discussion

Given the track record of the SRSI as a psychometrically sound tool for detecting symptom exaggeration, there is a need for more normative data on its different language versions. The present study builds upon prior research by Giger and Merten (2019) and Dandachi-FitzGerald et al. (in press). Giger and Merten found that the French and German SRSI produced highly comparable test scores in Swiss research participants, while Dandachi-FitzGerald and coworkers demonstrated the equivalence of the German and Dutch SRSI in bilingual participants. Our study contributes to this literature by showing that the Dutch and French SRSI have highly comparable properties across three different referral groups. Specifically, we found that the two versions exhibit internal reliabilities that are similar and produce patterns of differential prevalence that are highly alike. Taken together, our data are a first step in demonstrating the psychometric and conceptual equivalence of the Dutch and French SRSI. This implies that within reasonable bounds (see below), clinicians may draw upon normative French SRSI data when they want to interpret a Dutch SRSI protocol and vice versa. Given the ever-increasing diversity in clients' language backgrounds with which clinicians are confronted, this is good news. Our data contribute to a psychodiagnostic toolbox that is geared toward linguistic and cultural diversity in clients.

Our findings align with estimated base rates of symptom overreporting in both forensic and clinical contexts. Estimates of symptom exaggeration ranges between 30% and 50% of individuals involved in litigation or workers' compensation procedures (Larrabee et al., 2009; Sherman et al., 2020), which is consistent with our findings of compensation-seeking participants who failed the SRSI (37.25%–32.1%). Moreover, the SRSI failure rate of 10%–18.2% in psychotherapy clients is consistent with the estimated base rate of 15% for symptom overreporting across all clinical neuropsychological evaluations (Martin & Schroeder, 2020). Using the SIMS, Dandachi-FitzGerald et al. (2016) and Bodenburg et al. (2022) observed similar rates of symptom validity test failure in clinical patients: 14.1% and 17.6%, respectively. These findings indicate that symptom overreporting occurs in clinical (nonforensic) assessments. Note that symptom overreporting does not equal malingering as there might be many reasons why individuals distort their symptom presentation (Dandachi-FitzGerald et al., 2022). Apart from hidden agenda's for obtaining incentives (e.g., van Egmond & Kummeling, 2002), clients might, for example, exaggerate their symptoms out of fear that treatment will not be provided if symptoms are perceived as being to mild (e.g., Tracey et al., 2014).

In line with a priori expectations, job candidates endorsed on average less than one pseudosymptom, and none of them scored above the cutoff. Indeed, in order to be considered eligible for the job, this group had an incentive to appear psychologically healthy. Obviously, in this referral group, there is an increased risk of symptom underreporting and as such it makes sense to include validity instruments aimed at detecting faking good response distortions in the assessment battery (e.g., Brown & Sellbom, 2020). Still, the performance of this subgroup on the SRSI provides discriminant evidence for the construct validity of the SRSI as a measure for

symptom overreporting, and important for the present study, both language versions produced similar SRSI patterns in this specific group.

Several limitations of the present study deserve comment. First and foremost, although we compared clients with two distinct language backgrounds, the majority of them were of European origin. Thus, the psychometric and conceptual equivalence of the Dutch and French SRSI that we observed cannot be taken to mean that SRSI data can be generalized to Dutch- or French-speaking immigrants with a non-European background. Importantly, there is more to cross-cultural validity of psychometric tools than just their psychometric and linguistic equivalence. For example, test-taking attitudes may differ across cultures without this having to do with language per se (Ardila, 2005). Arguably, these test-taking attitudes affect normative data obtained with an instrument such as the SRSI.

Second, and related to inherent limitations of the differential prevalence approach (Nijdam-Jones & Rosenfeld, 2017), our data are silent about the sensitivity and specificity of the Dutch and French SRSI versions and whether they are comparable with regard to these features. Determining such accuracy parameters is essential, but requires an external criterion for defining symptom exaggeration. Thus, our findings are but a first, albeit promising, step in systematically comparing the Dutch and French SRSI. Follow-up research might want to specifically focus on sensitivity and specificity of the Dutch and French SRSI versions using criteria such as a failure on the SIMS and/or the presence of external gain motives as proxies of symptom exaggeration. In passing, we note that not all studies employing the SRSI have invariably found excellent specificity and sensitivity coefficients for this instrument. For example, Boskovic et al. (2019) observed a heightened false-positive level of the SRSI in people reporting aversive high impact experiences, an effect possibly mediated by fantasy proneness. Furthermore, in their study on instructed anxiety and pain feigning, Boskovic et al. (2020) found evidence to suggest that the SRSI is better at detecting those who fabricate anxiety symptoms than those who fabricate pain symptoms.

Third, and related to a more general omission in the SRSI literature, we required clients to complete a relatively lengthy subscale of genuine symptoms (50 items), but did not use the potentially clinically relevant information that this subscale may possess. Clearly, collecting more normative data on the genuine symptoms scale of the SRSI is imperative and can be done by relating these data to results of standard clinical tools (e.g., depressions and anxiety scales) in participants with psychological problems, but without incentive to exaggerate those problems.

Last, it is worth noting that many studies evaluating the SRSI, including our own, have involved authors of the SRSI. Therefore, future studies from independent research groups would be valuable to further validate the utility of the SRSI.

In sum, our study was a first step in establishing the psychometric and conceptual equivalence of the Dutch and French versions of the SRSI, which might serve as a complementary tool to widely used symptom validity tests such as the SIMS and the M-FAST. Our data are encouraging and indicate that it is worthwhile to conduct follow-up research. Yet, the lack of a statistical difference between the two language versions across referral groups is a first indication of language versions equivalence, and additional studies are needed to examine the evidence in favor of psychometric equivalence, for example, by using bilingual participants and employing a Bayesian inference approach (e.g., Dandachi-FitzGerald et al., in press). As well, such future research with different language versions of the SRSI might focus on the diagnostic accuracy parameters of the pseudosymptoms scale and the clinical information provided by the genuine symptoms across assessment settings and patient samples.

References

Ardila, A. (2005). Cultural values underlying psychometric cognitive testing. *Neuropsychology Review*, *15*(4), 185–195. https://doi.org/10.1007/s11065-005-9180-y

Ben-Porath, Y., & Tellegen, A. (2008). *Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF)*. APA PsycTests. https://doi.org/10.1037/t15121-000

Bianchini, K. J., Curtis, K. L., & Greve, K. W. (2006). Compensation and malingering in traumatic brain injury: A dose–response relationship? *The Clinical Neuropsychologist*, *20*(4), 831–847. https://doi.org/10.1080/13854040600875203

Bodenburg, S., Wendiggensen, J., & Kasten, E. (2022). Scores in self-report questionnaires assessing adult

ADHD can be influenced by negative response bias but are unrelated to performance on executive function and attention tests. *Psychological Injury and Law*, *15*(2), 189–199. https://doi.org/10.1007/s12207-022-09448-y

Boskovic, I., Hope, L., Ost, J., Orthey, R., & Merckelbach, H. (2019). Detecting feigned high impact experiences: A symptom over-report questionnaire outperforms the emotional Stroop task. *Journal of Behavior Therapy and Experimental Psychiatry*, *65*, Article 101483. https://doi.org/10.1016/j.jbtep.2019.101483

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*(5), 730–739. https://doi.org/10.1027/1015-5759/a000547

Brown, T. A., & Sellbom, M. (2020). The utility of the MMPI-2-RF validity scales in detecting underreporting. *Journal of Personality Assessment*, *102*(1), 66–74. https://doi.org/10.1080/00223891.2018.1539003

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Dandachi-FitzGerald, B., Merckelbach, H., & Merten, T. (2022). Cry for help as a root cause of poor symptom validity: A critical note. *Applied Neuropsychology. Adult*. Advance online publication. https://doi.org/10.1080/23279095.2022.2040025

Dandachi-FitzGerald, B., Pienkohs, S., Merten, T., & Merckelbach, H. (in press). Detecting symptom overreporting: Equivalence of the Dutch and German Self-Report Symptom Inventory. *Psychological Test Adaptation and Development*.

Dandachi-FitzGerald, B., van Twillert, B., van de Sande, P., van Os, Y., & Ponds, R. W. H. M. (2016). Poor symptom and performance validity in regularly referred Hospital outpatients: Link with standard clinical measures, and role of incentives. *Psychiatry Research*, *239*, 47–53. https://doi.org/10.1016/j.psychres.2016.02.061

Detullio, D., Messer, S. C., Kennedy, T. D., & Millen, D. H. (2019). A meta-analysis of the Miller Forensic Assessment of Symptoms Test (M-FAST). *Psychological Assessment*, *31*(11), 1319–1328. https://doi.org/10.1037/pas0000753

Franzen, S., Watermeyer, T. J., Pomati, S., Papma, J. M., Nielsen, T. R., Narme, P., Mukadam, N., Lozano-Ruiz, Á., Ibanez-Casas, I., Goudsmit, M., Fasfous, A., Daugherty, J. C., Canevelli, M., Calia, C., van den Berg, E., Bekkhus-Wetterberg, P., & the European Consortium on Cross-Cultural Neuropsychology (ECCroN). (2022). Cross-cultural neuropsychological assessment in Europe: Position statement of the European Consortium on Cross-Cultural Neuropsychology (ECCroN). *The Clinical Neuropsychologist*, *36*(3), 546–557. https://doi.org/10.1080/13854046.2021.1981456

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

Giger, P., & Merten, T. (2019). Equivalence of the German and the French versions of the Self-Report Symptom Inventory. *Swiss Journal of Psychology*, *78*(1–2), 5–13. https://doi.org/10.1024/1421-0185/a000218

Giromini, L., Young, G., & Sellbom, M. (2022). Assessing negative response bias using self-report measures: New articles, new issues. *Psychological Injury and Law*, *15*(1), 1–21. https://doi.org/10.1007/s12207-022-09444-2

Larrabee, G. J., Millis, S. R., & Meyers, J. E. (2009). 40 plus or minus 10, a new magical number: Reply to Russell. *The Clinical Neuropsychologist*, *23*(5), 841–849. https://doi.org/10.1080/13854040902796735

Lissitz, R. W., & Green, S. B. (1975). Effect of the number of scale points on reliability: A Monte Carlo approach. *Journal of Applied Psychology*, *60*(1), 10–13. https://doi.org/10.1037/h0076268

Martin, P. K., & Schroeder, R. W. (2020). Base rates of invalid test performance across clinical nonforensic contexts and settings. *Archives of Clinical Neuropsychology*, *35*(6), 717–725. https://doi.org/10.1093/arclin/acaa017

Merten, T., Dandachi-FitzGerald, B., Boskovic, I., Puente-López, E., & Merckelbach, H. (2022). The Self-Report Symptom Inventory. *Psychological Injury and Law*, *15*(1), 94–103. https://doi.org/10.1007/s12207-021-09434-w

Merten, T., Dandachi-FitzGerald, B., Puente-López, E., & Çetin, E. (in press). International Perspectives on Psychological Injury and Law: Cross-Cultural Aspects of Symptom and Performance Validity Assessment. In G. Young, T. Bailey, L. Giromini, L. Erdodi, R. Rogers, & B. Levitt (Eds.), *Handbook of Psychological Injury and Law*. Springer Nature.

Merten, T., Giger, P., Merckelbach, H., & Stevens, A. (2019). *Self-Report Symptom Inventory (SRSI)* [German version of the Self-Report Symptom Inventory]. 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. https://doi.org/10.1007/s12207-016-9257-3

Miller, H. A. (2001). *Miller-Forensic Assessment of Symptoms Test (M-FAST): Professional manual*. Psychological Assessment Resources.

Nijdam-Jones, A., & Rosenfeld, B. (2017). Cross-cultural feigning assessment: A systematic review of feigning instruments used with linguistically, ethnically, and culturally diverse samples. *Psychological Assessment*, *29*(11), 1321–1336. https://doi.org/10.1037/pas0000438

Sherman, E. M. S., Slick, D. J., & Iverson, G. L. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology*, *35*(6), 735–764. https://doi.org/10.1093/arclin/acaa019

Shura, R. D., Ord, A. S., & Worthen, M. D. (2022). Structured inventory of malingered symptomatology: A psychometric review. *Psychological Injury and Law*, *15*(1), 64–78. https://doi.org/10.1007/s12207-021-09432-y

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *The Journal of the American Academy of Psychiatry and the Law*, *25*(2), 183–189. https://doi.org/10.1037/t04573-000

Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist*, *69*(3), 218–229. https://doi.org/10.1037/a0035099

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

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*(6), 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*(8), 1336–1365. https://doi.org/10.1080/13854046.2014.984763

Weiss, R. A., & Rosenfeld, B. (2012). Navigating cross-cultural issues in forensic assessment: Recommendations for practice. *Professional Psychology, Research and Practice*, *43*(3), 234–240. https://doi.org/10.1037/a0025850

Weiss, R. A., & Rosenfeld, B. (2017). Identifying feigning in trauma-exposed African immigrants. *Psychological Assessment*, *29*(7), 881–889. https://doi.org/10.1037/pas0000381

Received November 16, 2022
Revision received May 22, 2023
Accepted May 22, 2023 ■