



# The Spectrum of Response Bias in Trauma Reports: Overreporting, Underreporting, and Mixed Presentation

Irena Boskovic<sup>1</sup> · Luciano Giromini<sup>2</sup> · Aria Katsouri<sup>1</sup> · Elizabet Tsvetanova<sup>1</sup> · Josephine Fonse<sup>1</sup> · Harald Merckelbach<sup>3</sup>

Received: 11 October 2023 / Accepted: 20 March 2024  
© The Author(s) 2024

## Abstract

Typically, research on response bias in symptom reports covers two extreme ends of the spectrum: overreporting and underreporting. Yet, little is known about symptom presentation that includes both types of response bias simultaneously (i.e., mixed presentation). We experimentally checked how overreporting, underreporting, and mixed reporting reflect in trauma symptom reports. Undergraduate students ( $N = 151$ ) were randomly allocated to one of four conditions: control group ( $n = 40$ ), overreporting ( $n = 37$ ), underreporting ( $n = 36$ ), or a mixed reporting ( $n = 38$ ) group. The control group was asked to be honest, whereas the simulation groups received instructions they were expected to be able to relate to. Participants were administered a PTSD checklist, the Supernormality Scale-Revised, which screens for symptom denial, and the Inventory of Problems-29, which gauges overreporting. For all three measures, group differences were significant ( $F(3, 147) > 13.78$ ,  $p_s < .001$ ,  $\eta^2 > .22$ ), with the overreporting group exhibiting the most severe symptom presentation and the underreporting group endorsing the fewest symptoms, while the mixed group almost perfectly imitated the results of the control group on all measures. Thus, instructing individuals to strategically exhibit both overreporting, to a functional degree, and underreporting results in a symptom presentation that is relatively similar to genuine performance. To optimize accuracy statistics, researchers might want to consider including such mixed group condition in future simulation studies.

**Keywords** PTSD · Trauma reports · Response bias · Overreporting · Underreporting · IOP-29 · Supernormality · SS-R

Response bias indicates distortion of symptom presentation, either in a negative (i.e., overreporting) or in a positive direction (i.e., underreporting), and it can be a product of conscious intent or a consequence of personality traits (e.g., fantasy proneness, Merckelbach, 2004). Intentional symptom overreporting can be driven by different types of motives, some internal (e.g., playing a sick role) and some external (e.g., compensation). When such behavior occurs due to the internal motives, it signals possible factitious disorder, whereas when external benefits are the motivators, this behavior is not a reflection of pathology but rather of deceptive behavior—malingering (Rogers & Bender, 2018).

However, oftentimes the type of incentives is unknown, in which case feigning is the preferred term.

Measures serving the purpose of screening for *negative* response bias based on self-reported information are often referred to as symptom validity tests (SVTs; for a review, see Giromini et al., 2022). A recently developed stand-alone SVT is the Inventory of Problems-29 (IOP-29; Viglione et al., 2017), which includes 29 wide-range items capturing both the invalid performance and the invalid symptom endorsement qualities in symptom presentation. The items pertain to different psychiatric and cognitive issues, such as post-traumatic stress disorder (PTSD), depression, schizophrenia, and cognitive impairment. Until now, IOP-29 has been well researched in different contexts and with different types of symptom presentations, and the research outcomes have been invigorating (for a review see Giromini & Viglione, 2022; for a meta-analysis see Puente-López et al., 2023).

The topic of *positive* response bias, also known as impression management or defensiveness (see Rogers, 2018), has been investigated mainly in the context of self-reported

✉ Irena Boskovic  
boskovic@essb.eur.nl

<sup>1</sup> Erasmus University Rotterdam, Rotterdam, the Netherlands

<sup>2</sup> University of Turin, Turin, Italy

<sup>3</sup> Maastricht University, Maastricht, the Netherlands

personality traits during pre-employment evaluations (Griffin & Wilson, 2012; Lavashina, 2018; see also Paulhus, 2012). The results of these studies suggest that the prevalence of positive response bias may be higher than that of negative response bias (Rogers, 2018), although it is less well researched (Faust, 2023). When it comes to the *symptom* presentation and underreporting, the majority of measures are scales embedded in clinical instruments, such as the social desirability (L) scale of MMPI instruments (Baer & Miller, 2002; for meta-analyses see Picard et al., 2023). Looking at the stand-alone measures, the only test developed specifically to allow testing in both general and forensic populations is the Supernormality Scale (SS; Cima et al., 2003). The term *supernormality* conceptualizes systematic denial of common everyday symptoms, regardless of social norms, therefore differing from social desirability (Cima et al., 2003, 2008). The revised version of this scale (SS-R) consists of 50 items, 34 of which relate to mood, dissociation, aggression, and obsessive issues in the broadest sense, while the remaining items are bogus items designed to obscure the true goal of the test (Cima et al., 2008). The psychometric properties of the SS-R have been shown to be adequate in both general and forensic populations (Cima et al., 2008). However, not many studies have been carried out using the SS-R, and further investigation would be beneficial.

An important problem in the study of response bias is that over- and under-reporting of symptoms have largely been treated as behavioral opposites, as dichotomous phenomena (Sherman et al., 2020; Walters et al., 2008). Typically, the instructions that experimental studies employ in these domains reflect this idea of bipolarity. In simulation studies relevant to the forensic domain, instructions to overreport are commonly contrasted with instructions to respond honestly (e.g., Boskovic et al., 2022), and a similar approach is typical for simulation of pre-employment evaluation, where instructions to underreport are contrasted with instructions to respond honestly (e.g., Edens et al., 2001). In real-life, however, it happens that people sometimes engage in both over- and under-reporting of symptoms (e.g., Whitman et al., 2023). More specifically, on some occasions, people strategically tailor their reports and presentations to exaggerate some problems (i.e., overreporting) while concealing some other complaints (i.e., underreporting). This type of behavior, which we refer to as “mixed feigning,” has been well-documented in organizational psychology research focusing on job applicants (Lavashina & Campion, 2007; Melchers et al., 2020) but has not been thoroughly investigated in the domain of symptom presentation.

## Trauma Reports and Response Bias

PTSD is a cluster of symptoms occurring after a traumatic event (see DSM-5, American Psychological Association (APA), 2013; DSM-5-TR; American Psychiatric Association,

2022). The prevalence of PTSD mostly depends on the type of traumatic exposure, so the highest frequency of such diagnosis is often found among victims of sexual assault (up to 80%, Hall & Hall, 2006) and war veterans (up to 58%, Guriel & Fremouw, 2003). In the general population, approximately 15% of individuals exhibit PTSD (Hall & Hall, 2006). However, these prevalence figures should be taken with caution, as the formulation of a PTSD diagnosis may be especially susceptible to distorted representations of symptoms due to high familiarity of the general public with traumatic experiences, and due to wide media coverage of this disorder. This is especially true because diagnoses of PTSD, as majority of psychological complaints listed in DSM-5, are often based largely on subjective, self-reported symptoms (Resnick et al., 2018), which are easily and frequently modified or embellished (Burgess & McMillan, 2001) in both clinical and forensic contexts.

The psychological assessment of PTSD is a commonly used context for the investigation of overreporting (Guriel & Fremouw, 2003). The estimated prevalence of deceptive symptom presentation ranges from 30 up to 50% of trauma reports (Freeman et al., 2008), and research involving students also confirmed the ease with which one could mispresent trauma-related symptomatology (Boskovic et al., 2019a, b). In contrast, positive response bias (i.e., underreporting) in trauma reports is a far less popular topic among researchers. This is confirmed by the lack of results when trying to find literature on the topic of symptom underreporting in PTSD.

While overreporting may lead to an inflated prevalence of trauma and PTSD in general, underreporting of symptoms, on the contrary, is likely to (falsely) lower it. The case in point are victims of sexual assault. As mentioned above, they are the most vulnerable to develop PTSD (Hall & Hall, 2006; Young, 2016), yet, sexual assault is commonly underreported trauma (REINN, 2016), especially among students (Boskovic et al., 2023; Wilson & Miller, 2016), which calls into question the prevalence of PTSD in this group. The forensic context includes many situations in which screening for underreporting could be of importance, such as child custody (see Baer & Miller, 2002) and parole hearings (Ruback & Hopper, 1986). Looking specifically for underreporting of PTSD symptoms may carry even more weight knowing that PTSD is highly comorbid with substance abuse (Brady et al., 2004), which is one of the most underreported problems (Lapham et al., 2001; Magura, 2010). Concealment of such behavior is important to detect, especially among professionals carrying weapons, for instance, police officers or military personal, who were already shown to exhibit underreporting of unfavorable personality traits (Jackson & Harrison, 2018; see also Lavashina, 2018).

As PTSD includes a variety of symptoms, a person does not necessarily have to either overreport or underreport all

of them. One could pick and choose which symptoms to exaggerate and which to underreport so as to find an optimal level of self-presentation that would appear convincing yet functional. For instance, exaggerating intrusive symptoms but underreporting high arousal or irritation might lead to the most supportive reaction of the environment, including the assessor. Further, some symptoms of PTSD mostly refer to physical complaints, which, in certain culture, might also be easier to acknowledge or exaggerate due to lower stigma than when disclosing psychological issues (e.g., Gilmoor et al., 2019). Due to the selective symptom endorsement in such cases, we expect that this more subtle type of response bias (i.e., mixed reporting) might be the most difficult to detect using current SVTs. Specifically, most SVTs are based on the notion that feigners' dominant response style is hyperbolism, i.e., a generalized form of symptom exaggeration (Boskovic, 2022). To our knowledge, no study so far has directly compared the detectability of the three types of response bias in trauma-related accounts. Therefore, this project was undertaken with a specific aim to address this gap.

## Current Study

To evaluate how well different forms of response bias can be detected, we employed a simulation design with four conditions. In the first phase of the study, all participants were screened for PTSD symptoms. Then, by random allocation, students were instructed to either respond honestly (i.e., control group), or they received a vignette depicting a situation in which exhibiting either (1) overreporting, or (2) underreporting, or (3) a mixed strategy (i.e., simultaneous over- and under-reporting) would be beneficial (see Materials). Participants were then assessed for PTSD symptoms (PTSD checklist; PCL-5; Weathers, 2008), overreporting (IOP-29) and underreporting (Supernormality Scale-Revised). Based on the research findings thus far, we expected that the responses of mixed strategy group would be well calibrated, hence, indistinguishable from the honest group, whereas the other two forms of responding (over- and under-reporting) will be detected to a higher degree. Conversely, we anticipated a higher degree of detectability for the other two response types (over- and under-reporting).

## Method

### Participants

An a priori *G\*power* analysis, with alpha set at .05, beta at .80, and opting for a medium-size effect ( $f=.25$ ), indicated a required sample size of 180 participants. Thus, our initial sample consisted of 189 undergraduates. However,

some participants had to be removed from the dataset due to our exclusion criteria: (a) failure to complete all questions ( $n_{\text{excluded}}=7$ ), (b) not giving permission to use data ( $n_{\text{excluded}}=2$ ), (c) failure to provide a detailed elaboration of the task at the end of the study ( $n_{\text{excluded}}=11$ ), (d) failure to pass attention checks within each measure ( $n_{\text{excluded}}=17^1$ ), (e) and being younger than 18 ( $n_{\text{excluded}}=1$ ; see Procedure). As such, the final sample consisted of 151 undergraduates in their 20s ( $M=20$ ,  $SD=2.75$ ), mostly women (81.5%). The most commonly reported nationalities were Dutch (45%), German (10%), and Polish (5%). Participants' self-reported English proficiency on a 5-point Likert scale was overall high ( $M=4.46$ ,  $SD=.67$ ).

As noted above, participants were randomly assigned to one of four groups: control ( $n=40$ ), overreporting ( $n=37$ ), underreporting ( $n=36$ ), and mixed presentation ( $n=38$ ). These groups did not differ in terms of age ( $F(3, 147)=.406$ ,  $p=.749$ ) or English proficiency ( $F(3, 147)=.806$ ,  $p=.492$ ).

## Measures and Materials

### Brief Symptom Inventory (BSI-18; Derogatis, 2001)

The BSI-18 includes 18 items that tap into symptoms of depression, anxiety, and somatization, measuring the individuals' overall level of psychological distress during the last 7 day. The response format is a 5-point scale, ranging from 0 (not at all applicable) to 4 (extremely). The range of the total score, hence, is from 0 to 72, with a higher score indicating the presence of a higher level of distress (Derogatis, 2001). The *Cronbach's alpha* for BSI-18 in this study was .93.

### Severity of Posttraumatic Stress Symptoms-Adult (National Stressful Events Survey PTSD Short Scale (NSESSS); Kilpatrick et al., 2013)

The NSESSS contains nine items and is used to evaluate the severity of individuals' PTSD-related symptoms during the past week based on its description in the DSM-5 (APA, 2013). The response format is a 5-point Likert-like scale, ranging from 0 (not at all) to 4 (extremely). The minimum score to be obtained is 0 and the maximum 36 with higher scores indicating increased severity of PTSD. The NSESSS is a reliable measure with proven high internal consistency and convergent validity in a non-clinical sample (LeBeau et al., 2014). In this study, this measure was employed in the pre-screening phase. The *Cronbach's alpha* of NSESSS in this study was .89.

<sup>1</sup> Some participants failed attention checks on multiple measures and some only on one of the measures.

### PTSD Checklist for DSM-5 (PCL-5; Weathers, 2008)

The PCL-5 consists of 20 items that measure the presence and severity of PTSD criteria according to the DSM-5. Participants do not have to provide any information regarding a traumatic event but just to respond to the list of symptoms and whether (and in which intensity) they were present during the last month. The response format is a 5-point Likert scale, ranging from 0 (not at all) to 4 (extremely). The total score hence ranges from 0 to 80, with higher score indicating higher severity of PTSD symptoms. A score higher than 33 is suggestive of probable PTSD within general population samples (Weathers, 2008; see also [www.ptsd.va.gov](http://www.ptsd.va.gov)). The *Cronbach's alpha* of PCL-5 was .96. In order to secure that participants were paying attention while filling out this questionnaire, we added two items that served as attention checks ("Please select Quite a bit/or Not at all/ if you are reading this"). Because seven participants failed these checks for inattentive responding, their data were removed from the dataset. We acknowledge that PCL-5 is rarely used as a stand-alone assessment measure and is mostly combined with Clinician-Administered PTSD Scale for DSM-5 (CAPS-5; Weathers et al., 2013). Importantly, research has confirmed appropriate psychometric properties of PCL-5 and strong association between the two measures (Boyd et al., 2021; Lee et al., 2022; Roberts et al., 2021).

### Supernormality Scale-Revised (SS-R; Cima et al., 2008)

The SS-R consists of 50 items, and it is employed to evaluate the tendency to underreport symptoms (i.e., supernormality). The items describe common everyday problems and it is expected that participants endorse the majority of them. Supernormality is detected when respondents endorse the "not applicable" option on a large scale and systematically deny everyday problems in an attempt to appear "supernormal." The response format is a 4-point Likert scale, with 1 (not applicable at all) to 4 (extremely applicable); lower scores are indicative of a stronger tendency toward supernormality. The total score is calculated as a sum of responses (with two items having revised coding). The SS-R was shown to have acceptable sensitivity and specificity, with a proposed cutoff score of 60 (Cima et al., 2008). Two checks for inattentive responding were added ("Please select Applicable/ or Not Applicable/ if you are reading this"), and 13 participants did not provide appropriate answers and were excluded from the dataset. The *Cronbach's alpha* of SS-R was .88.

### Inventory of Problems-29 (IOP-29; Viglione et al., 2017)

The IOP-29 is a relatively new measure designed to differentiate between genuine or credible symptom presentations

and symptom overreporting, related to a variety of psychological problems that include trauma-related complaints, cognitive/neuropsychological, psychotic, and depression-related symptoms. It includes 26 self-report items and three cognitive subtasks. For most of the items, the response options are in the form of "true," "false," and "doesn't make sense," the latter being a novelty for SVTs and a unique feature of the IOP-29 (Viglione et al., 2017). In this study, we also included one attention check on the basis of which three participants were excluded ("To this item respond with T"). Also, because the computation of the IOP-29 feigning score requires participants to be 18 or older, one participant was excluded due to their young age.

The scoring of the 29 IOP-29 items generates the False Disorder Probability Score (FDS), which in a recent quantitative literature review inspecting IOP-29 results from 3777 protocols yielded an average sensitivity of .86 and an average specificity of .92 when using the standard cutoff score of  $FDS \geq .50$  (Giromini & Viglione, 2022). However, it should be noted that most of the studies included in this review article used a simulation design, and almost half of these simulation studies included non-clinical controls rather than clinical controls. Therefore, these results might overestimate the true accuracy of the IOP-29. Still, Holcomb et al. (2023), studying a sample of 150 clinically referred individuals for neuropsychological assessment, found that the IOP-29 outperformed the Negative Impression Management (NIM) scale of the Personality Assessment Inventory (Morey, 1991, Morey et al., 2007) in predicting performance (in)validity ( $r_{IOP-29} = .34$  Vs.  $r_{PAI\ NIM} = .06$ ;  $z = 2.50$ ,  $p < .01$ ). Also, the results of a recently published criterion-groups study inspecting a dataset of 174 court-ordered psychological evaluations using the SIMS and MMPI-2-RF as criterion variables supported the effectiveness of the IOP-29. That study found Cohen's  $d$  effect size values ranging from 1.70 to 2.67, depending on the criterion (Roma et al., 2023). Accordingly, when designing our research project, we considered the IOP-29 to be an adequate measure to include in our study.

### Vignettes

Four different instructions were randomly administered to participants. One involved responding honestly (i.e., the control group), while the other three were created for the simulation groups: overreporting, underreporting, and mixed strategy (see Appendix). As our participants were psychology students, the instructions were created having in mind the type of experiences they could relate to. The vignette content was driven by prior research on response bias in student populations. Specifically, based on studies on symptom overendorsement among students (Boskovic, 2020), we decided that for the overreporting group

depicting an exam context would be highly motivational. For the underreporting group, we followed the work of Lavashina (2018), in which this type of responding is primarily connected to job applicants. Hence, we described a job-seeking situation. In order to elicit a mixed strategy, we stayed with the job-seeking context but included additional background information that would encourage students to overreport trauma-related complaints, at least to some degree.

The *overreporting* group received a vignette in which a protagonist is a student very likely to fail the exams unless they are given an extension. The extension might be achieved by pretending to experience trauma-related symptoms, which are exceptionally distressful to the protagonist. It should be noted that these instructions informed participants that they had the possibility of receiving an extension if they fabricated or exaggerated trauma-related problems, but they did not explicitly encourage them to engage in this behavior.

The *underreporting* group was asked to imagine being a protagonist who is fresh out of the university and wants to get employed in an “old school” trauma institution that has strict expectations from their employees, such as being professional first and human second. In this case, participants were explicitly invited to “find an optimal way to present [themselves] in order to impress them and score this job”.

The *mixed strategy* group received similar instructions as the underreporting group, except that the institution they are job hunting at is a modern “new age” trauma center, where having some personal experience and understanding of trauma is considered more important than being a professional. More specifically, they were told that in this institution, “empathy and personal difficult experiences are highly cherished”. As in the underreporting group, the mixed feigning group also was explicitly encouraged to find an optimal way to present themselves in order to be hired.

In all three vignettes, the end part referred to the assessment they needed to go through and requested participants to imagine that the tests given in this study were part of that official procedure. Participants were explicitly warned not to overdo their presentations as they would be caught lying. They were also told that the most convincing presentation would be rewarded with a €10 voucher.

## Procedure

This study was conducted online, using Qualtrics. The link for the study was provided on the university research participation platform from where they could sign up for the study, after which they received the Qualtrics link. After the information about the study and the informed consent,

participants were presented with demographic questions about their age, gender, education, and English proficiency. Then, participants received pre-screening measures (BSI-18 and NSESSS), enabling us to check for potential differences in mental health between the groups. After filling out the measures, participants were randomly (pre-set by Qualtrics configuration) assigned to the control group, the overreporting group (academic extension), the underreporting group (old school trauma institution), and the mixed strategy group (new age trauma center). The first group was just given the instructions to respond honestly, whereas the three remaining groups received vignettes that included information about the context they needed to imagine being in. Following the instructions, they were asked to complete the PCL-5, the SS-R, and the IOP-29. The presentation order of SS-R and IOP-29 was randomized. After completing these scales, participants were told to respond honestly to exit-questions, which were about their motivation, the clarity of instructions, and the difficulty of the task. Finally, participants were debriefed and rewarded student credits (0.5 credits for 30 min) for their participation, and randomly chosen participants received additional monetary award. The study was approved by the standing Ethical committee of Erasmus University Rotterdam, the Netherlands.

## Data Analyses

Potential differences between the conditions were inspected using Analyses of Variance (ANOVA), with Bonferroni post hoc tests. The data was also analyzed using the non-parametric alternatives, but, as the outcomes did not differ from the outcomes of ANOVA, we opted to present the parametric results. The data and the outputs of our analyses are uploaded on the Open Science Framework (OSF) platform (anonymous view link: [https://osf.io/huw67/?view\\_only=b73b5cea14b645ee8a2c58fc1ad80f5b](https://osf.io/huw67/?view_only=b73b5cea14b645ee8a2c58fc1ad80f5b)).

## Results

### Motivation, Clarity of Instructions, and Difficulty of the Task

The four groups reported moderate motivation ( $M = 3.34$ ,  $SD = .80$ ), with no differences between groups in this aspect,  $F(3, 147) = .89$ ,  $p = .449$ . They did, however, provide significantly different scoring on the clarity of instructions,  $F(3, 147) = 3.17$ ,  $p = .026$ , and on the difficulty of the task,  $F(3, 147) = 2.97$ ,  $p = .034$ . More specifically, post hoc Bonferroni tests indicated that significant differences in clarity and difficulty were evident when comparing the control and overreporting group ( $p = .038$ ,  $\eta^2 = .057$  and  $p = .036$ ,  $\eta^2 = .061$ , respectively) (For descriptives, see Table 1).

**Table 1** Mean scores of groups on motivation, clarity of instructions, and difficulty of the task

Group	Control <sup>1</sup>	Overreporting <sup>2</sup>	Underreporting <sup>3</sup>	Mixed strategy <sup>4</sup>
<i>N</i>	40	37	36	38
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
<b>Motivation</b>	3.45 (.85)	3.43 (.83)	3.22 (.80)	3.24 (.71)
<b>Clarity</b>	4.20 (.76)	3.68 (.91)	3.72 (.88)	3.89 (.76)
<b>Difficulty</b>	1.53 (.71)	2.00 (.81)	1.75 (.77)	1.61 (.68)

The responses were given on a 5-point Likert scale (1–5) with higher values indicating higher levels. Effect sizes of contrasts (*Hedge's g*): (a) motivation 1–2=.02, 1–3=.27, 1–4=.26, 2–3=.26, 2–4=.25, 3–4=.02; (b) clarity 1–2=.62, 1–3=.58, 1–4=.41, 2–3=.04, 3–4=.25; (c) difficulty 1–2=.62, 1–3=.30, 1–4=.11, 2–3=.31, 2–4=.52, 3–4=.19

## Pre-screening Measures: Distress and PTSD Symptoms

To control for potential group differences in terms of a priori psychopathology, we examined students' general distress levels (BSI-18) and PTSD-related symptoms (NSESSS). Overall distress was moderate ( $M = 15.87$ ,  $SD = 12.35$ ; range 0–68), and the level of PTSD symptoms was on the lower end ( $M = 8.18$ ,  $SD = 7.29$ ; range 0–33). The four groups did not significantly differ on these measures,  $F(3, 147) = .36$ ,  $p = .782$  and  $F(3, 147) = .501$ ,  $p = .682$ , respectively (for details, see Supplemental Table 1).

## Group Differences in Reported PTSD Symptoms (PCL-5)

Looking at the number of endorsed items on the PCL-5, the overall effect of group was found to be significant,  $F(3, 147) = 34.71$ ,  $p < .001$ ,  $\eta_p^2 = .41$ . More specifically, post hoc Bonferroni checks revealed statistically significant differences between the control and the overreporting groups ( $p < .001$ , *Cohen's d* = 1.42), whereas the difference between the control and underreporting groups attained exactly the  $p = 0.050$  level with an associated *Cohen's d* of .75. No statistically significant difference was found between the control and mixed strategy groups ( $p = 1.00$ , *Cohen's d* = .30). The overreporting group endorsed significantly more items than

the underreporting ( $p < 0.001$ , *Cohen's d* = 2.27) and mixed strategy groups ( $p < .001$ , *Cohen's d* = 1.20), and the underreporting group had a significantly lower scores than the mixed strategy group ( $p < .001$ , *Cohen's d* = 1.18; for details see Table 2). For results regarding groups' scores on the separate symptom domains, please see Supplemental Table 2.

Using the cutoff score of > 33, 20% of the control group ( $n = 8$ ) obtained scores indicative of risk for PTSD, against 76% of the overreporting condition ( $n = 28$ ), 2.7% of the underreporting group ( $n = 1$ ), and 26.3% of the mixed strategy participants ( $n = 10$ ).

## Group Differences in Supernormality (SS-R)

Significant group differences were also evident for the SS-R,  $F(3, 147) = 13.78$ ,  $p < .001$ ,  $\eta_p^2 = .22$ . More specifically, post hoc tests showed that the control group exhibited significantly more supernormality than the overreporting group ( $p = .004$ , *Cohen's d* = .85) but less than the underreporting group ( $p = .014$ , *Cohen's d* = .67). The control group and the mixed strategy group attained comparable supernormality levels ( $p = 1.00$ , *Cohen's d* = .12). Further, the overreporters presented significantly lower levels of supernormality than the underreporting ( $p < .001$ , *Cohen's d* = 1.43) and mixed strategy groups ( $p = .025$ , *Cohen's d* = .71). Finally, the underreporting and mixed strategy groups also differed significantly from each other ( $p = .003$ ,

**Table 2** Mean scores of groups and scores beyond the cutoff points on PTSD Checklist (PCL-5), Supernormality Scale (SS-R), and Inventory of Problems-29 (IOP-29)

	Control	Overreporting	Underreporting	Mixed strategy
<i>N</i>	40	37	36	38
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
PCL-5 total	19.07 (14.84)	43.43 (19.24)	10.17 (7.39)	23.37 (13.76)
% Cutoff: > 33	20%	76%	2.7%	26.3%
SS-R total*	78.47 (13.31)	89.67 (12.99)	68.39 (16.69)	80.16 (13.61)
% Cutoff < 60	12.5%	2.7%	39%	5.3%
IOP-29 FDS	.13 (.11)	.42 (.25)	.11 (.09)	.16 (.16)
% Cutoff $\geq .50$	2.5%	46%	2.7%	7.8%

\*SS-R, lower score indicates stronger underreporting tendencies

*Cohen's d* = .77), with the underreporters exhibiting higher levels of supernormality.

Using the SS-R cutoff of < 60, five participants in the control condition (12.5%) engaged in a supernormal presentation, and one participant in the overreporting group obtained score of 60 (2.7%). Among underreporting participants, 14 exhibited supernormality (39%), while in the mixed strategy condition two students did so (5.3%; see Table 2).

### Group Differences in Overreporting (IOP-29)

The FDS index was significantly different across the conditions,  $F(3, 147) = 30.10, p < .001, \eta_p^2 = .38$ . More specifically, the control group showed significantly lower scores than the overreporting group ( $p < .001, Cohen's d = 1.57$ ), but a very similar score to those of the underreporting ( $p \approx 1.00, Cohen's d = .14$ ) and mixed strategy ( $p \approx 1.00, Cohen's d = .14$ ) groups. The overreporting group, as expected, obtained higher FDS scores than the underreporting ( $p < .001, Cohen's d = 1.65$ ) and mixed strategy ( $p < .001, Cohen's d = 1.28$ ) groups, whereas underreporting participants and those employing mixed strategy did not differ from each other ( $p = 1.00, Cohen's d = .37$ ).

Applying the cutoff score of  $FDS \geq .50$ , one participant in the control group exhibited a noncredible symptom presentation (2.5%), 17 overreporters had FDS values  $\geq 0.50$  (46%), one participant in the underreporting condition obtained borderline score ( $FDS = .54; 2.7\%$ ), and three participants of mixed strategy group had FDS values  $\geq 0.50$  (7.8%; see Table 2).

### Correlation Between Pre-screening Scores and Post-manipulation PTSD Reporting

To inspect how preexisting levels of distress and trauma-related complaints reflected on students' responding post-manipulation, we ran Pearson product-moment correlations ( $r$ ) between the pre-screening scores (BSI-18 and NSESSS) and post-manipulation score (PCL-5) for each condition separately. For the control group ( $n = 40$ ), the correlation between the BSI-18 and PCL-5 were high and significant,  $r = .82, p < .001$ . For NSESSS and PCL-5, the correlations were similarly robust with  $r$  being .93,  $p < .001$ . For the overreporting condition ( $n = 37$ ), the correlation between BSI-18 and PCL-5 remained non-significant,  $r = .24, p = .15$ , whereas that between NSESSS and PCL-5 was significant, albeit of modest size,  $r = .35, p = .04$ . The pre-screening scores on BSI-18 and NSESSS of the underreporting group ( $n = 36$ ) were moderately and significantly related to the PCL-5,  $r = .37, p = .03$  and  $r = .44, p = .008$ , respectively. Lastly, for the mixed strategy condition ( $n = 38$ ), the correlation pattern resembled that found in the control condition, with BSI-18 and PCL-5 correlating at  $r = .65, p < .001$ , and NSESSS and PCL-5 correlating at  $r = .65, p < .001$ .

## Discussion

In this experimental simulation study, we examined a number of different response strategies and related results on some relevant test scores. In addition to the commonly tested extreme points of the response bias spectrum—overreporting and underreporting—we also included an often-overlooked type of response bias—simultaneous over- and under-reporting (i.e., mixed strategy condition). Our findings can be summarized as follows: first, it is worth noting that participants' scores on the PTSD checklist differed significantly in the order we expected, with the underreporting group obtaining the lowest scores, followed by the control and mixed strategy groups, and then by the overreporting group, which exhibited the highest levels of PTSD symptoms. Important to note is that, based on the pre-screening scores (BSI and NSESSS), there were no differences between groups in the levels of distress and PTSD-like symptoms prior to the instructions, meaning that the group differences on PCL-5 can be explained by our manipulation. Interestingly, 20% of the control group (i.e., honest participants) obtained scores indicative of clinical levels of PTSD, which fits well with the prevalence of trauma among students shown in a large multisite study (21%, Frazier et al., 2009; Sharp & Theiler, 2018; see also Stallman, 2010). The pattern in the mixed strategy condition was very similar to those of the control group, with 26% providing PCL-5 scores above the cutoff. This suggests that simultaneously using both over- and under-reporting might result in a clinical test profile that looks authentic. The majority of the overreporting group crossed the PCL-5 cutoff (76%), and, to our surprise, one participant in the underreporting condition obtained PCL-5 score above the screening point. As participants had to pass the attention checks and to provide a proper elaboration of their task to be kept in the dataset, it is unlikely that this participant was just responding randomly. It is, for instance, possible that this participant might have experienced a high baseline of PTSD symptoms.

Second, unlike all other group comparisons for this variable, the supernormality scores of the control and mixed strategy group were rather similar, again showing that instructing participants to be selective in the features they need to exaggerate and to hide might lead to a balanced presentation. Supernormality levels were, as expected, the highest for the underreporting group and the lowest for the overreporting group, while the control and mixed strategy groups stayed in the middle-range. The results of the control participants, of whom 12.5% attained a score indicative of underreporting, signal that a non-trivial minority of our students spontaneously engaged in supernormality. Our findings align with previous research showing that some students conceal their mental and emotional issues (Martin, 2010), thereby obscuring prevalence estimates of serious symptoms and disorders that this population might be particularly vulnerable to.

Third, with regard to the IOP-29 scores, the overreporting group obtained the highest FDS values, whereas all other groups did not significantly differ from each other. As this measure was specifically designed to detect feigned psychiatric and/or cognitive problems (Viglione et al., 2017; for a review, see Giromini & Viglione, 2022), it is not surprising that the control and underreporting groups did not differ from each other. Regarding the mixed strategy group, it is possible that the instruction to only partially overreport while still appear functional resulted in psychological portrayals that were too subtle for the IOP-29 to classify as noncredible. Indeed, the reference to “empathy and personal difficult experiences [being] highly cherished” is more likely to solicit the feigning of a history of PTSD rather than the feigning of ongoing PTSD complaints, which explains the absence of elevations on the IOP-29. This explanation is also consistent with the fact that the PCL-5 scores of this group were very similar to those of the control group.

With regard to the overreporting group, from which less than half crossed the FDS cutoff point, it should be emphasized that the PCL-5 and IOP-29 scores observed in this study were considerably lower than in many other published research studies (e.g., Blavier et al., 2023; Carvalho et al., 2021; Szogi & Sullivan, 2018). For example, in Szogi and Sullivan (2018), the average PCL-5 scores for the PTSD feigning groups ranged from 54.66 ( $SD = 12.66$ ) to 61.04 ( $SD = 12.09$ ), whereas our overreporting group had a notably lower mean PCL-5 score of 43.43, with a remarkably higher  $SD$  of 19.24. Along similar lines, while in our study the percentage of overreporters with an IOP-29 FDS  $\geq .50$  was only 46%, according to recent meta-analytic (Puente-López et al., 2023) and quantitative literature review (Giromini & Viglione, 2022) studies, the sensitivity of IOP-29 FDS  $\geq .50$  to feigned psychopathology is likely to range from 82 to 86%. Although we do not have a valid explanation for this unexpected finding, it is likely that the specific instructions we used in this study did not clearly enough convey the message that participants were supposed to feign PTSD. Indeed, as noted above, although our instructions informed participants that they had the option of receiving an extension if they faked or exaggerated trauma-related problems, they might have been too subtle, as we did not explicitly encourage them to engage in this behavior. Future studies with more explicit instructions would thus be beneficial to address this issue.

In any case, the main message of this study is that instructions leading to both over- and under-reporting of trauma-related symptoms are likely to result in a psychometric profile that is quite similar to the profile of a group instructed to respond honestly. Pending future replications, thus, it is possible that such mixed strategy may help individuals to evade detection on symptom validity tests. Accordingly, three new considerations are to be made: first, from a methodological

point of view, experimental studies on the diagnostic accuracy of symptom validity tests should consider including a mixed strategy group. Second, having clinical implications in mind, we need to know more about the settings that elicit this type of mixed strategy, so further investigation is necessary (see also Whitman et al., 2023). Third, diagnosticians and researchers in symptom validity domain are well advised to assess both over- and under-reporting, even if their primary goal is to detect feigned or exaggerated mental health problems. We further argue that, despite the room for improvement of currently available measures, inspecting a tendency to exaggerate and to conceal psychological and physical complaints significantly enhance the accuracy of any health assessment, and we encourage practitioners to include SVTs in their test battery.

## Limitations

The above presented results need to be considered in the context of limitations of this study. First, although our original sample met the requirements of power analyses, due to multiple exclusion criteria, our sample was slightly lower than needed. Therefore, future investigation should include larger samples. This is all the more relevant issue considering that the standard deviations on all of the measures indicated large variability of presentations. Also, it is important to note that our participants were highly proficient in English, but most of them were not native English speakers. Second, as we included young (mostly female), functional adults in our study, and given the experimental context of our study which notably limits its ecological validity, the generalization of our findings, especially to the forensic population, is severely constricted. Third, the study was conducted online and, despite multiple attention checks, we cannot be sure that our participants were not distracted during their participation. Fourth, the inclusion of the Supernormality Scale-Revised might not have been the best option to measure underreporting of PTSD, as its items are broad and do not specifically refer to this type of complaint. However, currently there are no other stand-alone validity measures for the detection of positive response bias in forensic context. Therefore, we urge our peers to devote more attention to this existing gap in our field. Fifth, although we used the task elaboration as an inclusion criterion, we cannot ensure that all of our participants understood or complied with given instructions. Due to the random allocation of participants, it is possible that some, already exhibiting high levels of PTSD-like symptoms, were also given instructions to exaggerate them, and then some participants without any complaints to underreport them. Further, our instructions might have been too flexible, giving participants the opportunity to decide what



is the best way to exhibit response bias, thereby introducing room for individual degrees of freedom in how they would approach over- or/and under-reporting. Further, we did not inspect the presence of response bias prior to the manipulation. Thus, we do not know whether a habitual individual responding style might have confounded our results (but see Van Helvoort et al., 2022). Yet, looking at the correlations between the pre-screening scores and post-manipulation test results suggests that the manipulation impacted participants' responding style. Specifically, the correlations for the manipulation groups were significantly lower than those in the control condition, with the exception of mixed strategy group for which the scores on two employed measures were associated to a higher degree. Finally, we failed to have in-depth exit interviews with participants in the mixed strategy condition. Therefore, it is difficult to determine whether they balanced their biased presentation or just opted to respond honestly in order to resolve their confusion about exhibiting both underreporting and overreporting tendencies. Future investigation of this type of responding might want to address the symptom profiles on the PCL-5 of mixed strategy groups. It may well be the case that certain PTSD symptoms (i.e., domains) lend themselves better to either under- or overreporting. Because of the limited sample size and a high degree of freedom in our instructions, we refrained from a thorough analysis of the interactions with symptom profiles, but it obviously is a topic that warrants further investigation. Still, as this topic has not been investigated before, we hope our findings encourage others to continue further investigation of different types of response bias that are exhibited simultaneously. It is likely that PTSD claims, including the reflection on the comorbidities, such as substance use and anger-control issues, might be particularly fitting for testing the mixed response strategy.

## Conclusion

This is the first attempt to systematically compare multiple types of response bias in trauma symptom reports (overreporting, underreporting and mixed strategy reporting). Our findings regarding symptom presentation showed the anticipated trend of results: overreporting group exhibited the worst symptom presentation, underreporters the best, and mixed strategy condition provided the most balanced symptom reports, very much like that of the control group. Thus, the simultaneous over- and under-reporting of symptoms can lead to presentations resembling those of honest responders, not only on clinical measures but also on symptom validity tests. Arguably, given the clinical implications, this is an issue that deserves more study, first and foremost outside the laboratory.

## Appendix

### Instructions

#### Control (Honest) Group

Thank you for responding to the demographic questions and our questionnaires. The main part of the study is about to start.

Now we need you to fill out three other questionnaires in the most honest manner, meaning that you should answer in a truthful manner so that your responses correspond to how you are feeling now and during the last week. Please, read each of the item carefully, as we will include checks for random responding. In order to see the questionnaires, please click "Next" in the bottom right corner.

#### Overreporting Group

Thank you for responding to the demographic questions and our questionnaires. The main part of the study is about to start.

Dear participant, in order to properly conduct this study, we will have to ask you to imagine being in a particular situation. Please, pay attention and carefully read the instructions. Later, we will ask you to describe your task in your own words, and if you fail to do so, your data will not be considered valid. Here are the instructions:

Things are simply not great right now. You are spending too much time binge-watching Netflix/Disney/AppleTv, and chatting with your friends. You are simply not motivated to study and deal with the uni stuff right now. Yet, the exams are just around the corner, and it is not looking great for you at the moment. You know that you are not going to pass the first exam that is coming, as you did not read anything and were barely taking any notes on those couple of lectures you attended. But, there is a possibility to make things a bit easier right now. Namely, you could just go to the student psychologist and claim that you are experiencing trauma-related symptoms right now. They will give you three tests to fill out, but if you are convincing enough, this will be official and with that you can ask for an extension or an additional exam period! However, keep in mind that you need to be believable and not to overdo it, otherwise they will know that you are lying. Still, remember that this would mean a great deal to you as you would have a chance to not only pass, but to score very well on the exam! So, imagine being in this situation and with that in mind go through our study. You went to the psychologists, said what is the issue and they are giving you some tests to fill out.

From this moment onward, this study is the assessment from the instructions. Please keep in mind the story you just read the instructions and fill out the tests accordingly. Pay attention to the items, as some of them only serve to check whether you are paying attention and if you fail them, your data will not be seen as valid.

Also, remember, that the most convincing report will be selected and will be awarded with a €10 voucher! In order to see the questionnaires, please click “Next” in the bottom right corner.

### Underreporting Group

Thank you for responding to the demographic questions and our questionnaires. The main part of the study is about to start. Dear participant, in order to properly conduct this study, we will have to ask you to imagine being in a particular situation. Please, pay attention and carefully read the instructions. Later, we will ask you to describe your task in your own words, and if you fail to do so, your data will not be considered valid. Here are the instructions: We want you to imagine that you are applying for a job, a position as a psychologist at a well-respected, old-school, institution. There are big names and famous people from our line of work there, so it would be brilliant if you scored that job. Your career would set off in the best way possible! They have a motto that every employee should be first a professional, and then a human being. Being precise and professional is highly cherished. Also, they take their job application process very seriously. They will even give you three tests to fill out. It is the best option for you to present yourself in the best light, with no history of any issues, including trauma, but still, keep in mind that you need to be believable and not to overdo it, otherwise they will know that you are lying. So, try to find an optimal way to present yourself in order to impress them and score this job!

From this moment onward, this study is the assessment from the instructions. Please keep in mind the story you just read the instructions and fill out the tests accordingly. Pay attention to the items, as some of them only serve to check whether you are paying attention and if you fail them, your data will not be seen as valid. Also, remember, that the most convincing report will be selected and will be awarded with a €10 voucher! In order to see the questionnaires, please click “Next” in the bottom right corner.

### Mixed Reporting Group

Thank you for responding to the demographic questions and our questionnaires. The main part of the study is about to start. Dear participant, in order to properly conduct this study, we will have to ask you to imagine being in a particular situation. Please, pay attention and carefully read the instructions. Later, we will ask you to describe your task in your own words, and if you fail to do so, your data will not be considered valid. Here are the instructions: We want you to imagine that you are applying for a job, a position as a psychotherapist at a famous, new age, trauma center. There are big names and

famous people from our line of work there, so it would be brilliant if you scored that job. Your career would set off in the best way possible! They have a motto that every employee should be first a human being and then a therapist. Empathy and personal difficult experiences are highly cherished. Also, they take their job application process very seriously. They will even give you three tests to fill out in order to screen you. It is the best option for you to present yourself in the best light, but still, keep in mind that you need to be relatable, especially to the patients you will be working with. It is a trauma center so some personal struggles with related issues would give you an advantage! Also, you need to be believable and not to overdo it, otherwise they will know that you are lying. So, try to find an optimal way to present yourself in order to impress them and score this job!

From this moment onward, this study is the assessment from the instructions. Please, keep in mind the story you just read the instructions and fill out the tests accordingly. Pay attention to the items, as some of them only serve to check whether you are paying attention and if you fail them, your data will not be seen as valid. Also, remember, that the most convincing report will be selected and will be awarded with a €10 voucher! In order to see the questionnaires, please click “Next” in the bottom right corner.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12207-024-09503-w>.

**Author Contribution** All listed authors significantly contributed to the execution of the study and to this manuscript.

**Data Availability** The data and the outputs of our analyses are uploaded on the Open Science Framework (OSF) platform (link: <https://osf.io/huw67/>).

### Declarations

**Ethics Approval** All procedures performed in this study were in accordance with the ethical standards of the Ethical committee of Erasmus University Rotterdam, the Netherlands, and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

**Competing Interests** The second author is a co-creator of one of the used measures, whereas the sixth author is a co-creator of the original version of the other used test.

**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

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th Edition). Author. <https://doi.org/10.1176/appi.books.9780890425596>
- American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders -5-TR* (5th Ed., text revision). <https://doi.org/10.1176/appi.books.9780890425787>
- Baer, R. A., & Miller, J. (2002). Underreporting of psychopathology on the MMPI-2: A meta-analytic review. *Psychological Assessment*, 14(1), 16. <https://doi.org/10.1037/1040-3590.14.1.16>
- Blavier, A., Palma, A., Viglione, D. J., Zennaro, A., & Giromini, L. (2023). A natural experiment design testing the effectiveness of the IOP-29 and IOP-M in assessing the credibility of reported PTSD symptoms in Belgium. *Journal of Forensic Psychology Research and Practice*. <https://doi.org/10.1080/24732850.2023.2203130>
- Bošković, I. (2022). Trust, doubt, and symptom validity. In S. Landstrom, P. A. Granhag, & P. J. van Koppen (Eds.), *The Future of Forensic Psychology: Core Topics and Emerging Trends*, 13. Routledge.
- Boskovic, I. (2020). Do motives matter? A comparison between positive and negative incentives in students' willingness to malingering. *Educational Psychology*, 40(8), 1022–1032. <https://doi.org/10.1080/01443410.2019.1704400>
- Boskovic, I., Akca, A. Y. E., & Giromini, L. (2022). Symptom coaching and symptom validity tests: An analog study using the structured inventory of malingered symptomatology, self-report symptom inventory, and inventory of problems-29. *Applied Neuropsychology: Adult*. <https://doi.org/10.1080/23279095.2022.2057856>
- Boskovic, I., Dibbets, P., Bogaard, G., Hope, L., Jelacic, M., & Orthey, R. (2019a). Verify the scene, report the symptoms: Testing the verifiability approach and SRSI in the detection of fabricated PTSD claims. *Legal and Criminological Psychology*, 24(2), 241–257. <https://doi.org/10.1111/lcrp.12149>
- Boskovic, I., Hope, L., Ost, J., Orthey, R., & Merckelbach, H. (2019b). Detecting feigned high impact experiences: A symptom over-report questionnaire outperforms the emotional Stroop task. *Journal of Behavior Therapy and Experimental Psychiatry*, 65, 101483. <https://doi.org/10.1016/j.jbtep.2019.101483>
- Boskovic, I., Orthey, R., Otgaar, H., Mangiulli, I., & Rassin, E. (2023). #StudentsToo: Prevalence of sexual assault reports among students of three European universities and their actions post-assault. *PLoS ONE*, 18(4), 1–18. <https://doi.org/10.1371/journal.pone.0283554>
- Boyd, J. E., Cameron, D. H., Shnaider, P., McCabe, R. E., & Rowa, K. (2021). Sensitivity and specificity of the Posttraumatic Stress Disorder Checklist for DSM-5 in a Canadian psychiatric outpatient sample. *Journal of Traumatic Stress*, 35, 424–433. <https://doi.org/10.1002/jts.22753>
- Brady, K. T., Back, S. E., & Coffey, S. F. (2004). Substance abuse and posttraumatic stress disorder. *Current Directions in Psychological Science*, 13(5), 206–209. <https://doi.org/10.1111/j.0963-7214.2004.00309>
- Burges, C., & McMillan, T. M. (2001). The ability of naive participants to report symptoms of post-traumatic stress disorder. *British Journal of Clinical Psychology*, 40(2), 209–214. <https://doi.org/10.1348/014466501163544>
- Carvalho, L., Reis, A., Colombarolli, M. S., Pasian, S. R., Miguel, F. K., Erdodi, L. A., Viglione, D. J., & Giromini, L. (2021). Discriminating feigned from credible PTSD symptoms: A validation of a Brazilian version of the Inventory of Problems-29 (IOP-29). *Psychological Injury and Law*, 14(1), 58–70. <https://doi.org/10.1007/s12207-021-09403-3>
- Cima, M., Merckelbach, H., Hollnack, S., Butt, C., Kremer, K., Schellbach-Matties, R., & Muris, P. (2003). The other side of malingering: Supernormality. *The Clinical Neuropsychologist*, 17(2), 235–243. <https://doi.org/10.1076/clin.17.2.235.16507>
- 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. <https://doi.org/10.1111/j.1556-4029.2008.00740.x>
- Derogatis, L. R. (2001). Brief symptom inventory 18. *Johns Hopkins University*. <https://doi.org/10.1037/t07502-000>
- Edens, J. F., Buffington, J. K., & Tomacic, T. L. (2001). Effects of positive impression management on the psychopathic personality inventory. *Law and Human Behaviour*, 25, 235–256. <https://doi.org/10.1023/A:1010793810896>
- Faust, D. (2023). Invited Commentary: Advancing but not yet Advanced: Assessment of Effort/Malingering in Forensic and Clinical Settings. *Neuropsychology Review*, 33, 628–642. <https://doi.org/10.1007/s11065-023-09605-3>
- Frazier, P., Anders, S., Perera, S., Tomich, P., Tennen, H., Park, C., & Tashiro, T. (2009). Traumatic events among undergraduate students: Prevalence and associated symptoms. *Journal of Counseling Psychology*, 56(3), 450–460. <https://doi.org/10.1037/a0016412>
- Freeman, T., Powell, M., & Kimbrell, T. (2008). Measuring symptom exaggeration in veterans with chronic posttraumatic stress disorder. *Psychiatry Research*, 158(3), 374–380. <https://doi.org/10.1016/j.psychres.2007.04.002>
- Gilmoor, A. R., Adithy, A., & Regeer, B. (2019). The cross-cultural validity of post-traumatic stress disorder and post-traumatic stress symptoms in the Indian context: A systematic search and review. *Frontiers in Psychiatry*, 10, 439–450. <https://doi.org/10.3389/fpsy.2019.00439>
- Giromini, L., & Viglione, D. J. (2022). Assessing negative response bias with the Inventory of Problems-29 (IOP-29): A quantitative literature review. *Psychological Injury and Law*, 15(1), 79–93. <https://doi.org/10.1007/s12207-021-09437-7>
- Giromini, L., Young, G., & Sellbom, M. (2022). Assessing negative response bias using self-report measures: Introducing the special issue. *Psychological Injury and Law*, 15, 1–21. <https://doi.org/10.1007/s12207-022-09444-2>
- Griffin, B., & Wilson, I. G. (2012). Faking good: Self-enhancement in medical school applicants. *Medical Education*, 46(5), 485–490. <https://doi.org/10.1111/j.1365-2923.2011.04208.x>
- Guriel, J., & Fremouw, W. (2003). Assessing malingered posttraumatic stress disorder: A critical review. *Clinical Psychology Review*, 23(7), 881–904. <https://doi.org/10.1016/j.cpr.2003.07.001>
- Hall, R. C., & Hall, R. C. (2006). Malingering of PTSD: Forensic and diagnostic considerations, characteristics of malingerers and clinical presentations. *General Hospital Psychiatry*, 28(6), 525–535. <https://doi.org/10.1016/j.genhosppsych.2006.08.011>
- Holcomb, M., Pyne, S., Cutler, L., Oikle, D. A., & Erdodi, L. A. (2023). Take their word for it: The inventory of problems provides valuable information on both symptom and performance validity. *Journal of Personality Assessment*, 105(4), 520–530.
- Jackson, R. L., & Harrison, K. S. (2018). Assessment of law enforcement personnel: The role of response styles. In R. Rogers & D. S. Bender (Eds.), *Clinical assessment of malingering and deception* (pp. 552–570). Guilford Publications.
- Kilpatrick, D. G., Resnick, H. S., Milanak, M. E., Miller, M. W., Keyes, K. M., & Friedman, M. J. (2013). National estimates of exposure to traumatic events and PTSD prevalence using DSM-IV and DSM-5 criteria. *Journal of Traumatic Stress*, 26(5), 537–547. <https://doi.org/10.1002/jts.21848>

- Lapham, S. C., Smith, E., C'de Baca, J., Chang, I., Skipper, B. J., Baum, G., & Hunt, W. C. (2001). Prevalence of psychiatric disorders among persons convicted of driving while impaired. *Archives of General Psychiatry*, *58*(10), 943–949. <https://doi.org/10.1001/archpsyc.58.10.943>
- Lavashina, J. (2018). Evaluating deceptive impression management in personnel selection and job performance. In R. Rogers & D. S. Bender (Eds.), *Clinical assessment of malingering and deception* (pp. 530–551). Guilford Publications.
- LeBeau, R., Mischel, E., Resnick, H., Kilpatrick, D., Friedman, M., & Craske, M. (2014). Dimensional assessment of posttraumatic stress disorder in DSM-5. *Psychiatry Research*, *218*(1-2), 143–147. <https://doi.org/10.1016/j.psychres.2014.03.032>
- Lee, D. J., Weathers, F. W., Thompson-Hollands, J., Sloan, D. M., & Marx, B. P. (2022). Concordance in PTSD symptom change between DSM-5 versions of the Clinician-Administered PTSD Scale (CAPS-5) and PTSD Checklist (PCL-5). *Psychological Assessment*, *34*(6), 604–609. <https://doi.org/10.1037/pas0001130>
- Levashina, J., & Campion, M. A. (2007). Measuring faking in the employment interview: Development and validation of an interview faking behavior scale. *Journal of Applied Psychology*, *92*, 1638–1656. <https://doi.org/10.1037/0021-9010.92.6.1638>
- Magura, S. (2010). Validating self-reports of illegal drug use to evaluate National Drug Control Policy: A reanalysis and critique. *Evaluation and Program Planning*, *33*(3), 234–237. <https://doi.org/10.1016/j.evalproplan.2009.08.004>
- Martin, J. M. (2010). Stigma and student mental health in higher education. *Higher Education Research & Development*, *29*(3), 259–274. <https://doi.org/10.1080/07294360903470969>
- Melchers, K. G., Roulin, N., & Buehl, A. K. (2020). A review of applicant faking in selection interviews. *International Journal of Selection and Assessment*, *28*(2), 123–142. <https://doi.org/10.1111/ijsa.12280>
- Merckelbach, H. (2004). Telling a good story: Fantasy proneness and the quality of fabricated memories. *Personality and Individual Differences*, *37*(7), 1371–1382. <https://doi.org/10.1016/j.paid.2004.01.007>
- Morey, L. C. (1991). Personality assessment inventory. Odessa, FL: Psychological Assessment Resources.
- Morey, L. C., Warner, M. B., & Hopwood, C. J. (2007). The personality assessment inventory: Issues in legal and forensic settings. *Forensic Psychology: Emerging Topics and Expanding Roles*, 97–126.
- Paulhus, D. L. (2012). Overclaiming on personality questionnaires. In M. Ziegler, C. MacCann, & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 151–164). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195387476.003.0045>
- Picard, E., Aparcero, M., Nijdam-Jones, A., & Rosenfeld, B. (2023). Identifying positive impression management using the MMPI-2 and the MMPI-2-RF: A meta-analysis. *The Clinical Neuropsychologist*, *37*(3), 545–561. <https://doi.org/10.1080/13854046.2022.2077237>
- Puente-López, E., Pina, D., López-Nicolás, R., Iguacel, I., & Arce, R. (2023). The inventory of problems–29 (IOP-29): A systematic review and bivariate diagnostic test accuracy meta-analysis. *Psychological Assessment*, *35*(4), 339–352. <https://doi.org/10.1037/pas0001209>
- REINN, 2016. The criminal justice system: Statistics. <https://www.rainn.org/statistics/criminal-justice-system>
- Resnick, P. J., West, S. G., & Wooley, C. N. (2018). The malingering of posttraumatic disorders. In R. Rogers & D. S. Bender (Eds.), *Clinical assessment of malingering and deception* (pp. 188–211). Guilford Publications.
- Roberts, N. P., Kitchiner, N. J., Lewis, C. E., Downes, A. J., & Bisson, J. I. (2021). Psychometric properties of the PTSD Checklist for DSM-5 in a sample of trauma exposed mental health service users. *European Journal of Psychotraumatology*, *12*, 1863578. <https://doi.org/10.1080/20008198.2020.1863578>
- Rogers, R., & Bender, S. D. (Eds.). (2018). *Clinical assessment of malingering and deception (4th Edition)*. Guilford Publications.
- Rogers, R. (2018). An introduction to response style. In R. Rogers & D. S. Bender (Eds.), *Clinical assessment of malingering and deception* (pp. 3–17). Guilford Publications.
- Roma, P., Giromini, L., Sellbom, M., Cardinale, A., Ferracuti, S., & Mazza, C. (2023). The ecological validity of the IOP-29: A follow-up study using the MMPI-2-RF and the SIMS as criterion variables. *Psychological Assessment*, *35*, 868–879. <https://doi.org/10.1037/pas0001273>
- Ruback, R. B., & Hopper, C. H. (1986). Decision making by parole interviewers: The effect of case and interview factors. *Law and Human Behavior*, *10*, 203–214. <https://doi.org/10.1007/BF01046210>
- Sharp, J., & Theiler, S. (2018). A review of psychological distress among university students: Pervasiveness, Implications and potential points of intervention. *International Journal for the Advancement of Counselling*, *40*, 193–212. <https://doi.org/10.1007/s10447-018-9321-7>
- Sherman, E. M., 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/accaa019>
- Stallman, H. M. (2010). Psychological distress in university students: A comparison with general population data. *Australian Psychologist*, *45*(4), 249–257. <https://doi.org/10.1080/00050067.2010.482109>
- Szogi, E. G., & Sullivan, K. A. (2018). Malingered posttraumatic stress disorder (PTSD) and the effect of direct versus indirect trauma exposure on symptom profiles and detectability. *Psychological Injury and Law*, *11*, 351–361. <https://doi.org/10.1007/s12207-018-9315-0>
- Van Helvoort, D., Merckelbach, H., Van Nieuwenhuizen, C., & Otgaar, H. (2022). Traits and distorted symptom presentation: A scoping review. *Psychological Injury and Law*, *15*(2), 151–171. <https://doi.org/10.1007/s12207-022-09446-0>
- Viglione, D. J., Giromini, L., & Landis, P. (2017). 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. <https://doi.org/10.1080/00223891.2016.1233882>
- Walters, G. D., Rogers, R., Berry, D. T., Miller, H. A., Duncan, S. A., McCusker, P. J., & Granacher, R. P., Jr. (2008). Malingering as a categorical or dimensional construct: The latent structure of feigned psychopathology as measured by the SIRS and MMPI-2. *Psychological Assessment*, *20*(3), 238–247. <https://doi.org/10.1037/1040-3590.20.3.238>
- Weathers, F. W., Blake, D. D., Schnurr, P. P., Kaloupek, D. G., Marx, B. P., & Keane, T. M. (2013). The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5). Available from [www.ptsd.va.gov](http://www.ptsd.va.gov)
- Weathers, F. W. (2008). Posttraumatic stress disorder checklist. In G. Reyes, J. D. Elhai, & J. D. Ford (Eds.), *Encyclopedia of psychological trauma* (pp. 491–494). Wiley.
- Whitman, M. R., Gervais, R. O., & Ben-Porath, Y. S. (2023). Virtuous victims: Disability claimants who over-and under-report. *The Clinical Neuropsychologist*. <https://doi.org/10.1080/13854046.2023.2185686>
- Wilson, L. C., & Miller, K. E. (2016). Meta-analysis of the prevalence of unacknowledged rape. *Trauma, Violence, & Abuse*, *17*, 149–159. <https://doi.org/10.1177/1524838015576391>
- Young, G. (2016). PTSD in court I: Introducing PTSD for court. *International Journal of Law and Psychiatry*, *49*, 238–258. <https://doi.org/10.1016/j.ijlp.2016.10.012>

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