Discriminating Feigned from Credible PTSD Symptoms: a Validation of a Brazilian Version of the Inventory of Problems-29 (IOP-29)

Lucas de Francisco Carvalho[1] · Ana Reis[1] · Maíra Stivaleti Colombarolli[2,3] · Sonia Regina Pasian[2] ·
Fabiano Koich Miguel[4] · Laszlo A. Erdodi[5] · Donald J. Viglione[6] · Luciano Giromini[3]

Abstract

The Inventory of Problems-29 (IOP-29) is a recently introduced free-standing symptom validity test (SVT) with a rapidly growing evidence base. Its classification accuracy compares favorably with that of the widely utilized Structured Inventory of Malingered Symptomology (SIMS), and it provides incremental validity when used in combination with other symptom and performance validity tests. This project was designed to cross-validate the IOP-29 in a Brazilian context. Study 1 focused on specificity and administered the IOP-29 and a PTSD screening checklist to 154 Brazilian firefighters who had been exposed to one or more potentially traumatic stressors. Study 2 implemented a simulation/analogue research design and administered the IOP-29, together with a new IOP-29 add-on memory module, to nonclinical volunteers; 101 asked to respond honestly, 100 instructed to feign PTSD. Taken together, the results of both study 1 (specificity = .96) and study 2 (Cohen's d = 2.15; AUC = .92) support the validity, effectiveness, and cross-cultural applicability of the IOP-29. Additionally, study 2 provides preliminary evidence for the incremental utility of the newly introduced, IOP-29 add-on memory module. Despite the encouraging findings, we highlight that the determination of feigning or malingering should never be made off a single test alone.

Keywords PTSD · Malingering · Simulation · IOP-29 · Inventory of Problems

Posttraumatic stress disorder (PTSD) is one of the most relevant conditions to forensic work (Young, 2016). In the DSM-5 (APA, 2013), it is classified as a trauma and stressor-related disorder, and it is characterized by four symptom clusters: trauma-related re-experiencing, avoidance, negative alterations in cognition or mood, and hyperarousal. Since PTSD is typically attributed to specific events, it is one of the several psychological injuries that may lead to litigation (Young, 2016, 2017a, 2017b). Common examples of situations in which the forensic assessor is asked to evaluate the possible presence of PTSD are workers' compensation claims, service-related disability claims in the military, and personal injury lawsuits.

A crucial component of any PTSD-related forensic evaluations is the assessment of negative response bias (Bass & Hallian, 2014; Merten et al., 2009). Indeed, in the presence of external incentives to appear impaired (personal injury compensation, avoiding service), forensic assessors should examine the credibility of the presentation (Boone, 2013; Resnick et al., 2018; Young, 2017b), using objective psychometric evidence. Symptom validity tests (SVTs) are interviews and self-report measures designed to evaluate the veracity of various symptoms. In contrast, performance validity tests (PVTs) are designed to estimate the extent to which scores on cognitive tests reflect the examinee's true ability level. Recent evidence suggests that combining SVTs and PVTs in an assessment battery likely improves signal detection accuracy over using SVTs only or PVTs only (Boone, 2013; Fox & Vincent, 2020; Giromini et al., 2020; Larrabee, 2008; Rogers & Bender, 2018).

A novel instrument to assess the credibility of emotional and cognitive complaints is the Inventory of Problems-29 (IOP-29; Viglione & Giromini, 2020). Designed as

✉ Luciano Giromini
luciano.giromini@unito.it

1 São Francisco University, Bragança Paulista, Brazil

2 University of São Paulo, São Paulo, Brazil

3 Department of Psychology, University of Turin, Via Verdi 10, 10123 Torino, TO, Turin, Italy

4 State University of Londrina, Londrina, Brazil

5 University of Windsor, Windsor, Canada

6 Alliant International University, San Diego, USA

an omnibus test to evaluate the credibility of self-reported symptoms of PTSD, depression, schizophrenia, and cognitive decline, it is comprised of 29 self-administered items intermixing self-report (i.e., SVT-like) with a few cognitive (i.e., PVT-like) items. Its chief feigning scale, the False Disorder Probability Score (FDS), is not normed on healthy, non-clinical controls. Instead, it is derived from the IOP-29′s ability to differentiate between two reference groups: bona fide patients and experimental feigners (i.e., individuals instructed to feign). According to Viglione and Giromini (2020), this approach better discriminates between credible and non-credible profiles. Derived from a logistic regression equation, the FDS provides practitioners with an estimated likelihood of obtaining a given IOP-29 from a group of experimental feigners versus a group of individuals with genuine psychiatric disorders. As such, the FDS ranges from 0 to 1, with higher scores reflecting increasing likelihood of non-credible presentations.

In their developmental work conducted in the USA, Viglione et al. (2017) tested the IOP-29 with several hundreds of individuals with presumably bona fide (often severe) symptoms[1] and a comparable sample size of healthy individuals instructed to appear mentally ill (i.e., experimental feigners). The classification accuracy of the IOP-29 approached or exceeded that of the validity scales of more complex measures such as the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher et al., 1989; Green, 1991) or Personality Assessment Inventory (PAI; Morey, 2007). Subsequently, Giromini et al. (2018) conducted an independent clinical comparison simulation/analogue study on 216 bona fide patients and 236 experimental feigners and found that the classification accuracy of the IOP-29 significantly outperformed that of the substantially longer Structured Inventory of Malingered Symptoms (SIMS; Smith & Burger, 1997; Widows & Smith, 2005). Indeed, when considering the entire sample ($N = 452$), the receiver operator characteristic curve (area under the curve (AUC)) was 0.89 (SE = 0.02) for the IOP-29 and 0.83 (SE = 0.02) for the SIMS; Cohen's d were 1.93 for the IOP-29 and 1.39 for the SIMS.

These two initial studies inspired several subsequent replications. During the 2019–2020 academic year alone, 10 research articles focused on the IOP-29 were published, covering various geographical regions and languages (Gegner et al., 2021; Giromini, Barbosa et al., 2019; Giromini, Carfora Lettieri et al., 2019; Giromini, Viglione

et al., 2019a; Giromini, Viglione et al., 2019b; Giromini, Viglione et al., 2020; Ilgunaite et al., 2020; Roma et al., 2020; Viglione et al., 2019; Winters et al., 2020). Taken together, the results of these research efforts demonstrated that (a) the clinical utility of the IOP-29 when evaluating the credibility of PTSD-, depression-, mild traumatic brain injury-, and schizophrenia-related presentations is reliably replicated (Giromini, Viglione et al., 2019a); (b) using the IOP-29 together with the MMPI-2 or Test of Memory Malingering (Tombaugh, 1996) improves signal detection compared with using either test alone (Giromini, Barbosa et al., 2019; Giromini, Carfora Lettieri et al., 2019); (c) the performance of the IOP-29 in a real-life forensic context is similar to that achieved in an experimental context (Roma et al., 2020); and (d) the IOP-29 generalizes to cultures outside of the USA, including Italy (Roma et al., 2020), Australia (Gegner et al., 2021), Portugal (Giromini, Barbosa et al., 2019), Lithuania (Ilgunaite et al., 2020), or the UK (Winters et al., 2020).

More recently, an add-on PVT module of the IOP-29 was introduced: the Inventory of Problems-Memory (IOP-M; Giromini et al., 2020). The IOP-M consists of a 34 forced-choice memory recognition items that are administered immediately after the IOP-29. The examinee is presented with pairs of words or brief phrases, and is instructed to identify the ones that were part of the IOP-29 item content. According to Giromini et al. (2020), using the IOP-M in combination with the IOP-29 has the potential to significantly improve signal detection over using the IOP-29 alone, because it allows combining key SVT with key PVT detection strategies. However, their findings lack the extensive replication available for the IOP-29.

Overview of the Current Project

Brazil has a population larger than Russia with over 200 million citizens where an evidence-based psychological assessment is becoming a standard practice. This is achieved through local guidelines (e.g., Brazilian Institute of Psychological Assessment (IBAP)) based on international recommendations (e.g., AERA, APA, & NCME, 2013). As is the case in many other countries, in Brazil, the diagnosis of PTSD, established through psychological assessment, can lead to legal actions in court. For instance, when a causal link between a work-related traumatic event and subsequent psychological illness is established, the employee is eligible for various benefits (Souto et al., 2012). In such contexts, forensic practitioners are asked to use evidence-based procedures to assess the causal link between the precipitating traumatic event and resulting impairment, including issues such as temporary or permanent disability and estimating the future cost of treatments. Naturally, the legitimacy of the

---

[1] Viglione et al. (2017) report that patients included in their original samples were recruited through inpatient hospitals, outpatient clinics, day treatment centers, or criminal investigation units, and that their diagnoses were verified via various approaches, including structured and unstructured interviews, psychological testing, etc

claim rests on the assumption that the presentation is credible and is causally related to the claimed work-related injury (Bass & Hallian, 2014; Young, 2016). To our knowledge, however, no validated SVTs are available for use within the Brazilian population at this time. In fact, the first author of this article is currently working on introducing a Brazilian Portuguese version of the MMPI and PAI, which have not been published in Brazil. Similarly, the authors are not aware of any other well-validated SVTs being available in Brazil. For these reasons, Brazilian forensic practitioners would likely benefit from the translation and validation of a Brazilian Portuguese version of IOP-29.

Although the IOP-29 was investigated in various cultural contexts, there is no research available on its clinical/forensic utility in Latin America. Thus, the current project was initiated to further investigate the cross-cultural adaptability of the IOP-29, and test its psychometric properties in a Brazilian sample. While the IOP-29 was being administered to Brazilian firefighters exposed to a number of potentially traumatic/destabilizing events (study 1), the IOP-M, was introduced. Therefore, it was added to a subsequent investigation focused on the incremental validity of the Brazilian IOP-M (study 2).

Study 1, conducted before the publication of the IOP-M, had two goals: (1) develop and validate the Brazilian Portuguese version of the IOP-29 via a translation/back-translation procedure (Brislin, 1980; Geisinger, 2003; Van de Vijver & Hambleton, 1996) and (2) test the newly developed Brazilian Portuguese IOP-29 within a credible sample of 154 firefighters previously exposed to one or more potentially traumatic/destabilizing events while carrying out their duty. Study 2, conducted after the publication of the IOP-M, also had two research goals: (1) develop an accurate Brazilian Portuguese version of the IOP-M via a translation/back-translation procedure and (2) investigate the classification accuracy of the Brazilian Portuguese IOP-29-M. In study 2, both the IOP-29 and IOP-M were administered to 101 non-clinical volunteers instructed to respond honestly and 100 non-clinical volunteers instructed to feign PTSD. Both study 1 and study 2 also included a DSM-5-based PTSD self-report screening measure, the PTSD Checklist for DSM-5 (PCL-5; Weathers et al., 2013).

Study 1

The primary purpose of study 1 was to develop a culturally informed Brazilian Portuguese translation of the IOP-29. The first step consisted of the translation and cultural adaptation of the IOP-29 from English to Brazilian Portuguese. Although a European Portuguese version of the IOP-29 had been developed and tested before (Giromini, Barbosa et al., 2019), Brazilian and European Portuguese are two different languages that differ on several aspects. Therefore, the European Portuguese version of the IOP-29 had to be revised, following a translation/back-translation procedure (Brislin, 1980; Geisinger, 2003; Van de Vijver & Hambleton, 1996). That is, the original (English) version of the IOP-29 was first translated into Brazilian Portuguese by a native Brazilian Portuguese speaker who was also fluent in English. Next, the newly generated Brazilian Portuguese IOP-29 was translated back to English by someone who was blind to the original English version of the IOP-29. Lastly, a group of researchers who fluent in both English and Brazilian Portuguese reviewed, together with the IOP-29 authors, both the original, translated, and back-translated versions of the test to identify and resolve any inconsistencies.

Method

The major goal of study 1 was to test the vulnerability of the Brazilian IOP-29 to false positives. More specifically, we were interested in investigating whether exposure to traumatic events could attenuate the signal detection performance of the IOP-29. Therefore, we recruited a sample of firefighters who had recently been exposed to various potentially traumatic/destabilizing events, and asked them to complete the IOP-29 under standard instructions. In other words, study 1 tested the specificity of the IOP-29 in a sample of Brazilian firefighters with legitimate reason to have PTSD-related symptoms. The project received ethical approval by the relevant Institutional Review Board (IRB).

Participants A sample of 154 firefighters aged 29 to 56 years old ($M = 43.1$; SD = 5.2) were recruited, who had been working in the field for 4 to 30 years ($M = 17.4$; SD = 5.3). The majority (59%) were men. In terms of education, 50.6% completed a high school, 12.3% initiated but did not finish a college, and 37.0% obtained a college degree. The majority (60.4%) were Afro-descendants, followed by White (23.4%) and Pardo (16.2%). All participants included in the study had been exposed to at least one of the 20 potentially traumatic/destabilizing events described in Table 1. However, more than 90% had been exposed to at least three events, with an average of 6.5 (SD = 2.9) exposures. It should be underscored that none had any incentive to appear impaired.

Measures All participants were administered the IOP-29 and PCL-5 along with a demographic form and a brief questionnaire inquiring on whether or not they had been recently exposed to a series of potentially traumatic events.

The Inventory of Problems-29 (IOP-29; Viglione & Giromini, 2020) As reviewed in the Introduction, the IOP-29 is a

Table 1 Point biserial correlations of IOP-29 FDS to experienced stressors (*N* = 154)

Experienced stressor	% of Incidence	Correlation w/IOP-29 FDS
Flood	41.6%	−0.06
Landslide	72.1%	−0.22**
Storms	9.1%	0.07
Dry	10.4%	0.01
Epidemic	10.4%	−0.03
Chemical accidents	21.4%	0.00
Industrial accidents	20.1%	−0.07
Electrical accidents	40.9%	−0.13
Boat accidents	27.9%	−0.24**
Road accidents	90.9%	−0.16*
Train accidents	4.5%	0.18*
Urban fires	90.3%	−0.14
Rural fires	46.1%	−0.05
Forest fires	66.9%	−0.02
Air crash	11.7%	0.16*
Urban violence	3.9%	−0.07
Occurrence of attempted suicide	29.9%	−0.20*
Divorce in the past year	13.1%	−0.08
Lost someone close or significant in the past year	29.4%	−0.13
Diagnosed with a serious illness in the last year	9.1%	0.04

[*]Uncorrected *p* < .05

[**]Uncorrected *p* < .01

brief, self-administered test designed to evaluate the credibility of presenting complaints. The majority of its 29 items focus on various psychological problems and experiences, offering three possible response options, i.e., *true*, *false*, and *does not make sense*. The remaining few items involve mathematical and/or logical problems solving. The responses provided to these 29 items are processed via the official IOP-29 website (www.iop-test.com) to generate the False Disorder probability Score (FDS), a probabilistic value that informs on the overall credibility of the IOP-29 presentation. Like all probabilities, it ranges from 0 to 1, with higher scores indicating less credible outcomes. According to the IOP-29 manual (Viglione & Giromini, 2020), a cutoff of FDS ≥ 0.50 offers the best balance between sensitivity and specificity, yielding an overall correct classification (OCC) rate of about 80%. A more conservative cutoff score of FDS ≥ 0.65 achieves a specificity of 90%, whereas a more liberal cut-score of FDS ≥ 0.30 has a sensitivity of 90%.

**The PTSD Checklist for DSM–5 (PCL-5; Weathers et al., 2013)** We used the PCL-5 Standard Form, "Past-Month" version without PTSD diagnosis criterion A (i.e., exposure to trauma). It is a 20-item screening self-report inventory designed to measure PTSD symptoms described in the DSM-5. Eight of the items refer to difficulties with a stressful experience. The other 12 refer to negative affects related to symptoms and experiences of PTSD, including depression, anxiety, irritability, and hyper-arousal. For each PCL-5 item, the respondent is asked to rate how much they have been bothered by the described problem during the last month, using a 5-point scale (0–4). The sum of individual item scores measures the severity of PTSD symptoms. According to Weathers et al. (2013), a total score between 31 and 33 is indicative of probable PTSD across samples. However, the PCL-5 does not establish causality between a particular trauma and subsequent PTSD symptoms—a key legal component in forensic assessment. More importantly, it should be underscored that the PCL-5 is a screening measure.

Osorio et al. (2017) introduced a Brazilian adaptation for the PCL-5. A second Brazilian study by Pereira-Lima et al. (2019) confirmed the strong psychometric properties of the instrument. In our study, Cronbach's alpha was 0.96 for the total PCL-5 score.

Procedures This study followed the ethical research procedures following the Declaration of Helsinki (WMA, 2013) and was approved by a Research Ethics Committee. First, researchers contacted the head of a Brazilian fire

Fig. 1 Graphical representation of IOP-29 FDS and PCL-5 scores (study 1; $N = 154$)
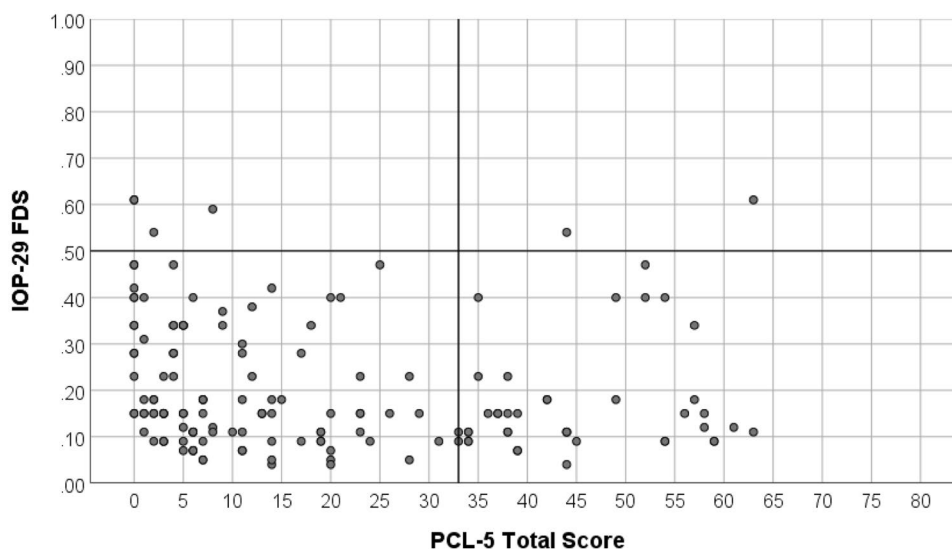
department, who provided contact information for all active duty firefighters in that department. All of them agreed to participate in the study, and completed the IOP-29 and PCL-5, along with a demographic form and a questionnaire about history of trauma exposure.

Data Analysis The main purpose of study 1 was to test the specificity of the IOP-29, and its vulnerability to false positive errors in credible test-takers who had been exposed to various potentially traumatizing/destabilizing stressors, the majority of which were work-related. The specificity of the IOP-29 was operationalized as the base rate of failure at the following cutoff scores FDS ≥ 0.65, ≥ 0.50, and ≥ 0.30. These analyses were computed first on the entire sample and then once again in individuals who scored ≥ 33 on the PCL-5 [a relatively conservative cutoff indicative of probable PTSD according to Weathers et al. (2013)]. Finally, we calculated the correlation between the IOP-29 FDS and the PCL-5 total score, as well as its point biserial correlations to each of the traumatic/destabilizing events (exposed = 1; not exposed = 0).

Results

The average PCL-5 score was 19.3 (SD = 18.5; range = 0–63); the average IOP-29 FDS was 0.20 (SD = 0.14; range = 0.04–0.61). Of the 154 firefighters, 43 (27.9%) had a PCL-5 score ≥ 33 indicative of a probable PTSD (Weathers et al., 2013). Within this subsample, the average IOP-29 FDS was 0.18 (SD = 0.14; range = 0.04–0.61).

A perfect specificity (1.00) was observed at the conservative cutoff of FDS ≥ 0.65, when considering the entire sample. At the standard cutoff of FDS ≥ 0.50, specificity was still very high (0.96), but it declined to 0.76 at the liberal cutoff of FDS ≥ 0.30, which is typically used for screening purposes only. When considering the subsample of firefighters who were at higher risk of PTSD (the 43 individuals with a PCL-5 score ≥ 33), specificity was 0.95 at the standard cutoff of FDS ≥ 0.50. However, a perfect specificity (1.00) was achieved at the conservative cutoff of FDS ≥ 0.65. Noteworthy, the IOP-29 FDS did not correlate with the total PCL-5 score, $r = -0.12$, $p = 0.14$ (Fig. 1).

Examination of Table 1 reveals that none of the stressors considered in this study influenced the IOP-29 FDS value. In fact, of the 20 tested correlations, none remained statistically significant after applying a Bonferroni correction (*corrected p* = 0.05/20 = 0.0025). Two correlations were positive and statistically significant at the uncorrected *p* value of 0.05, i.e., exposure to train ($r = 0.18$) or airplane ($r = 0.16$) crashes, with *small* to *medium* effect sizes (explaining only 2–3% of the variance; Cohen, 1988). Another four correlations were significant at the uncorrected *p* value of 0.05, but with a negative sign. Specifically, exposure to landslides ($r = -0.22$) and boat accidents ($r = -0.24$) yielded an uncorrected *p* value < 0.01, and exposure to the occurrence of an attempted suicide ($r = -0.20$) and to a road accident ($r = -0.16$) yielded an uncorrected *p* value < 0.05. The correlation of the IOP-29 FDS with the average number of stressors the participants had been exposed to was $r = -0.18$ ($p = 0.03$), thus confirming that exposure to stress did not increase the FDS; in fact, it perhaps even decreased it, though with a very small effect size. Conversely, the average number of stressors the participants had been exposed to correlated *positively* (as expected) with the total PCL-5 score, $r = 0.19$, $p = 0.02$.

Table 2 Demographic composition of study 2 (N = 201)

	Honest (n = 101)	Simulator (n = 100)
Age [$t(70) = 0.63$, $p = .53$, $d = 0.15$][a]		
M	29.0	30.4
SD	9.4	9.0
Education [$t(199) = 1.83$, $p = .07$, $d = 0.26$]		
M	10.2	10.9
SD	2.7	2.6
Gender [$Phi = 0.31$, $p < .01$]		
M	20 (19.8%)	49 (49.0%)
F	81 (80.2%)	51 (51.0%)

[a]Information on age was missing for 129 participants. Accordingly, for this analysis n = 72

Discussion

This study investigated the extent to which individuals with credible risk for PTSD-related conditions could be mistakenly classified as simulators by the Brazilian IOP-29. Taken together, the results of this study are consistent with the conclusion that exposure to a single or multiple potentially traumatic/destabilizing events does not elevate the IOP-29 FDS. Indeed, the standard cutoff score of the IOP-29 (FDS ≥ 0.50) achieved a specificity ≥ 0.95, both when considering the entire sample and when restraining the analyses to a subsample of 43 firefighters with probable PTSD according to PCL-5 results (i.e., PCL-5 ≥ 33; Weathers et al., 2013). Likewise, the IOP-29 FDS did not correlate with the total PCL-5 score.

On the other hand, it should be pointed out that (a) self-reported exposure to traumatic events on a brief screener alone would not meet criteria for PTSD; (b) this study did not perform a comprehensive assessment of posttraumatic conditions; (c) the determination of feigning or malingering should never be made off a single test alone, regardless of its apparent effectiveness; and (d) participants were active duty firefighters, who likely differ from civilians who sustained physical and emotional injuries severe enough to seek compensation (i.e., they were likely screened for mental and physical fitness; voluntarily exposed themselves to potentially traumatic events; are regularly compensated for doing so). Therefore, the relationship between exposure to traumatic events and IOP-29 FDS may be stronger in individuals assessed in clinical or forensic settings. As such, the perfect specificity observed in this sample may not generalize to more complex presentations.

Study 2

Given that study 1 focused on specificity, the primary purpose of study 2 was to examine the sensitivity of the IOP-29 in the detection of feigned PTSD. Additionally, study 2 also aimed at developing and testing a Brazilian version of the recently introduced IOP-M. A total of 201 non-clinical volunteers completed the IOP-29, IOP-M, and PCL-5. Half of them were randomly assigned to the control group (i.e., administered the tests under standard instructions; HON); half were instructed to feign PTSD (SIM).

Method

The first step of study 2 entailed the development of a Brazilian Portuguese version of the IOP-M, which was accomplished via a translation/back-translation procedure (Brislin, 1980; Geisinger, 2003; Van de Vijver & Hambleton, 1996). Unlike study 1, study 2 was conducted via the Internet.

Participants A sample of 201 non-clinical adult Brazilian volunteers was recruited via social media. Eligibility criteria included being aged 18 to 70, and reporting absence of current or past diagnosis of PTSD. Participants who failed to respond correctly to the post-test manipulation check were excluded from the study and their data were removed from the data set. Data collection was terminated when both the HON and SIM groups reached the desired sample size of n = 100.

Table 2 presents the demographic characteristics of the two samples. The majority (129 of 201) of the participants did not report their age. For those who did (n = 72), there were no notable differences between the HON and SIM groups, $t(70) = 0.63$, $p = 0.53$, $d = 0.15$. The level of education also did not differ by group, $t(199) = 1.83$, $p = 0.07$, $d = 0.26$. Conversely, the two samples were not balanced on gender, as a significantly higher percentage of women were assigned to the HON (80.2%) rather than in the SIM (51.0%) group: $phi = 0.31$, $p < 0.01$. Follow-up analyses, however, tested two two-way ANOVAs with group (HON vs. SIM) and gender (M vs. F) as between-subject factors and the IOP-29 FDS and IOP-M scores as dependent variables, and

found that there was no interaction between group and gender, $F(1, 197) \leq 0.93$, $p \geq 0.34$, partial $\chi^2 \leq 0.01$. As such, we concluded that the unequal distribution of genders across the two groups did not influence the IOP-29 or IOP-M scores, consistent with previous reports on the independence of gender and performance validity (Abeare et al., 2020; Erdodi et al., 2019).

Measures In addition to the IOP-29 and PCL-5 as described in study 1, the current study also administered the IOP-M. It should be noted that although study 2 primarily focused on the IOP-29 and IOP-M, we also administered the PCL-5 to document the proportion of individuals who would be classified as having PTSD in the presence of credible SVT results.

The Inventory of Problems-Memory module (IOP-M; Giromini et al., 2020) The IOP-M is a 34-item memory recognition test that was recently introduced by Giromini et al. (2020) to improve the classification accuracy of the IOP-29. It is administered immediately after the examinee has completed the IOP-29 under standard instructions. That is, the examinee is not warned about a subsequent memory test. As such, the IOP-M is an incidental memory test. In the two-alternative, forced-choice recognition task, the examinee is presented with two sets of words or phrases (a target and a foil), and is asked to identify the one that was part of the IOP-29 item content. According to Giromini et al. (2020), the total number of correct responses is expected to be ≥ 30 for bona fide responders unaffected by severe cognitive problems. Consequently, IOP-M scores < 30 are indicative of possible non-credible presentation.

Although it is probably more suitable to identifying cognitive feigning, the initial IOP-M research (Giromini et al., 2020) found incremental validity in detecting feigned mental disorders, including PTSD. When used together with the IOP-29, indeed, it could improve signal detection as it combines an SVT component (i.e., the IOP-29) with a PVT component (i.e., the IOP-M) in a time- and cost-effective manner. In other words, some feigners present with exaggerated memory problems as a part of depression, schizophrenia, or PTSD. The effect sizes produced by the IOP-M with feigned PTSD, however, are expected to be milder than those produced by the IOP-29, as the IOP-M essentially focuses on recognition memory.

To combine the results from the two tests, Giromini et al. (2020) recommended the following classification algorithm. If the test-taker passes both the IOP-29 and IOP-M, the presentation is highly likely to be valid; if one fails the IOP-29 (regardless of their performance on the IOP-M), then the presentation is likely to be invalid; if one passes the IOP-29 but *fails* the IOP-M, then the presentation is likely to be invalid, unless a moderate to severe cognitive impairment

can be established. This classification algorithm is referred to, here, as the "IOP-29-M."

Procedures Participants were recruited through social media and responded to the instruments using the Qualtrics platform for online surveys. After reading an informed consent form and accepting to proceed with the research, demographic information (socio-educational status and current or past diagnoses of PTSD) was collected, along with psychiatric history. Next, participants were randomly assigned to the HON or SIM condition by the platform.

Honest condition (HON) consisted of instructions to respond honestly to the instruments, followed by the administration of the Brazilian version of the PCL-5, IOP-29, and IOP-M. Conversely, the simulation condition (SIM) consisted of a detailed instruction on how to fake PTSD, which included a description of symptoms of this disorder, as well as the presentation of a vignette describing a traumatic situation, to be used as a resource for to help participants successfully simulate PTSD. After that, the participants were asked to respond to the PCL-5, IOP-29, and IOP-M pretending to be the character that suffered the traumatic event described in the vignette. At the end of the questionnaire, both groups were asked to confirm the condition to which they had been assigned, responding validation questions on whether they responded honestly of pretended to be mentally ill while taking the survey (manipulation check).

Data Analysis First, we performed a series of *t* tests to compare the average IOP-29 and IOP-M scores generated by the HON sample against those produced by the SIM. Next, we focused on AUC and classification accuracy, and examined the sensitivity and specificity of both tests, as well as for the IOP-29 and IOP-M classification algorithm described above (i.e., the IOP-29-M). To evaluate whether simulators who were classified as "credible" by both the IOP-29 and IOP-M did make an adequate effort to feign PTSD, their PCL-5 scores were inspected. Lastly, to quantify the extent to which the IOP-M has the potential to yield incremental validity, a hierarchical logistic regression was tested with group (0 = honest; 1 = simulator) as the criterion variable, the IOP-29 FDS score as the predictor variable entered in the first step, and the IOP-M score the predictor variable entered in the second step.

Results

When compared with HON, simulators scored significantly higher on the IOP-29 FDS and significantly lower on the IOP-M, as expected (Table 3). The effect size generated by the IOP-29 ($d = 2.15$) was over double than that of the

Table 3 IOP-29 and IOP-M scores of honest responders versus simulators (N = 201)

	Honest (n = 101)		Simulator (n = 100)		t	df	p	d
	M	SD	M	SD				
IOP-29 FDS	0.31	0.20	0.73	0.19	15.26	199.0	< .01	2.15
IOP-M	33.3	0.9	28.8	6.5	6.86	102.8[a]	< .01	0.97

[a]Because homoscedasticity could not be assumed, the Welch–Satterthwaite method was used to adjust degrees of freedom

IOP-M (d = 0.97). In terms of classification accuracy, AUC was 0.92 (SE = 0.02; 95% CI = 0.88 to 0.96) for the IOP-29 and 0.74 (SE = 0.04; 95% CI = 0.67 to 0.81) for the IOP-M. Thus, although both tests efficiently discriminated credible from non-credible presentations, the IOP-29 outperformed the IOP-M, as expected.

Table 4 provides some additional details on diagnostic efficiency statistics. Worth noting, at standard cutoff of FDS ≥ 0.50, the IOP-29 generated 0.87 sensitivity at 0.78 specificity, whereas a more liberal cutoff (FDS ≥ 0.30) yielded a 0.97 sensitivity. Conversely, a more conservative cutoff of FDS ≥ 0.65 yielded a specificity of 0.92. These findings are consistent with data presented in the IOP-29 manual (Viglione & Giromini, 2020). As for the IOP-M, at the suggested cutoff of < 30 (Giromini et al., 2020), it generated a notably lower overall correct classification (OCC) rate (0.69), with a sensitivity of 0.37 at a perfect specificity (1.00). Finally, when compared with using the IOP-29 alone at standard cutoff (FDS ≥ 0.50), combining the IOP-29 with the IOP-M slightly improved sensitivity (0.89 vs. 0.87), at no costs for specificity (0.78 vs. 0.78), as indicated in Giromini et al. (2020)

We then examined the extent to which simulators who looked credible on both the IOP-29 and IOP-M produced elevated PCL-5 scores. This analysis essentially served as an additional manipulation check aimed at confirming whether simulators undetected by the IOP-29-M managed to score in the impaired range. Of the 11 simulators (out of 100)

who *passed* both the IOP-29 and IOP-M, only one had a PCL-5 score below 33, whereas the other ten had PCL-5 scores ≥ 35. Importantly, the average PCL-5 score within this subgroup was 42.3 (SD = 10.5), which is almost one standard deviation above the suggested cut-off score of 33. Taken together, these data indicate that the majority of simulators who were not detected by the IOP-29-M did look like individuals affected by PTSD-related problems on the PCL-5 (i.e., they beat the validity check, while managing to appear impaired).

Lastly, we focused on whether using the IOP-M together with the IOP-29 improved classification accuracy compared with using the IOP-29 alone. As shown in Table 5, when the IOP-M was entered in the second step of a logistic regression aimed at predicting group membership (HON vs. SIM), the model significantly improved, $\mathcal{X}^2(1) = 15.9$, $p < 0.01$. Thus, the IOP-M did yield some incremental validity (OCC = 0.86 vs. OCC = 0.83), with this sample.

Discussion

Study 2 compared the IOP-29 and IOP-M scores produced by 101 honest responders against those produced by 100 feigners of PTSD. In line with previous research (Viglione & Giromini, 2020), the IOP-29 effectively discriminated credible from non-credible presentations, with a very large effect size of d = 2.15 and an excellent AUC of 0.92. Furthermore,

Table 4 Classification accuracy of IOP-29, IOP-M, and IOP-29-M

	Sensitivity	Specificity	Overall correct classification
IOP-29			
FDS ≥ .65	.69	.92	.81
FDS ≥ .50	.87	.78	.83
FDS ≥ .30	.97	.54	.76
IOP-M			
# of correct < 30	.37	1.00	.69
IOP-29-M			
Fail IOP-29 *or* IOP-M	.89	.78	.84

Table 5 Incremental validity of the IOP-M: results from logistic regression analyses

	Step 1 (IOP-29 only)	Step 2 (IOP-29 and IOP-M)
$\mathcal{X}^2$ model	132.8**	148.7**
$\Delta \mathcal{X}^2$	132.8**	15.9**
B IOP-29 (SE)	8.36 (1.07)**	7.63 (1.11)**
B IOP-M (SE)	–	−0.37 (0.14)*
Sensitivity	.87	.83
Specificity	.78	.88
Overall correct classification	.83	.86

the standard cutoff score of FDS ≥ 0.50 yielded sensitivity and specificity values of about 80% or more (specifically, sensitivity = 0.87 and specificity = 0.78). Taken together, these findings thus suggest that the Brazilian version of the IOP-29 likely is an accurate adaptation of the original IOP-29.

As for the IOP-M, consistent with previous research (Giromini et al., 2020), it produced perfect specificity, but low sensitivity (= 0.37) in the detection of feigned PTSD, at the cutoff of < 30. What is most noteworthy, however, is that the feigners who failed the IOP-M were not the same as those who failed the IOP-29. Indeed, when the IOP-M was entered after the IOP-29 in a logistic regression aimed at predicting group membership (HON vs. SIM), the model did improve. As such, one may conclude that—pending future replication—the IOP-M does have the potential to yield some incremental validity when used in combination with the IOP-29, even when assessing for psychiatric problems (and not just for feigned cognitive deficits).

The fact that the IOP-M provided useful and unique validity is a particularly important finding. Indeed, although PVTs have been typically used to assess the credibility of neuropsychological problems, recent evidence suggests that they could also be useful in evaluating the credibility of various psychiatric conditions. For instance, Fox and Vincent (2020) instructed 411 volunteers to respond to a number of SVTs and PVTs as if they were experiencing common emotional, behavioral, and cognitive symptoms of PTSD, and found that both types of validity checks provided useful and unique information. As such, they specifically recommended practitioners performing PTSD-related evaluations "to utilize more than one measure of malingering, including both PVT and SVT approaches, when PTSD is alleged" (p. 90). Along similar lines, Giromini, Barbosa et al. (2019) administered the IOP-29 and TOMM to 50 feigners of mTBI and 50 feigners of depression, and found that the TOMM provided useful and unique information not only when investigating feigned mTBI but also when examining feigned depression. In line with these and other emerging findings (Erdodi et al., 2017, 2018; Giromini et al., 2020; Tyson et al., 2018), our study thus provides some further support to the use of PVTs for the assessment not only of neuropsychological but also psychiatric conditions.

Final Remarks

This project aimed at developing and validating a Brazilian version of the IOP-29 (Viglione & Giromini, 2020). Additionally, it also sought to provide some additional evidence that the recently developed IOP-M (Giromini et al., 2020) could contribute to improving signal detection, when used in combination with the IOP-29. All in all, the results of

this investigation may be summarized as follows: Study 1 showed that IOP-29 elevations cannot be automatically attributed to previous exposures to traumatic events or destabilizing stressors. Study 2 demonstrated that feigned PTSD instead does elevate the IOP-29—and, to a lesser degree, the number of IOP-M errors too. Importantly, this research was conducted in Brazil, a country in which the IOP-29 and IOP-M have not been previously available or validated.

With regard to the IOP-29, previous research has demonstrated that it offers excellent classification accuracy regardless of whether it is used to evaluate the possible presence of PTSD (e.g., Giromini et al., 2018), depression (e.g., Giromini, Carfora Lettieri et al., 2019; Ilgunaite et al., 2020), schizophrenia (e.g., Giromini et al., 2018; Winters et al., 2020), or cognitive problems (Gegner et al., 2021; Viglione et al., 2017), and that it is similarly valid in remarkably different countries such as the USA (e.g., Viglione et al., 2019), Australia (Gegner et al., 2021), Portugal (Giromini, Barbosa et al., 2019), Lithuania (Ilgunaite et al., 2020), Italy (e.g., Giromini et al., 2018), and the UK (Winters et al., 2020). By independently replicating these findings within a Brazilian context, the current article contributes to the growing research base for using the IOP-29 in applied clinical and forensic settings. Additionally, it also provides initial evidence that the classification accuracy of the IOP-29 may generalize to Latin American countries. This finding is consistent with previous reports that the IOP-29 functions similarly across cultures and languages.

As for the IOP-M, this is the second study to independently replicate the results described by Giromini et al. (2020), and the first focused on feigned PTSD. Given that our research design did not target cognitive impairment, the sensitivity of the IOP-M was far from optimal (0.37), at its suggested cutoff of < 30. Nonetheless, study 2 did support the hypothesis that the IOP-M has the potential to further improve the already excellent classification accuracy of the IOP-29. Furthermore, it also showed that, consistent with Giromini et al. (2020), none of the 101 adult controls in study 2 was mistakenly classified as a simulator by the IOP-M (i.e., specificity was 1.00). Because failure of PVTs is more likely to occur when feigning neuropsychological rather than psychiatric symptoms (Van Dyke et al., 2013; Whiteside et al., 2020), additional research should test the incremental validity of the IOP-M with other types of feigned complaints too, such as in the case of mild traumatic brain injury presentations and similar conditions. In fact, previous research demonstrated the "domain specificity effect" (Erdodi, 2019): the similarity between the target construct of the criterion and predictor variable influences classification accuracy (Abeare, Sabelli et al., 2019; Gaasedelen et al., 2019, Rai & Erdodi, 2019; Schroeder et al., 2019).

It should be pointed out that our investigation focused on PTSD because this disorder is among the most relevant mental health conditions to psychological injury evaluations in Brazil and elsewhere (Young, 2016), and because it is particularly easy to feign (Resnick et al., 2018; see also Burges & McMillan, 2001; Lees-Haley & Dunn, 1994; Slovenko, 1994). According to Young (2016), about 70% individuals are exposed to traumatic events over their lifetime, and traumatic reactions such as PTSD occur in about 10% of cases, after traumatic exposure. From this perspective, the fact that none of our 154 firefighters from study 1 who had been exposed on average to more than six destabilizing or potentially traumatic events during the past few months had an IOP-29 FDS value above 0.61, and only six (i.e., 4%) had an FDS greater than 0.50 is particularly encouraging. On the other hand, the specificity of the IOP-29 surprisingly dropped to 0.92 (for FDS ≥ 0.65) and 0.78 (for FDS ≥ 0.50) when considering the non-clinical volunteers of study 2. Although at face value this appears to be an unexpected finding, two aspects of the study's methodology may be accounted for the observed outcome. First, unlike study 1, study 2 was conducted online. Thus, there was significantly less experimental control (no opportunity to check for comprehension, ask questions, or ensure that small, but consequential misunderstandings are promptly corrected). Second, the experimental malingering paradigm contains numerous well-known confounds, such as lack of experiential basis for the target condition (Lau et al., 2017), low motivation to produce credible impairment (i.e., demonstrate deficits, but avoid detection; Erdal et al., 2004; Hurtubise et al., 2020), or lack of compliance with instructions in general (Abeare, Messa et al., 2019; Abeare, Sabelli et al., 2019; An et al., 2012).

A limitation inherent in experimental feigning designs is that group assignment (i.e., HON or SIM) is a quasi-independent variable: the researcher only controls the instructions given to participants, not the fidelity of their execution (Rai et al., 2019). As a result, criterion contamination either in the form of controls failing to demonstrate intact performance (An et al., 2019; Roye et al., 2019) or simulators failing to demonstrate credible impairment (Abeare et al., 2020) can undermine the internal validity of the design. Potential explanations for this phenomenon include non-contingent reinforcement (i.e., individuals are rewarded for participation, not performance; An et al., 2017; Niesten et al., 2017; Tan et al., 2002), lack of knowledge and experiential basis for the neuropsychiatric condition requested to emulate (Lau et al., 2017) as well as the vastly different type and magnitude of reward for successful malingering in research versus real-life settings (Jurick et al., 2020; Lace et al., 2020; Merten & Merckelbach, 2013; Shura et al., 2016).

A few other potential limitations of this project should also be pointed out. The most obvious one is the lack of participants with established PTSD diagnoses. Indeed, although study 1 focused on individuals who had previously been exposed to various potentially traumatic/destabilizing stressors, and more than 40 of them in fact scored above the suggested cutoff for probable PTSD on the PCL-5, no test-taker had been formally evaluated for possible PTSD. As such, the specificity levels observed in this investigation might overestimate the real specificity one might find in real-life examinations involving patients with severe PTSD or complex trauma. Also, there are likely significant differences in incentives to evade detection while feigning PTSD between active duty firefighters and individuals who sustained significant injuries in a work-related accident, medical malpractice, or assault. The present findings may not generalize to the latter populations.

Along similar lines, it should be noted that study 1 included firefighters as participants because we had easy access to that specific population and for their being frequently exposed to potentially traumatic or destabilizing events. Yet, firefighter selection and training protocols have the potential to influence study participants' experiences and reports of PTSD symptoms, so that further replication studies with other populations would be beneficial. Indeed, because study 2 did not ask participants to report whether or not they had been exposed to any potentially traumatizing or destabilizing events, we could not address this issue with our study 2 data, unfortunately. Another limitation is that—as is the case for all analogue/simulation studies—the external validity of our study 2 may be questioned, given that feigning in a high-stake, real-life context likely poses different challenges and involves different mental processes from feigning within an experimental context. Although we took several methodological cautions to maximize external validity, i.e., we warned simulators "not to over-do it" and we gave them a description of typical PTSD symptoms (Rogers & Bender, 2018; Viglione et al., 2001), the "Achilles heel" of the experimental malingering paradigm is not the quality or type of instructions, but participant compliance (Abeare et al., 2020; An et al., 2017; Hurtubise et al., 2020). Therefore, future replications with real-life clinical and forensic samples would be beneficial. Additionally, as we did not administer any other SVTs or PVTs, aside from the IOP-29 and IOP-M, comparative and convergent validity could not be assessed.

Despite these limitations, our research still has the merit to be the first to investigate the IOP-29 within a Brazilian sample, and the first to provide an independent cross-validation of the IOP-M focused on feigned PTSD. Taken together, the results of the two studies described in this article further extend the growing literature on the IOP-29 and suggest that it may be used validly in various different assessment and cultural contexts. Besides, they also provide some additional support for the newly developed IOP-M as a useful add-on to improve

the already excellent classification accuracy of the IOP-29. Despite these encouraging findings, however, we would like to highlight once again that the determination of feigning or malingering should never be based on a single test alone, regardless of its apparent effectiveness. Using multiple SVTs and PVTs to evaluate the credibility of a clinical presentation is not only advisable (Boone, 2013; Fox & Vincent, 2020; Giromini et al., 2020; Larrabee, 2008; Rogers & Bender, 2018), but the emerging standard of practice (Bush et al., 2005, 2014; Chafetz, 2011; Chafetz et al., 2015; Heilbronner et al., 2009; Schutte et al., 2015).

Declarations

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Copyright Statement The rights to both Inventory of Problems-29 (IOP-29) and Inventory of Problems-Memory (IOP-M) are possessed by IOP-Test, LLC.

Conflict of Interest The seventh and eighth authors declare that they own a share in the corporate (LLC) that possesses the rights to Inventory of Problems. The other six authors declare that they have no conflict of interest to report.

References

Abeare, C. A., Hurtubise, J., Cutler, L., Sirianni, C., Brantuo, M., Makhzoun, N., & Erdodi, L. (2020). Introducing a forced choice recognition trial to the Hopkins Verbal Learning Test-Revised. *The Clinical Neuropsychologist*. https://doi.org/10.1080/13854046.2020.1779348

Abeare, C., Messa, I., Whitfield, C., Zuccato, B., Casey, J., & Erdodi, L. (2019). Performance validity in collegiate football athletes at baseline neurocognitive testing. *Journal of Head Trauma Rehabilitation, 34*(4), 20–31. https://doi.org/10.1097/HTR.0000000000000451

Abeare, C., Sabelli, A., Taylor, B., Holcomb, M., Dumitrescu, C., Kirsch, N., & Erdodi, L. (2019). The importance of demographically adjusted cutoffs: age and education bias in raw score cutoffs within the Trail Making Test. *Psychological Injury and Law, 12*(2), 170–182.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2013).The standards for educational and psychological testing DC, AERA.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders 5. American Psychiatric Association.

An, K. Y., Charles, J., Ali, S., Enache, A., Dhuga, J., & Erdodi, L. A. (2019). Re-examining performance validity cutoffs within the Complex Ideational Material and the Boston Naming Test-Short Form using an experimental malingering paradigm. *Journal or Clinical and Experimental Neuropsychology, 41*(1), 15–25.

An, K. Y., Kaploun, K., Erdodi, L. A., & Abeare, C. A. (2017). Performance validity in undergraduate research participants: a comparison of failure rates across tests and cutoffs. *The Clinical Neuropsychologist, 31*(1), 193–206.

An, K. Y., Zakzanis, K. K., & Joordens, S. (2012). Conducting research with non-clinical healthy undergraduates: does effort play a role in neuropsychological test performance? *Archives of Clinical Neuropsychology, 27,* 849–857.

Bass, C., & Halligan, P. (2014). Factitious disorders and malingering: challenges for clinical assessment and management. *The Lancet, 383,* 1422–1432.

Boone, K. B. (2013). Clinical practice of forensic neuropsychology. Guilford.

Brislin, R. W. (1980). Translation and content analysis of oral and written material HC Triandis JW Berry Eds Handbook of cross-cultural psychology Allyn & Bacon 389–444.

Burges, C., & McMillan, T. (2001). The ability of naïve participants to report symptoms of post-traumatic stress disorder. *British Journal of Clinical Psychology, 40*(2), 209–214.

Bush, S. S., Heilbronner, R. L., & Ruff, R. M. (2014). Psychological assessment of symptom and performance validity, response bias, and malingering: official position of the Association for Scientific Advancement in Psychological Injury and Law. *Psychological Injury and Law, 7*(3), 197–205.

Bush, S. S., Ruff, R. M., Troster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., & Silver, C. H. (2005). Symptom validity assessment: practice issues and medical necessity (NAN Policy and Planning Committees). *Archives of Clinical Neuropsychology, 20,* 419–426.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A. M., Kaemmer, B. (1989) Minnesota Multiphasic Personality Inventory-2 (MMPI-2): manual for administration and scoring University of Minnesota Press.

Chafetz, M. D. (2011). The psychological consultative examination for Social Security Disability. *Psychological Injury and Law, 4,* 235–244. https://doi.org/10.1007/s12207-011-9112-5

Chafetz, M. D., Williams, M. A., Ben-Porath, Y. S., Bianchini, K. J., Boone, K. B., Kirkwood, M. W., et al. (2015). Official position of the American Academy of Clinical Neuropsychology Social Security Administration policy on validity testing: guidance and recommendations for change. *The Clinical Neuropsychologist, 29*(6), 723–740.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences 2. Erlbaum.

Erdal, K. (2004). The effects of motivation, coaching, and knowledge of neuropsychology on the simulated malingering of head injury. *Archives of Clinical Neuropsychology*, *19*(1), 73–88.

Erdodi, L. A. (2019). Aggregating validity indicators: the salience of domain specificity and the indeterminate range in multivariate models of performance validity assessment. *Applied Neuropsychology: Adult, 26*(2), 155–172.

Erdodi, L. A., Abeare, C. A., Medoff, B., Seke, K. R., Sagar, S., & Kirsch, N. L. (2018). A single error is one too many: The Forced Choice Recognition trial on the CVLT-II as a measure of performance validity in adults with TBI. *Archives of Clinical Neuropsychology, 33*(7), 845–860. https://doi.org/10.1093/arclin/acx110

Erdodi, L. A., Seke, K. R., Shahein, A., Tyson, B. T., Sagar, S., & Roth, R. M. (2017). Low scores on the Grooved Pegboard Test are associated with invalid responding and psychiatric symptoms. *Psychology and Neuroscience, 10*(3), 325–344. https://doi.org/10.1037/pne0000103

Erdodi, L. A., Taylor, B., Sabelli, A., Malleck, M., Kirsch, N. L., & Abeare, C. A. (2019). Demographically adjusted validity cutoffs in the Finger Tapping Test are superior to raw score cutoffs. *Psychological Injury and Law, 12*(2), 113–126.

Fox, K. A., & Vincent, J. P. (2020). Types of malingering in PTSD: evidence from a psychological injury paradigm. *Psychological Injury and Law, 13,* 90–104.

Gaasedelen, O. J., Whiteside, D. M., Altmaier, E., Welch, C., & Basso, M. R. (2019). The construction and the initial validation of the

Cognitive Bias Scale for the Personality Assessment Inventory. *The Clinical Neuropsychologist, 33*(8), 1467–1484.

Gegner, J., Erdodi, L. A., Giromini, L., Viglione, D. J., Bosi, J., & Brusadelli, E. (2021). An Australian study on feigned mTBI using the Inventory of Problems-29 (IOP-29), its Memory Module (IOP-M), and the Rey Fifteen Item Test (FIT). Applied Neuropsychology: Adult. Advance online publication. https://doi.org/10.1080/23279095.2020.18643752020.1864375

Geisinger, K. F. (2003). Testing and assessment in cross-cultural psychology JR Graham JA Naglieri Eds Handbook of psychology: Assessment psychology. John Wiley & Sons Inc. 95–117

Giromini, L., Viglione, D. J., Pignolo, C., & Zennaro, A. (2018). A clinical comparison, simulation study testing the validity of SIMS and IOP-29 with an Italian sample. *Psychological Injury and Law, 11*(4), 340–350.

Giromini, L., Barbosa, F., Coga, G., Azeredo, A., Viglione, D. J., & Zennaro, A. (2019). Using the Inventory of Problems-29 (IOP-29) with the Test of Memory Malingering (TOMM) in symptom validity assessment: A study with a Portuguese sample of experimental feigners. *Applied Neuropsychology: Adult*, [Epub ahead of print], https://doi.org/10.1080/23279095.2019.1570929

Giromini, L., Carfora Lettieri, S., Zizolfi, S., Zizolfi, D., Viglione, D.J., Brusadelli, E., Perfetti, B., di Carlo, D.A., & Zennaro, A. (2019). Beyond rare-symptoms endorsement: a clinical comparison simulation study using the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) with the Inventory of Problems (IOP-29). *Psychological Injury and Law,* [Epub ahead of print], https://doi.org/10.1007/s12207-019-09357-7

Giromini, L., Viglione, D. J., Pignolo, C., & Zennaro, A. (2019a). An Inventory of Problems (IOP-29) sensitivity study investigating feigning of four different symptom presentations via malingering experimental paradigm. *Journal of Personality Assessment,* [Epub ahead of print]. https://doi.org/10.1080/00223891.2019.1566914

Giromini, L., Viglione, D. J., Pignolo, C., & Zennaro, A. (2019b). An Inventory of Problems (IOP-29) study on random responding using experimental feigners, honest controls, and computer-generated data. *Journal of Personality Assessment*, [Epub ahead of print]. https://doi.org/10.1080/00223891.2019.1639188

Giromini, L., Viglione, D. J., Zennaro, A., Maffei, A., & Erdodi, L. (2020). SVT meets PVT: development and initial validation of the Inventory of Problems-Memory (IOP-M). *Psychological Injury and Law [Epub ahead of print]*, https://doi.org/10.1007/s12207-020-09385-8

Green, R. L. (1991). MMPI-2/MMPI: an interpretative manual. Allyn & Bacon.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Participants, C. (2009). American Academy of Clinical Neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist, 23,* 1093–1129. https://doi.org/10.1080/13854040903155063

Hurtubise, J., Baher, T., Messa, I., Cutler, L., Shahein, A., Hastings, M., Carignan-Querqui, M., & Erdodi, L. (2020). Verbal fluency and digit span variables as performance validity indicators in experimentally induced malingering and real world patients with TBI Applied Neuropsychology. Child 1–18. https://doi.org/10.1080/21622965.2020.17194092020.1719409

Ilgunaite, G., Giromini, L., Bosi, J., Viglione, D. J., & Zennaro, A. (2020). A clinical comparison simulation study using the Inventory of Problems-29 (IOP-29) with the Center for Epidemiologic Studies Depression Scale (CES-D) in Lithuania. *Applied Neuropsychology: Adult,* [Epub ahead of print]. https://doi.org/10.1080/23279095.2020.17255182020.1725518

Jurick, S. M., Crocker, L. D., Merritt, V. C., Hoffman, S. N., Keller, A. V., Eglit, G. M. L., et al. (2020). Psychological symptoms and rates of performance validity improve following trauma-focused treatment in veterans with PTSD and history of mild-to-moderate TBI. *Journal of International Neuropsychological Society, 26*(1), 108–118.

Lace, J. W., Grant, A. F., Kosky, K. M., Teague, C. L., Lowell, K. T., & Gfeller, J. D. (2020). Identifying novel embedded performance validity test formulas within the Repeatable Battery for the Assessment of Neuropsychological Status: a simulation study. *Psychological Injury and Law, 13,* 303–315.

Larrabee, G. J. (2008). Aggregation across multiple indicators improves the detection of malingering: relationship to likelihood ratios. *The Clinical Neuropsychologist, 22,* 666–679.

Lau, L., Basso, M. R., Estevis, E., Miller, A., Whiteside, D. M., Combs, D., & Arentsen, T. J. (2017). Detecting coached neuropsychological dysfunction: a simulation experiment regarding mild traumatic brain injury. *The Clinical Neuropsychologist, 31*(8), 1412–1431.

Lees-Haley, P. R., & Dunn, J. T. (1994). The ability of naive subjects to report symptoms of mild brain injury, post-traumatic stress disorder, major depression and generalized anxiety disorder. *Journal of Clinical Psychology, 50*(2), 252–256.

Merten, T., & Merckelbach, H. (2013). Symptom validity in somatoform and dissociative disorders: a critical review. *Psychological Injury and Law, 6*(2), 122–137.

Merten, T., Thies, E., Schneider, K., & Stevens, A. (2009). Symptom validity testing in claimants with alleged posttraumatic stress disorder: comparing the Morel emotional numbing test, the structure inventory of malingered symptomatology, and the word memory test. *Psychological Injury and Law, 2,* 284–293.

Morey, L. C. (2007). Personality Assessment Inventory (PAI). Professional manual 2 Psychological Assessment Resources.

Niesten, I. J. M., Muller, W., Merckelbach, H., Dandachi-FitzGerald, B., & Jelicic, M. (2017). Moral reminders do not reduce symptom over-reporting tendencies. *Psychological Injury and Law, 10,* 368–384.

Osório , F.L. , Silva , T.D.A. , Santos , R.G. , Chagas , M.H.N. , Chagas , N.M. S. , Sanches , R.F. , & Crippa , J.A.S. (2017). Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): transcultural adaptation of the Brazilian version. *Archives of Clinical Psychiatry (São Paulo)*, *44*(1), 10-19. https://doi.org/10.1590/0101-60830000000107

Pereira-Lima, K., Loureiro, S. R., Bolsoni, L. M., Apolinario da Silva, T. D., & Osório, F. L. (2019). Psychometric properties and diagnostic utility of a Brazilian version of the PCL-5 (complete and abbreviated versions). *European Journal of Psychotraumatology*, *10*(1), 1581020. https://doi.org/10.1080/20008198.2019.1581020

Rai, J., An, K. Y., Charles, J., Ali, S., & Erdodi, L. A. (2019). Introducing a forced choice recognition trial to the Rey Complex Figure Test. *Psychology and Neuroscience., 12*(4), 451–472.

Resnick, P. J., West, S., Wooley, C. N. (2018). The malingering of posttraumatic disorders R Rogers SD Bender Eds Clinical assessment of malingering and deception Guilford Press 188–211.

Rogers, R., Bender, D. (2018) Clinical assessment of malingering and deception. Guilford Press.

Roma, P., Giromini, L., Burla, F., Ferracuti, S., Viglione, D. J., & Mazza, C. (2020). Ecological validity of the Inventory of Problems-29 (IOP-29): an Italian study of court-ordered, psychological injury evaluations using the Structured Inventory of Malingered Symptomatology (SIMS) as criterion variable. *Psychological Injury and Law, 13,* 57–65.

Roye, S., Calamia, M., Bernstein, J. P., De Vito, A. N., & Hill, B. D. (2019). A multi-study examination of performance validity in undergraduate research participants. *The Clinical Neuropsychologist, 33*(6), 1138–1155.

Schroeder, R. W., Martin, P. K., Heindrichs, R. J., & Baade, L. E. (2019). Research methods in performance validity testing studies: criterion grouping approach impacts study outcomes. *The Clinical Neuropsychologist, 33*(3), 466–477.

Schutte, C., Axelrod, B. N., & Montoya, E. (2015). Making sure neuropsychological data are meaningful: use of performance validity

testing in medicolegal and clinical contexts. *Psychological Injury and Law, 8*(2), 100–105.

Shura, R. D., Miskey, H. M., Rowland, J. A., Yoash-Gatz, R. E., & Denning, J. H. (2016). Embedded performance validity measures with postdeployment veterans: cross-validation and efficiency with multiple measures. *Applied Neuropsychology: Adult, 23,* 94–104.

Slovenko, R. (1994). Legal aspects of post-traumatic stress disorder. *Psychiatric Clinics of North America, 17,* 439–446.

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy on Psychiatry and Law, 25,* 180–183.

Souto, L. S., de Oliveira, L. M., & Haag, C. K. (2012). Transtorno de estresse pós-traumático decorrente de acidente de trabalho: implicações psicológicas, socioeconômicas e jurídicas. *Estudos de Psicologia, 17*(2), 329–336.

Tan, J. E., Slick, D. J., Strauss, E., & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist, 16*(4), 495–5050.

Tombaugh, T. N. (1996). Test of Memory Malingering (TOMM) Multi Health Systems.

Tyson, B. T., Baker, S., Greenacre, M., Kent J, K., Lichtenstein, J. D., Sabelli, A., & Erdodi, L. A. (2018) Differentiating epilepsy from psychogenic nonepileptic seizures using neuropsychological test data Epilepsy & Behavior *87,* 39–45.

Van de Vijver, F., & Hambleton, R. K. (1996). Translating tests. *European Psychologist, 1*(2), 89–99.

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.

Viglione, D. J., Giromini, L., Landis, P., McCullaugh, J. M., Pizitz, T. D., O'Brien, S., et al. (2019). Development and Validation of the False Disorder Score: The Focal Scale of the Inventory of Problems. *Journal of Personality Assessment, 101,* 653–661.

Viglione D. J., Giromini, L. (2020). Inventory of Problems-29: Professional Manual IOP-Test LLC.

Viglione, D. J., Wright, D., Dizon, N. T., Moynihan, J. E., DuPuis, S., & Pizitz, T. D. (2001). Evading detection on the MMPI–2: does caution produce more realistic patterns of responding? *Assessment, 8,* 237–250.

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). The PTSD Checklist for DSM–5 (PCL-5). Retrieved from http://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp

Whiteside, D. M., Hunt, I., Choate, A., Caraher, K., & Basso, M. R. (2020). Stratified performance on the Test of Memory Malingering (TOMM) is associated with differential responding on the Personality Assessment Inventory (PAI). *Journal of Clinical and Experimental Neuropsychology, 42*(2), 131–141.

Widows, M. R., Smith, G. P. (2005). SIMS-Structured Inventory of Malingered Symptomatology. Professional manual. Psychological Assessment Resources.

Winters, C. L., Giromini, L., Crawford, T. J., Ales, F., Viglione, D. J., & Warmelink, L. (2020). An Inventory of Problems–29 (IOP–29) study investigating feigned schizophrenia and random responding in a British community sample. *Psychiatry, Psychology and Law,* [Epub ahead of print]. https://doi.org/10.1080/13218719.2020.17677202020.1767720

van Dyke, S. A., Millis, S. R., Axelrod, B. N., & Hanks, R. A. (2013). Assessing effort: differentiating performance and symptom validity. *The Clinical Neuropsychologist, 27*(8), 1234–1246.

World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA. 310 (20),* 2191–2194.

Young, G. (2016). PTSD in Court I: Introducing PTSD for Court. *International Journal of Law and Psychiatry, 49,* 238–258.

Young, G. (2017a). PTSD in Court II. Risk factors, endophenotypes, and biological underpinnings. *International Journal of Law and Psychiatry, 51,* 1–21.

Young, G. (2017b). PTSD in Court III: malingering, assessment, and the law. *International Journal of Law and Psychiatry, 52,* 81–102.