IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Detection of Feigning of Different Symptom Presentations With the PAI and IOP-29

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1823277 since 2022-07-01T09:26:57Z

Published version:

DOI:10.1177/10731911211061282

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

15 October 2023

Assessment

Detection of Feigning of Different Symptom Presentations with the PAI and IOP-29

Journal:	*Assessment*
Manuscript ID	ASMNT-20-0549.R2
Manuscript Type:	Original Research Article
Keywords:	PTSD, depression, schizophrenia, non-clinical, inmates, malingering

SCHOLARONE™
Manuscripts

Detection of Feigning of Different Symptom Presentations with the PAI and IOP-29

For Peer Review

Abstract

This study examined the effectiveness of the negative distortion measures from the Personality Assessment Inventory (PAI) and Inventory of Problems-29 (IOP-29), by investigating data from a community and a forensic sample, across three different symptom presentations (i.e., feigned depression, PTSD, and schizophrenia). The final sample consisted of 513 community-based individuals and 288 inmates (total N = 801); all were administered the PAI and the IOP-29 in an honest or feigning conditions. Statistical analyses compared the average scores of each measure by symptom presentation and data source (i.e., community vs. forensic sample), and evaluated diagnostic efficiency statistics. Results suggest that the PAI Negative Impression Management scale and the IOP-29 are the most effective measures across all symptom presentations, whereas the PAI Malingering Index and Rogers Discriminant Function generated less optimal results, especially when considering feigned PTSD. Practical implications are discussed.

Keywords: PTSD; depression; schizophrenia; non-clinical; inmates; malingering

For Peer Review

During the past few decades, a variety of instruments and scales have been developed to assist practitioners in assessing the credibility of symptom presentations in both clinical and forensic populations. These measures vary in format (e.g., self-report scales, structured interviews, or performance-based tests), purpose (e.g., screening *vs.* comprehensive evaluation), and effort level required of the test-taker. Although numerous studies support the effectiveness of various symptom validity tests (SVTs) in the detection of feigning of mental health problems, most SVTs are known to perform differently across symptom presentations, situations, and contexts. As such, it is important for practitioners to be aware of these subtle caveats so as to formulate more accurate interpretations. To that goal, the current study investigated the effectiveness of the negative distortion measures of the broad-band Personality Assessment Inventory (PAI; Morey, 1991; 2007) and of a brief, stand-alone SVT, the Inventory of Problems-29 (IOP-29; Viglione, Giromini, & Landis, 2017; Viglione & Giromini, 2020), across different symptom presentations in both a community and a forensic sample.

The PAI provides three measures for detecting over-reporting of psychopathology: the Negative Impression Management (NIM; Morey, 1991), the Malingering Index (MAL; Morey, 1996), and the Rogers Discriminant Function (RDF; Rogers et al., 1996). In the most recent meta-analysis on their validity, Hawes and Boccaccini (2009) reported that with NIM $\geq$ 81, sensitivity (Se) and specificity (Sp) were .73 and .83 respectively, whereas NIM $\geq$ 110 produced Se = .33 and Sp = .98. As for MAL, a cut score of $\geq$ 3 produced Se = .58 and Sp = .86, whereas MAL $\geq$ 4 yielded Se = .35 and Sp = 1.00. These authors did not examine RDF cut scores' performance because only a few studies reported effects for the same cut score. Finally, they found that the PAI negative distortion measures are likely to perform differently depending on the symptom presentation under consideration. Indeed, all measures under

For Peer Review

investigation seemed to be more efficient in detecting feigned severe mental disorders[1] (Cohen's *d* ranged from 1.82 to 2.32) rather than feigned mood or anxiety disorders (Cohen's *d* values ranged from 0.90 to 1.25).

A few subsequent studies further investigated the performance of the PAI negative distortion measures in detecting feigned PTSD and depression. These newer studies found that when comparing PAI scores of a sample of healthy individuals instructed to feign PTSD symptoms with a genuine patient group, the PAI negative distortion scales showed *d* effect sizes ranging from 0.68 to 1.24 for NIM, from 0.73 to 1.28 for MAL, and from 0.60 to 0.82 for RDF (Russel & Morey, 2019; Thomas et al., 2012). When using healthy responders as honest controls instead of bona fide patients, Cohen's *d* effect sizes were: $d = 1.74$ for NIM, $d = 1.73$ for MAL, and $d = 0.72$ for RDF when considering feigning of PTSD; and $d = 2.28$ for NIM, $d = 2.23$ for MAL, and $d = 1.61$ for RDF when considering feigning of depression (Lange et al., 2010).

Finally, with regard to cut scores, NIM ≥ 80 produced Se = .33 and Sp = 1.00, MAL ≥ 3 produced Se = .53 and Sp = 1.00, and RDF ≥ 0.124 produced Se = .40 and Sp = .80 for feigned PTSD (Lange et al., 2010), whereas sensitivity values were higher for NIM and RDF when considering feigned depression (NIM ≥ 80: Se = .64, Sp = 1.00; MAL ≥ 3: Se = .50, Sp = 1.00; RDF ≥ 0.124: Se = .86, Sp = .70; Lange et al., 2010). Thus, additional research would be beneficial to reach an agreement on what cutoffs to use in different evaluation contexts and to shed light on the different results found by different authors in the previous literature.

Another under investigated topic concerns the extent to which the PAI validity scales and indices would preserve their effectiveness in forensic and correctional contexts

[1] In their meta-analysis, Hawes and Boccaccini (2009) consider severe mental disorders all those disorders related to the psychotic spectrum. They stated: "In the more severe mental disorder grouping, we included studies in which participants were feigning psychosis, attempting to receive inpatient psychiatric treatment, or attempting to be found incompetent to stand trial or not guilty by reason of insanity." (Hawes, & Boccaccini, 2009, p. 116).

For Peer Review

(Duellman & Bowers, 2004; Edens et al., 2001; Gaines et al., 2012; Morey & Quigley, 2002; Morey et al., 2007). This would be important to know, because Pierre and colleagues (2004) found that almost 30% of inmates seen in psychiatric services of a large American prison report having malingered symptoms in order to obtain psychoactive drugs. Moreover, feigning a mental disorder can be especially appealing to inmates charged with serious crimes, as evidence of mental illness can result in mitigated sentencing, including avoiding capital punishment (Resnick, 1999). In one of the few PAI studies on this subject, Edens et al. (2007) evaluated the classification accuracy of the PAI validity scales and indexes in a psychiatric unit and general population of prison inmates. More specifically, in an inmate subsample of experimental simulators, NIM ≥ 77T produced Se = .76, MAL ≥ 5 produced Se = .52, and RDF ≥ 70T produced Se = .90, whereas the sensitivity of these same measures was remarkably lower in a subsample of real-life suspected malingerers (NIM ≥ 77T: Se = .58; MAL ≥ 5: Se = .46; RDF ≥ 70T: Se = .39). In a control sample, the same three scales showed Sp = .97, .10, and .80, respectively. Finally, in a patient sample, NIM, MAL and RDF yielded Sp = .50, .73, and .73, respectively. The Overall Correct Classification (OCC) was .70 (AUC = .75) for NIM scale, .69 (AUC = .76) for MAL, and .70 (AUC = .77) for RDF. Given the paucity of data on this topic, yet, additional research would be beneficial also to appreciate the extent to which the effectiveness of the PAI validity scales vary as a function of the context (non-clinical *vs.* forensic sample) in which they are investigated.

As for the IOP-29, it was recently developed by Viglione et al. (2017) as a brief, self-administered measure to evaluate the credibility of various symptom presentations. Its key measure, the False Disorder probability Score (FDS), reflects the likelihood of finding a given IOP-29 profile within a group of bona fide patients *vs.* a group of feigners, so that higher scores indicate a lower credibility of the presented complaints, whereas lower scores suggest that the symptom presentation is more credible. In their introductory article, Viglione

For Peer Review

et al. (2017) inspected data from various samples of bona fide patients and experimental simulators and found that with an a-priori established cut-off score of FDS ≥ .50, sensitivity and specificity were about 80% (Se = .81; Sp = .79). According to the IOP-29 manual (Viglione & Giromini, 2020), a cut score of FDS ≥ .50 is thus described as the 'standard' IOP-29 cut score in that it minimizes both false positive and false negative classifications. When using the IOP-29 for screening purposes, however, sensitivity might be more important than specificity, because only positive classifications will be followed up with additional testing. In these contexts, using FDS cut scores of ≥ .15 and ≥ .30 might thus be preferable, so as to seek for sensitivity levels of 95% and 90%, at specificity levels of about 30% and 60%, respectively. Conversely, in high-stakes forensic evaluations, where specificity is typically more important than sensitivity, the suggested cut scores would be FDS ≥ .65 and FDS ≥ .70, in that they presumably offer specificity levels of about 90% and 95%, at sensitivity of about 70% and 65%, respectively.

The first independent clinical comparison, simulation study reported after the first publication on the IOP-29 (Viglione et al., 2017) showed that similar diagnostic accuracy results were found for feigning of both mild psychopathology, such as anxiety, depression, and/or trauma related symptomatology (Se = .81, Sp = .83), and severe psychopathology, such as psychosis (Se = .82, Sp = .81) when applying the same cut score of FDS ≥ .50 (Giromini et al., 2018). Subsequently, a series of additional simulation studies were conducted to evaluate the classification accuracy of the IOP-29 in detecting feigning of different diagnostic categories, i.e., depression, PTSD, and schizophrenia. Comparing the IOP-29 scores of feigners with healthy, honest responders, Cohen's *d* was 2.50 (AUC = .94) for the PTSD subgroup, 3.16 (AUC = .96) for the schizophrenia subgroup, and 4.32 (AUC = .99) for the depression subgroup in a study by Giromini, Viglione et al. (2020). Conversely, when clinical groups were used to contrast the IOP-29 results against those of experimental

For Peer Review

feigners, Cohen's *d* ranged between 1.80 and 1.95 (AUC between .89 and .92) for psychosis-related presentations (Giromini et al., 2018; Viglione et al., 2017) and between 1.64 and 3.31 (AUC = .88-.98) for depression (Ilgunaite et al., 2020; Giromini, Carfora Lettieri, et al., 2019). Finally, when considering a cut score of FDS ≥ .50, sensitivity values ranged from .70 to .94 for feigned schizophrenia (Banovic et al., 2021; Giromini, Viglione et al., 2020; Winters et al., 2020), from .86 to .87 for feigned PTSD (Carvalho et al., 2021; Giromini, Viglione et al., 2020), and from .75 to .96 for feigned depression (Giromini, Barbosa, et al., 2020; Giromini, Carfora Lettieri, et al., 2019; Giromini, et al., 2020; Ilgunaite et al., 2020; Viglione, et al. 2017).

In their initial validation paper, Viglione et al. (2017) also asked 128 adult volunteer offenders on community based county and federal probation to respond genuinely or to simulate mental health symptoms. The IOP-29 showed a sensitivity of .72, a specificity of 1.00, an OCC of .86, and an AUC of .94. Although this first study was conducted in a specific forensic context, it was a simulation study, nonetheless. Another study that evaluated the ecological validity of the IOP-29 inspected data from a sample of 74 individuals evaluated in the context of a lawsuit involving psychological injury (Roma et al., 2019). In that study, Roma et al. (2019) classified participants into two groups (i.e., credible *vs.* non-credible presentations) based on the scores of the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997; Widows & Smith, 2005). To do so, they used two different criteria, i.e., criterion 1 used SIMS Total < 17 *vs.* SIMS Total ≥ 17; criterion 2 used SIMS Total < 25 *vs.* SIMS Total ≥ 25. The first criterion produced 32 cases with a non-credible symptom presentation and 43 individuals below the cut-off: with FDS ≥ .50, Se = .81, Sp = .98, OCC = .91, *d* = 2.39, AUC = .98. The second criterion produced 20 non-credible cases and 41 credible cases: Se = .85, Sp = .98, OCC = .93, *d* = 3.59, AUC = .99. Although these findings are encouraging, the IOP-29 is relatively new, so that additional

For Peer Review

research on its effectiveness across different contexts and symptom presentations would be beneficial, too.

This Study

Currently, practitioners might find it challenging to decide which PAI cut scores would be more appropriate to use in different evaluation settings. Indeed, available literature differs remarkably with regards to the classification accuracy of the PAI negative distortion measures (i.e., NIM, MAL, and RDF) across disorders and contexts. Besides, all aforementioned PAI studies are characterized by a multiplicity of study designs and procedures. For instance, with regards to experimental simulation/analogue studies, the description of symptoms given to the simulators, the instructions on how to feign in a successful manner, and the warning to not exaggerate the presented symptoms when faking mental illness were very different across studies. As such, new studies comparing different samples and the feigning of different disorders using the same procedures are needed.

With regard to the IOP-29, even though its research foundation is growing rapidly, to date very few studies included test-takers recruited in a forensic context. Besides, except for Viglione et al.'s (2017) introductory article, no study has yet informed on how the classification accuracy of the IOP-29 compares to that of the more consolidated and widely accepted, validity scales of the PAI. More broadly, as the IOP-29 is relatively new, more research on its concurrent validity would be beneficial.

To fill these literature gaps, the aim of the current study was to test the diagnostic accuracy of the PAI and IOP-29 to detect feigning of three different categories of symptom presentations, i.e., depression, PTSD, and schizophrenia, in both a community and a forensic sample. More specifically, we intended to inspect the extent to which the sensitivity of these instruments would vary as a function of the disorder(s) being feigned, and the context in which these tests are administered (i.e., non-clinical *vs.* correctional settings). Below we

For Peer Review

report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. Interested researchers may contact the first or corresponding author to receive our Supplementary Materials and/or the data set used for the statistical analyses.

Method

Participants

Simulation studies vary greatly in sample sizes, typically ranging from a minimum of 20-30 cases to a few hundred cases per group (e.g., see Rogers et al., 2003, pp. 164-166). For this project, we aimed to reach a minimum of 50 cases per each subgroup, anticipating that about 20% to 30% could yield invalid data. However, when applicable (e.g., if the site allowed us to do so, if the student in charge of data collection could find a larger group of volunteers, etc.), we attempted to extend this number to about 150, so as to allow more stable statistical estimation. Ultimately, a total of 1,039 participants were recruited for this study, 608 individuals were from a community sample and 458 were inmates. After excluding invalid protocols (see the Procedure section for a detailed description), the final sample consisted of 801 participants, 513 community-based individuals (community sample) and 288 inmates (forensic sample). Experimental simulators included 139 individuals instructed to feign schizophrenia (SCZ Feigner), 150 individuals instructed to feign PTSD (PTSD Feigner), and 178 individuals instructed to feign depression (DEP Feigner). Honest responders (HON) were 334 (Table 1).

The community sample was comprised of 513 non-clinical participants ranging in age from 18 to 85 years old (M = 37.3, SD = 14.3). Almost all were Italians (99%), except for 3 participants who were non-Italian Europeans. In this sample, there were significant differences between the honest and the three simulator groups on age ($F_{(3,533)} = 8.015$, $p <$ .001), relationship status ($X^2 = 25.13$, $p < .001$), and gender ($X^2 = 23.84$, $p < .001$). More specifically, looking at standardized residuals for the chi square analyses and at the

For Peer Review

Bonferroni post-hoc analyses for the ANOVA, when compared to honest responders and feigners of depression, feigners of schizophrenia were younger, less frequently in a relationship, and more frequently male-gendered. Also, feigners of depression included a higher proportion of women.

The participants in the forensic sample were recruited from eight different prisons in four Italian regions. All the 288 participants were men, ranging in age from 18 to 81 years old (M = 40.9, SD = 12.7). The great majority (89.9%) were Italians, whereas 5.9% had a European citizenship different from Italian, 3.1% were Africans, 1.0% were South Americans, and only 1 participant (0.3%) was Asian. One participant did not provide this information. In this forensic sample, no differences were found for age, relationship status, and education level between honest responders and the three groups of simulators.

Measures

Personality Assessment Inventory (PAI; Morey, 1991, 2007)

The PAI is a 344-item self-report personality inventory aimed at measuring psychopathology that provides clinically useful information about an array of critical client variables in professional settings. It includes four Validity scales, eleven Clinical scales, five Treatment scales, and two Interpersonal scales. To cover all contents investigated by the multidimensional PAI scales, Morey created three to four conceptual subscales for nine Clinical scales, and for one Treatment scale (Morey, 1991, 2007). He (Morey, 1996) also developed a specialized index, the Mean Clinical Elevation (MCE), which is the mean score (i.e., the average) of the clinical scales. High (> 60 T) MCE scores provide an excellent measure of global distress, psychiatric severity, and generalized pathology. Recently, an Italian version of the PAI was published (Zennaro et al., 2015) and its psychometric properties were carefully examined (Pignolo et al., 2018). That same Italian version was used in this study.

For Peer Review

Inventory of Problems-29 (IOP-29; Viglione et al., 2017; Viglione & Giromini, 2020)

The IOP-29 was recently introduced to assist practitioners assessing fabrication, feigning, or exaggeration of psychiatric or cognitive complaints. It is a self-administered test that includes 26 questions about a range of emotional, cognitive, and social experiences and 3 items dealing with simple mathematical or logical problems. Rather than focusing on rare-symptom endorsement, it addresses the test-takers' subjective experience regarding their ability to cope with their symptoms and problems. Differently from most popular tests (e.g., the MMPI instruments), the IOP-29 offers three possible response options, i.e., "True", "False", and "Doesn't make sense". The latter option was introduced to possibly detect resistance to the evaluation context and/or feigned confusion or cognitive deficits (Rogers, 2008). The chief feigning measure of the IOP-29 is the False Disorder probability Score (FDS), an easy-to-use probability score, which indicates whether a given IOP-29 protocol is more similar to that of a feigner or of a genuine patient.

Procedure

For both the community and forensic samples, the research project was approved by the University Ethical Committee. Participants from the community were recruited through advertising of the study in community settings and via snowball sampling. They were then asked to go to the university lab to complete the study. As for the inmates, we asked the collaboration of different prisons in Italy. Once the study was approved by the institution, it was advertised among inmates by the personnel. Participation in the study was voluntary and every participant had to sign an informed consent form prior to take part to the study. Participants were administered the IOP-29 and PAI in counterbalanced order, along with a socio-demographic form. In the forensic sample, we administered the tests in groups, whereas, in the community sample, we administered the tests individually. Moreover, we administered the paper-and-pencil version of the PAI in both samples; in contrast, we

For Peer Review

administered the paper-and-pencil version of the IOP-29 in prison, whereas we used both the in-person, paper-and-pencil and computerized versions of the IOP-29 in the community sample. It should be noted, however, that the administration format is known to have no influence on IOP-29 testing results (Giromini et al., 2021).

Participants in the community sample were assigned to the control or simulation condition randomly. As for the forensic settings, the assignment to one of the two conditions was randomized by prison; however, because we administered the tests in groups, each group received the same set of instructions (i.e., all participants in a given group were to take to the tests either honestly or feigning). In the control condition, participants were asked to take the tests under standard instructions. In the experimental condition, feigners received a vignette and a list of symptoms, and they were asked to imagine themselves in a situation in which they would wish to feign a certain mental disorder. Three vignettes were used in both samples: one for feigning schizophrenia, one for feigning depression, and one for feigning PTSD symptoms (Supplementary Materials). All three vignettes were derived from previous research (Viglione et al., 2017) and intended to facilitate feigning (Rogers, & Gillard, 2011; Viglione et al., 2001) and therefore provided a realistic and contextualized means of feigning. Feigners were asked to take the tests presenting symptoms according to the vignette (i.e., feigning schizophrenia, depression, or PTSD symptoms). Feigners were also instructed to respond in a manner that would convince the examiner, i.e., without over-exaggerating so as not to be easily detected as feigners. Finally, participants were informed that if they could produce test results that would look like those of a true schizophrenic, depressed, or PTSD patient, they would receive a small compensation (amazon card for community-based simulators and games for prison for inmate simulators).

Participants were not included in the final sample if they had more than 18 items missing on the PAI or more than 3 missing items on the IOP-29, or if they had elevated the

For Peer Review

PAI ICN (Inconsistency) scale (i.e., ICN ≥ 73T; Boccaccini et al., 2006). Additionally, to ensure that participants included in the feigning groups did make an effort to look mentally ill, we excluded feigners who did not produce any meaningful clinical elevations on the PAI as detected by a MCE < 60T. This last step notably reduced the available data for the forensic group in that a conspicuous number of participants failed to fulfill this manipulation check. Table 2 shows the frequency of invalid protocols by group.

Statistical Analyses

Before analyzing the data, we evaluated the normality of the distribution of the variables under investigation and found that none departed substantially from normality (i.e., skewness > 2 and kurtosis > 7; West et al., 1995). To examine the performance of the PAI and IOP-29 negative distortion measures across symptom presentations and data source, we computed a series of ANOVAs with the NIM, MAL, RDF, and IOP-29 FDS, one at a time, as dependent variables, and symptom presentation (i.e., honest group and schizophrenia, PTSD, and depression feigning groups) and sample (i.e., community and forensic) as between-subject factors. Consistent with Rogers et al. (2003), because simulation studies are known to produce substantial effect sizes, we have adopted the following descriptive terms based on Cohen's *d*: ≥ .75 for "moderate," ≥ 1.25 for "large," and ≥ 1.75 for "very large."

Next, we computed a series of point bi-serial correlations of the PAI and IOP-29 scores to group membership (0 = honest responder; 1 = simulator) and examined AUC. Finally, we analyzed diagnostic efficiency statistics. As for the PAI, we considered the cut scores suggested by Hawes and Boccaccini (2009): NIM ≥ 81T, NIM ≥ 110T, MAL ≥ 3, MAL ≥ 4, and RDF ≥ 0. As for the IOP-29, we considered the cut scores suggested by Giromini et al. (2018) and IOP-29 manual (Viglione & Giromini, 2020): ≥ .15, ≥ .30, ≥ .50, ≥ .65 and ≥ .70.

For Peer Review

With regard to diagnostic efficiency statistics, it should be pointed out that from an applied and clinical perspective, knowing the positive (PPP) and negative (NPP) predictive power of investigated measures would be particularly useful, as forensic assessors performing symptom validity assessment typically make their determinations based on whether their evaluee(s) did or did not score above a given cut score. However, differently from Se and Sp, PPP and NPP highly depend on the base rate of the condition being tested (Meehl & Rosen, 1955), and no universally accepted estimate of malingering prevalence in forensic assessments exists. Available estimates indeed vary dramatically from one study to another (Rogers, 2008), mainly because different authors use different approaches and terms to define and investigate the phenomenon. For instance, different authors use different malingering attribution rules (e.g., based on clinical impression, expert rating using specified criteria, use of multiple tests, etc.), different decision rules (e.g., based on a score in a single instrument versus multiple scores on a whole set of instruments, etc.), and so forth (Merten & Merckelbach, 2020). Nevertheless, Young (2015, 2017, 2019) has recently made a strong case for considering 15 +/ 15% a reasonable estimate for the prevalence of malingering in most forensic assessments. As such, when calculating our diagnostic efficiency statistics, we applied Streiner's (2003) formulas and calculated PPP and NPP values by considering a base rate of 15%.

For Peer Review

Results

Descriptive statistics for all variables and subgroups under investigation are reported in Table 3. Results from the ANOVAs revealed that the interaction effect of symptom presentation by sample was statistically significant for all scales (i.e., PAI NIM: $F(3,793) = 10.60$, $p < .001$; PAI MAL: $F(3,793) = 6.50$, $p < .001$; IOP-29 FDS: $F(3,793) = 8.48$, $p < .001$), except for the PAI RDF ($F(3,793) = 1.71$, $p = .163$). Thus, separate analyses were performed by symptom presentation and by sample, i.e., a series of one-way ANOVAs tested

the effects of symptom presentation within each sample, and a series of *t*-tests tested the effects of data source (i.e., type of sample: community *vs.* forensic) within each symptom presentation subgroup.

For all measures under investigation, a statistically significant difference by symptom presentation was observed in both the community ($F(3,509) \geq 88.85$, $p < .001$) and forensic ($F(3,284) \geq 53.59$, $p < .001$) samples. Bonferroni-corrected post hoc tests, more specifically, revealed that in all cases feigners scored significantly higher than honest responders on all measures, with Cohen's *d* values ranging from 1.01 to 3.43 within the community sample and from 0.56 to 1.97 within the forensic sample (Table 4, top part). Additionally, when focusing on the scores produced by the three community-based groups of feigners, the PAI NIM did not differ by symptom presentation, whereas the PAI MAL and PAI RDF yielded higher scores in the SCZ than in the PTSD group, and the PAI RDF and IOP-29 FDS yielded higher scores in the DEP than in the PTSD group. When focusing on the scores produced by the forensic groups of feigners, the only measure that produced significant differences across different symptom presentations is the PAI RDF, whose scores were significantly lower in the PTSD compared to the DEP and SCZ groups.

Three out of 16 results were statistically significant, after applying a Bonferroni correction, when comparing the scores observed in the community versus forensic samples (Table 4, bottom part). Specifically, honest responders in the forensic sample scored higher than those in the community sample on PAI NIM, PAI MAL and IOP-29 FDS; no other differences remained statistically significant after Bonferroni correction.

All point bi-serial correlations of the PAI and IOP-29 scores to group membership (0 = honest responder; 1 = simulator) were positive and statistically significant ($p < .01$; Table 5). The PAI NIM scale and IOP-29 FDS produced relatively stable correlation values across all three symptom presentations in both the community (PAI NIM: *range*: .80-.82; IOP-29

For Peer Review

FDS: *range*: .77-.86) and forensic (PAI NIM: *range*: .48-.54; IOP-29 FDS: *range*: .61-.64) samples. There were, instead, some differences from one symptom presentation to another for the other indexes under investigation. Specifically, the correlations within the SCZ and DEP groups were substantially stronger than those found in the PTSD group, for both the PAI RDF and, to a lesser extent, PAI MAL. As for the AUC (Table 5), the values for PAI NIM and IOP-29 FDS exceeded those of PAI MAL and PAI RDF in most cases, with the exception of the DEP feigning group in the forensic sample, in which the PAI RDF performed better than all other measures.

Finally, we analyzed diagnostic efficiency statistics by calculating Se, Sp, and OCC using the observed data (Table 6), and by estimating PPP and NPP based on a malingering base rate of 15% (Table 7). In the community sample, all PAI and IOP-29 scores produced similar sensitivity values across all disorders. More specifically, with NIM ≥ 81T, sensitivity ranged from .65 to .68 in the community sample and from .53 to .71 in the forensic sample; with FDS ≥ .50, it ranged from .73 to .86 and from .71 to .77 in the community and forensic samples, respectively. Results were remarkably stable when going from one condition to another also with regard to PPP and NPP values. Indeed, with NIM ≥ 81T, PPP was .92 in all conditions within the community sample and it ranged from .42 to .49 within the forensic sample; with FDS ≥ .50, PPP ranged from .91 to .93 within the community sample and from .57 to .59 within the forensic sample.

Additional analyses

Because we found statistically significant differences among the honest and feigner groups in the community sample on age, marital status, and gender (Table 1), we performed additional analyses to evaluate the extent to which this heterogeneity could influence our results. More specifically, we computed a series of partial correlations of the negative distortion variables to condition (dummy variable, 0 = honest; 1 = experimental simulator)

For Peer Review

removing, one at a time, the effects of each of those demographic variables. The results showed that the partial correlation values were nearly identical to the point bi-serial correlations reported in Table 4, with differences equal or lower than |.02|. Therefore, the differences between the groups in the non-clinical sample did not notably affect our findings.

Discussion

The aim of this study was to test the classification accuracy of the PAI and IOP-29 negative distortion scores to detect feigning of depression, PTSD, and schizophrenia in both a community and a forensic sample. Very similar procedures were followed with all feigners, so we could accurately compare the effectiveness of all measures under investigation across different symptom presentations (SCZ, PTSD, and DEP) and samples (i.e., community *vs.* forensic). The ultimate goal was to appreciate the extent to which the performance of the PAI and IOP-29 negative distortion scales would differ from one evaluation context to another, and whether or not different cut scores should be implemented in different situations.

Overall, our initial analyses (i.e., the one-way ANOVAs and *t*-tests) suggested that both the PAI and IOP-29 were effective in differentiating honest responders and experimental simulators. Focusing on the PAI, the NIM scale produced consistent results across symptom presentations and samples; however, we found significant differences on its scores in the honest group when comparing the community versus forensic samples, so that forensic, honest participants produced higher scores than community-based, honest ones. The MAL scores were stable across symptom presentations only in the forensic sample, while in the community sample lower scores were found in the group instructed to feign PTSD. Similar to the NIM, the MAL also showed higher scores in the forensic, honest group compared to the community-based, honest group. These different scores between the community and forensic honest conditions suggest that specificity values of the NIM and MAL may decrease when switching from community-based to forensic settings. Conversely, the RDF was the only

For Peer Review

measure to not show a significant interaction between symptom presentation and data source, thus revealing a comparable pattern of results in either sample. More specifically, with the RDF, feigners of PTSD scored lower than other feigners in both the community and forensic sample, but honest responders from either sample did not differ from each other. As for the IOP-29 FDS, it also produced stable effect sizes across all symptom presentations in the forensic sample; however, when considering the groups of honest responders, the forensic sample produced higher scores than the community sample, thus suggesting a possibly lower specificity in forensic settings.

Our study has some important implications for PAI users. First, when comparing the feigning against the honest groups, the NIM scale produced stable effect sizes across all different symptom presentations and overall higher effect sizes than the MAL or RDF. However, notably larger effect sizes were observed within the community sample, where Cohen's *d* ranged from 2.76 to 2.95; in the inmate sample, instead, Cohen's *d* ranged from 1.38 to 1.47. As for the MAL and RDF, they produced very large effect sizes in the SCZ and DEP groups, but yielded unexpected low values in the PTSD groups. Especially in the inmate group, Cohen's *d* was 0.91 for MAL and 0.56 for RDF. Detecting feigned PTSD thus seems to be more difficult with the PAI, confirming previous findings reported in the literature (Lange et al., 2010; Russel & Morey, 2019; Thomas et al., 2012). In general, the effect size values compare favorably with the non-clinical comparisons reported in the literature.

Considering the meta-analytic findings by Hawes and Boccaccini (2009) for NIM ≥ 81T (Se = .73, Sp = .83, OCC = .79), our total sample produced a slightly lower sensitivity (.65), but higher specificity (.94) and similar OCC (.77). As for MAL ≥ 3, Hawes and Boccaccini (2009) reported that it produced Se = .58, Sp = .83, OCC = .79; our total sample produced similar sensitivity (.50), higher specificity (.95), but lower OCC (.69). Given that in forensic settings one needs to favor specificity over sensitivity (Sherman et al., 2020), it is

For Peer Review

worth noting that the cut-off scores analyzed in our study produced relatively high specificity values (i.e., ≥ .88). Nevertheless, analyzing sensitivity, MAL and RDF produced very low sensitivity scores in detecting feigned PTSD. While a relatively poor performance of the RDF was expected (Eakin et al., 2006; Lange et al., 2010; Scragg et al., 2000; Thomas et al., 2012), MAL ≥ 3 produced a surprisingly low sensitivity, ranging from .31 to .43. On this matter, Thomas et al. (2012) suggested to use more conservative cut-off scores to correctly classify feigners of PTSD: NIM > 69T, MAL > 2, and RDF > 59T. Along similar lines, other authors tested and recommended other cut-scores (e.g., Russell & Morey, 2019), making it difficult to establish whether the sensitivity values we found across the symptom presentations are expected or not. Future studies may want to consider publishing diagnostic efficiency results by using the indication by Hawes and Boccaccini (2009), i.e., NIM ≥ 81T and MAL ≥ 3, and RDF ≥ 0, so to allow straightforward comparisons.

As for the IOP-29, its effectiveness in discriminating the feigning from the honest responding groups was remarkably similar across the three feigning conditions, in line with previous publications (e.g., Banovic et al., 2021; Carvalho et al., 2021; Giromini et al., 2018; Giromini, Viglione et al. 2020; Viglione et al., 2017; Winters et al., 2020). However, as is often the case in feigning studies, results were generally more satisfactory within the community-based than within the forensic samples. In fact, Cohen's *d* was very large in both the community ($2.52 \leq d \leq 3.43$) as well as in the forensic ($1.75 \leq d \leq 1.97$) groups, based on Rogers et al.'s (2003) standards. Along similar lines, AUC ranged from .93 to .98 in the community sample, and from .84 to .87 in the forensic sample. Besides, with the standard IOP-29 cut-off of FDS ≥ .50, OCC ranged from .87 to .92 within the community sample, and from .85 to .87 within the forensic sample. Taken together, these results compare favorably to those reported by IOP-29 authors in their initial studies, in which sensitivity, specificity, and OCC approximated .80 with the standard cut score of FDS ≥ .50 (Viglione et al., 2017;

For Peer Review

Giromini et al., 2018), although these differences are likely due to the fact that our study did not include a clinical sample. Indeed, it is known that specificity estimates tend to be inflated when simulation studies use non-clinical volunteers rather than genuinely impaired participants as their control groups (Rogers & Gillard, 2011).

Our study also suggests that despite its being comprised of 29 items only, the IOP-29 FDS is likely similarly effective as the more established PAI NIM scale, and that it perhaps outperforms the PAI MAL and RDF scales. Indeed, statistics reported in Table 5 showed that when considering the entire combined sample, both PAI NIM and IOP-29 FDS yielded an AUC value of .92 (*SE* = .01), whereas PAI MAL and PAI RDF yielded lower AUC values of .85 (*SE* = .01) and .82 (*SE* = .01) respectively. When focusing on the forensic sample more specifically, which is probably more relevant from an applied perspective, PAI NIM and IOP-29 FDS yielded AUC values of .84 (*SE* = .02) and .86 (*SE* = .02) respectively, whereas PAI MAL and PAI RDF yielded lower AUC values of .77 (*SE* = .03) and .78 (*SE* = .03) respectively. Likewise, similar conclusions may be drawn when looking at the various diagnostic efficiency statistics reported in Table 6.

Lastly, examination of PPP values reported in Table 7 also suggests some further considerations that are particularly relevant to forensic practitioners. Extant guidelines on how to perform symptom validity assessment suggest that empirically derived cutoffs should be set at a false-positive rate of 10% or lower, meaning that 90% of credible evaluees should be classified as credible by administered symptom validity measures. That is, the best cut scores for a given symptom validity measure are those that yield a specificity of .90 (Sherman et al., 2020). In our study, a-priori identified PAI and IOP-29 cut scores of NIM ≥ 81T and FDS ≥ .65 approximated or exceeded similarly high specificity levels, and with these same cut scores PPP values ranged from .42 to .92 for PAI NIM and from .59 to .93 for IOP-29 FDS. Accordingly, when using optimal, a-priori identified, PAI and IOP-29 cut scores, the

For Peer Review

likelihood that an individual would have an invalid presentation given a positive result would range from 8% to 58% for PAI NIM ≥ 81T and from 7% to 41% for IOP-29 FDS ≥ .65, based on our data and our estimated base rate of malingering of 15%. This finding should thus remind forensic practitioners that the determination of the validity of a presentation should never been made using one single measure, and that failing one single SVT should never be taken as sufficient evidence that a symptom presentation is invalid (Larrabee, 2008; Heilbronner et al., 2009). Indeed, the likelihood of false positive findings obviously decreases when multiple positive results are obtained, so it is often recommended that at least two noncredible results be observed prior to determining that a given presentation is invalid (Sherman et al., 2020).

Results reported in Table 7 also suggest that the IOP-29 is likely to be a very useful tool, when used as a screening measure. With the standard cut score of FDS ≥ .50, indeed, NPP was ≥ .95 both within the community and within the forensic sample. As such, the probability that a presentation is not invalid, given a negative test result (i.e., FDS < .50), is equal to or greater than 95%. Given that in a screening context only positive classifications are followed up with further testing, it is crucial for a screening measure to offer a similar level of precision, in ruling out possibly invalid presentations.

In conclusion, taken together these findings suggest that both the PAI and IOP-29 proved to be useful in detecting feigned disorders across contexts, confidently with the feigning of schizophrenia and depression, more cautiously with the feigning of PTSD, when considering the MAL and RDF of the PAI. Consistent with the extant literature, the PAI NIM proved to be the most effective negative distortion measure of the PAI and performed similarly across symptom presentations and contexts. The PAI MAL showed the best performance in detecting feigned schizophrenia and the worst in detecting feigned PTSD, with some – albeit small – differences between the forensic and community samples. It is

For Peer Review

worth noting, however, that the PAI MAL ≥ 3 produced the lowest sensitivity values of all measures under investigation. For both the PAI NIM and MAL, the specificity values were lower in the forensic setting, in line with the higher mean scores observed in the forensic, honest group. The PAI RDF appeared particularly suited to detect feigned depression in both samples and feigned schizophrenia in the community sample. As such, contrary to Hawes and Boccaccini (2009), our findings provide support for the use of the PAI also in detecting feigning of depression. Also the IOP-29 FDS performed similarly across symptom presentations and data source, with perhaps a slightly higher accuracy in detecting feigned depression, especially in the community sample.

When considering the implications of these findings, however, it should be noted that our study had some practical limitations. First, the ecological motivations present in real-life malingering conditions are inevitably different from those present in empirical studies. In order to motivate feigners in forensic samples, Rogers (2008) recommends to use both an external compensation (e.g., games for the prison) and an internal motivation (e.g., asking the participants whether they were skilled enough to beat the test). Although the internal motivation was elicited verbally by the researchers during the experimental procedure, the external incentives in our study (i.e., gift card and games) may be not sufficient, especially in the inmate sample. Secondly, our participants' compliance to our feigning instructions is difficult to evaluate. In fact, malingering studies operate on the assumption that test outcomes are primarily linked to the absence or presence of motivation. Therefore, respondents' non-compliance is a constant threat to validity in malingering research (Rai et al., 2019; Walls et al., 2017). In this study, we relied on the PAI MCE score to exclude feigners who did not provide an impaired PAI profile, which is in line with the principle that if an examinee looks credible but not impaired, they will not receive the compensation or reduced culpability they are seeking to obtain. On the other hand, future studies should strive to employ more rigorous

For Peer Review

manipulation checks. Thirdly, we did not have access to clinical samples of patients with depression, schizophrenia, or PTSD. As such, the effect size values overestimated the differences in the mean values of honest responders and feigners, as well as specificity. It should be pointed out, however, that in previous meta-analytic research, the effect sizes generated by different MMPI-2 validity scales strongly correlated with each other, when considering simulation studies conducted with non-clinical *versus* clinical controls (see, for instance, Table 4 of Rogers et al., 2003). As such, it is reasonable to believe that our results will generalize to clinical comparison simulation studies, to some extent, albeit likely with smaller effect sizes. Lastly, the community sample was characterized by differences among the subgroups in terms of age, gender, and marital status, and the forensic sample was composed of only men. Our additional analyses suggested that those demographic variables had null or minimal effects on the effectiveness of the discriminant capacity of the PAI and IOP-29 negative distortion variables. Nonetheless, future studies should address this potential limitation by recruiting more homogenous samples. Finally, because the base rate of malingering is basically unknown, our PPP and NPP results should be taken as tentative. And, for the same reason, they could not be compared against previous research given that, for instance, Hawes and Bocaccini (2009) did not report any data on these statistics, in their meta-analytic study.

For Peer Review

References

Boccaccini, M. T., Murrie, D. C., & Duncan, S. A. (2006). Screening for malingering in a criminal-forensic sample with the Personality Assessment Inventory. *Psychological Assessment*, *18*, 415-423. https://doi.org/10.1037/1040-3590.18.4.415

Duellman, R. M., & Bowers, T. G. (2004). Use of the Personality Assessment Inventory in forensic and correctional settings: Evidence for concurrent validity. *International Journal of Forensic Psychology*, *1*, 42-57.

Eakin, D. E., Weathers, F. W., Benson, T. B., Anderson, C. F., & Funderburk, B. (2006). Detection of feigned posttraumatic stress disorder: A comparison of the MMPI-2 and PAI. *Journal of Psychopathology and Behavioral Assessmen*t, 28, 145–155. https://doi.org/10.1007/s10862-005-9006-5

Edens, J. F., Cruise, K. R., & Buffington-Vollum, J. K. (2001). Forensic and correctional applications of the Personality Assessment Inventory. *Behavioral Sciences and the Law*, *19*, 519-543. https://doi.org/10.1002/bsl.457

Edens, J. F., Poythress, N. G., & Watkins-Clay, M. M. (2007). Detection of malingering in psychiatric unit and general population prison inmates: A comparison of the PAI, SIMS, and SIRS. *Journal of Personality Assessment*, *88*, 33-42. https://doi.org/10.1080/00223890709336832

Gaines, M. V., Giles, C. L., Morgan, R. D., (2012). The Detection of Feigning Using Multiple PAI Scale Elevations: A New Index. *Assessment*, *20*(4), 437-447. https://doi.org/10.1177/1073191112458146

Giromini, L., Barbosa, F., Coga, G., Azeredo, A., Viglione, D.J., Zennaro, A. (2020). 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

For Peer Review

experimental feigners. *Applied Neuropsychology: Adult*, *27*(6), 504-516. https://doi.org/10.1080/23279095.2019.1570929

Giromini, L., Carfora Lettieri, S., Zizolfi, S., Zizolfi, D., Viglione, D. J., Brusadelli, E., 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-29 (IOP-29). *Psychological Injury and Law*, *12*, 212-224. https://doi.org/10.1007/s12207-019-09357-7

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,* 340-350. https://doi.org/10.1007/s12207-018-9314-1

Giromini, L., Viglione, D. J., Pignolo, C., & Zennaro, A. (2020). An Inventory of Problems-29 sensitivity study investigating feigning of four different symptom presentations via malingering experimental paradigm. *Journal of Personality Assessment*, *102*(4), 563-572. https://doi.org/10.1080/00223891.2019.1566914

Hawes, S.W., & Boccaccini, M.T. (2009). Detection of Overreporting of Psychopathology on the Personality Assessment Inventory: A Meta-Analytic Review. *Psychological Assessment*, *21*(1), 112-124. https://doi.org/10.1037/a0015036

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference Participants. (2009). American academy of clinical neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, *23*(7), 1093-1129. https://doi.org/10.1080/13854040903155063

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

For Peer Review

Center for Epidemiologic Studies Depression Scale (CES-D) in Lithuania. *Applied Neuropsychology: Adult*. https://doi.org/10.1080/23279095.2020.1725518

Lange, R. T., Sullivan, K. A., Scott, C. (2010). Comparison of MMPI-2 and PAI validity indicators to detect feigned depression and PTSD symptom reporting. *Psychiatry Research*, *176*(2-3), 229-235. https://doi.org/10.1016/j.psychres.2009.03.004

Larrabee, G. J. (2008). Aggregation across multiple indicators improves the detection of malingering: Relationship to likelihood ratios. *The Clinical Neuropsychologist*, *22*(4), 666-679. https://doi.org/10.1080/13854040701494987

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, *52*(3), 194-216. https://doi.org/10.1037/h0048070

Merten, T., & Merckelbach, H. (2020). Factitious disorder and malingering. In Geddes, J. R., Andreasen, N. C., & Goodwin, G. M. (Eds.), The New Oxford Textbook of Psychiatry (3rd ed., pp. 1342-1349). Oxford, UK: Oxford University Press. https://doi.org/10.1136/practneurol-2018-001950

Morey, L. (1991). *Personality Assessment Inventory: Professional manual*. Tampa, FL: Psychological Assessment Resources.

Morey, L. C. (1996). *An interpretive guide to the Personality Assessment Inventory (PAI)*. Odessa, FL: Psychological Assessment Resources.

Morey, L. C. (2007). *Personality Assessment Inventory (PAI) professional manual (2nd ed.)*. Odessa, FL: Psychological Assessment Resources.

Morey, L.C., Quigley, B.D. (2002). The use of the Personality Assessment Inventory (PAI) in assessing offenders. *International Journal of Offender Therapy and Comparative Criminology*, *46*(3), 333-349. https://doi.org/10.1177/0306624X02463007

For Peer Review

Morey, L., Warner, M. B., & Hopwood, C. J. (2007). The Personality Assessment Inventory: Issues in legal and forensic settings. In A. M. Goldstein (Ed.), *Forensic psychology: Emerging topics and forensic roles* (pp. 97-126). Hoboken, NJ: Wiley.

Pignolo, C., Di Nuovo, S. Fulcheri, M., Lis, A. Mazzeschi, C., Zennaro, A. (2018). Psychometric Properties of the Italian Version of the Personality Assessment Inventory (PAI). *Psychological Assessment*, *30*(9), 1226-1236. https://doi.org/10.1037/pas0000560

Rai, J. K., An, K. Y., Charles, J., Ali, S., & Erdodi, L. A. (2019). Introducing a forced choice recognition trial to the Rey Complex Figure Test. *Psychology & Neuroscience*, *12*(4), 451-472. https://doi.org/10.1037/pne0000175

Resnick, P. J. (1999). The detection of malingered psychosis. *Psychiatric Clinics of North America*, *22*(1), 159-172. https://doi.org/10.1016/S0193-953X(05)70066-6

Rogers, R. (2008). Detection strategies for malingering and defensiveness. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 14-35). New York, NY: Guilford Press.

Rogers, R. (2018a). Researching response styles. In R. Rogers & S. D. Bender (Eds.), Clinical assessment of malingering and deception, 4th ed. (pp. 592-614). New York: Guilford.

Rogers, R., & Bender, D. (2018). *Clinical assessment of malingering and deception. (3rd ed.)*. New York: Guilford.

Rogers, R., & Gillard, N. D. (2011). *Research methods for the assessment of malingering.* In B. Rosenfeld & S. D. Penrod (Eds.), *Research methods in forensic psychology* (pp. 174-188). John Wiley & Sons Inc.

For Peer Review

Rogers, R., Ornduff, S. R., Sewell, K. W. (1993). Feigning Specific Disorders: A Study of the Personality Assessment Inventory (PAI). *Journal of Personality Assessment*, *60*(3), 554-560. https://doi.org/10.1207/s15327752jpa6003_12

Rogers, R., Sewell, K. W., Martin, M. A., & Vitacco, M. J. (2003). Detection of feigned mental disorders: A meta-analysis of the MMPI-2 and malingering. *Assessment*, *10*, 160-177. https://doi.org/10.1177/1073191103010002007

Rogers, R., Sewell, K. W., Morey, L. C., & Ustad, K. L. (1996). Detection of feigned mental disorders on the Personality Assessment Inventory: A discriminant analysis. *Journal of Personality Assessment*, *67*, 629-640. https://doi.org/10.1207/s15327752jpa6703_15

Roma, P., Giromini, L., Burla, F., Ferracuti, S., Viglione, D. J., & Mazza, C. (2019). 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*. https://doi.org/10.1007/s12207-019-09368-4

Russell, D., Morey, L. C. (2019). Use of Validity Indicators on the Personality Assessment Inventory to Detect Feigning of Post-Traumatic Stress Disorder. *Psychological Injury and Law*, *12*, 204-211. https://doi.org/10.1007/s12207-019-09349-7

Scragg, P., Bor, R., & Mendham, M. (2000). Feigning posttraumatic stress disorder on the PAI. *Clinical Psychology and Psychotherapy*, *7*, 155-160. https://doi.org/10.1002/(SICI)1099-0879(200005)7:2<155::AID-CPP237>3.0.CO;2-Z

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

For Peer Review

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.

Thomas, K. M., Hopwood, C. J., Orlando, M. J., Weathers, F. W., & McDevitt-Murphy, M. E. (2012). Detecting feigned PTSD using the personality assessment inventory. *Psychological Injury and Law*, *5*(3-4), 192-201. https://doi.org/10.1007/s12207-011-9111-6

Veltri, C. O. C., Williams, J. E. (2012). Does the Disorder Matter? Investigating a Moderating Effect on Coached Noncredible Overreporting Using the MMPI-2 and PAI. *Assessment*, 20(2), 199-209. https://doi.org/10.1177/1073191112464619

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

Viglione, D. J., Wright, D. M., 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*(3), 237-250. https://doi.org/10.1177/107319110100800301

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

Walls, B. D., Wallace, E. R., Brothers, S. L., & Berry, D. T. R. (2017). Utility of the Conners' Adult ADHD Rating Scale validity scales in identifying simulated attention deficit hyperactivity disorder and random responding. *Psychological Assessment*, *29*(12), 1437-1446. https://doi.org/10.1037/pas0000530

For Peer Review

Widows, M. R., & Smith, G. P. (2005). *SIMS-Structured Inventory of Malingered Symptomatology. Professional manual*. Lutz, FL: 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*. https://doi.org/10.1080/13218719.2020.1767720

Young, G. (2015). Malingering in forensic disability-related assessments: Prevalence 15±15%. *Psychological Injury and Law*, *8*(3), 188-199. https://doi.org/10.1007/s12207-015-9232-4.

Young, G. (2017). PTSD in court. Part III. Malingering, assessment/ testing, and the law. *International Journal of Law and Psychiatry*, *52*, 81-102. https://doi.org/10.1016/j.ijlp.2017.03.001

Young, G. (2019). The cry for help in psychological injury and law: Concepts and review. *Psychological Injury and Law*, *12*(3-4), 225-237. https://doi.org/10.1007/s12207-019-09360-y

Zennaro, A., Di Nuovo, S., Fulcheri, M., Lis, A., & Mazzeschi, C. (Eds.) (2015). *Personality Assessment Inventory (PAI). Manuale.* Firenze, IT: Hogrefe Editore.

For Peer Review

Table 1.

Composition of the sample.

	HON	SCZ Feigner	PTSD Feigner	DEP Feigner	Total
Community Sample					
N	188	81	114	130	513
Age [$F(3,533) = 8.015$, $p < .001$]					
M	39.6	31.5	35.4	39.2	37.3
SD	16.1	10.1	13.7	13.1	14.3
Relationship Status ($X^2 = 25.13$, $p < .001$)					
Not in a relationship (single, divorced, …)	99	64	80	66	309
In a relationship (married, cohabiting)	82	16	33	61	192
Education ($X^2 = 5.45$, $p = .142$)					
High school or less	125	43	72	74	314
Bachelor or more	60	38	42	55	198
Gender ($X^2 = 23.84$, $p < .001$)					
M	82	53	67	47	249
F	106	28	47	83	264
Forensic Sample					
N	146	58	36	48	288
Age [$F(3,283) = 1.50$, $p = .214$]					
M	40.6	42.9	42.4	38.1	40.9
SD	13.2	12.7	10.9	12.3	12.7
Relationship Status ($X^2 = 1.65$, $p = .648$)					
Not in a relationship (single, divorced, …)	82	37	21	27	167
In a relationship (married, cohabiting)	58	18	15	21	112
Education ($X^2 = 0.69$, $p = .875$)					
High school or less	132	51	33	46	262
Bachelor or more	10	4	3	2	19
Entire Sample					
N	334	139	150	178	801
Age [$F(3,769) = 3.13$, $p = .025$]					
M	40.1	36.3	37.1	38.9	38.6
SD	14.9	12.6	13.4	12.8	13.8
Relationship Status ($X^2 = 21.13$, $p < .001$)					
Not in a relationship (single, divorced, …)	181	101	101	93	476
In a relationship (married, cohabiting)	140	34	48	82	304
Education ($X^2 = 8.02$, $p = .046$)					
High school or less	257	94	105	120	576
Bachelor or more	73	42	45	57	217
Gender ($X^2 = 25.79$, $p < .001$)					
M	228	111	103	95	537
F	106	28	47	83	264

For Peer Review

Table 2.

Frequency of valid and invalid protocols.

	HON	ALL Feigners	SCZ Feigner	PTSD Feigner	DEP Feigner	Total
Community Sample						
Total	204	404	105	141	158	608
Missing items	0	6	4	0	2	6
ICN ≥ 73T	16	55	17	22	16	71
Feigner with MCE ≤ 60T	0	18	3	5	10	18
Valid protocols	188	325	81	114	130	513
Forensic Sample						
Total	209	249	104	50	95	458
Missing items	15	2	2	0	0	17
ICN ≥ 73T	48	82	32	10	40	130
Feigner with MCE ≤ 60T	0	23	12	4	7	23
Valid protocols	146	142	58	36	48	288
Entire Sample						
Total	413	626	209	164	253	1039
Missing items	15	8	6	0	2	23
ICN ≥ 73T	64	137	49	32	56	201
Feigner with MCE ≤ 60T	0	41	15	9	17	41
Valid protocols	334	467	139	150	178	801

For Peer Review

Table 3.

PAI and IOP-29 scores by symptom presentations and sample.

	Honest		ALL Feigners		SCZ Feigner		PTSD Feigner		DEP Feigner	
	M	SD	M	SD	M	SD	M	SD	M	SD
Community Sample										
PAI NIM	52.0	10.8	90.3	17.0	93.1	19.4	89.9	17.5	88.9	14.8
PAI MAL	0.53	0.72	2.76	1.57	3.17	1.60	2.47	1.58	2.76	1.50
PAI RDF	-1.04	0.91	0.52	1.24	0.68	1.06	0.08	1.37	0.81	1.12
IOP-29 FDS	0.19	0.15	0.73	0.24	0.73	0.26	0.69	0.26	0.78	0.20
Forensic Sample										
PAI NIM	62.1	17.3	87.2	18.7	88.6	22.4	85.8	16.9	86.6	14.8
PAI MAL	0.99	1.12	2.44	1.56	2.67	1.79	2.03	1.25	2.48	1.44
PAI RDF	-1.21	1.07	0.14	1.41	0.16	1.49	-0.55	1.51	0.65	1.02
IOP-29 FDS	0.27	0.18	0.66	0.27	0.65	0.29	0.66	0.26	0.66	0.27
Entire Sample										
PAI NIM	56.4	14.8	89.4	17.6	91.2	20.7	88.9	17.4	88.3	14.8
PAI MAL	0.73	0.94	2.67	1.57	2.96	1.69	2.37	1.52	2.69	1.49
PAI RDF	-1.11	0.99	0.41	1.31	0.46	1.28	-0.07	1.42	0.76	1.09
IOP-29 FDS	0.22	0.17	0.71	0.25	0.70	0.28	0.68	0.26	0.75	0.23

For Peer Review

Table 4.

Cohen's d effect sizes by symptom presentation and sample.

	PAI			IOP-29
	NIM	*MAL*	*RDF*	*FDS*
Differences by Symptom Presentation				
Community sample				
SCZ Feigner vs. HON	2.95**	2.48**	1.80**	2.85**
PTSD Feigner vs. HON	2.76**	1.73**	1.01**	2.52**
DEP Feigner vs. HON	2.93**	2.01**	1.85**	3.43**
SCZ Feigner vs. PTSD Feigner	0.17	0.44**	0.48**	0.15
SCZ Feigner vs. DEP Feigner	0.25	0.27	-0.12	-0.22
PTSD Feigner vs. DEP Feigner	0.06	-0.19	-0.59**	-0.39**
Forensic sample				
SCZ Feigner vs. HON	1.40**	1.25**	1.14**	1.75**
PTSD Feigner vs. HON	1.38**	0.91**	0.56*	1.97**
DEP Feigner vs. HON	1.47**	1.24**	1.76**	1.90**
SCZ Feigner vs. PTSD Feigner	0.14	0.40	0.47*	-0.04
SCZ Feigner vs. DEP Feigner	0.10	0.12	-0.38	-0.04
PTSD Feigner vs. DEP Feigner	-0.05	-0.33	-0.96**	0.00
Differences by Sample				
HON				
Community vs. Forensic	-0.72**	-0.50**	0.17	-0.49**
SCZ Feigner				
Community vs. Forensic	0.22	0.30	0.41	0.29
PTSD Feigner				
Community vs. Forensic	0.24	0.29	0.45	0.12
DEP Feigner				
Community vs. Forensic	0.16	0.19	0.15	0.54

Notes. Significance of differences by symptom presentations were obtained via Bonferroni-corrected post-hoc tests following one-way ANOVAs; significance of differences by sample (data source) were obtained via independent-samples t-test comparisons, with a Bonferroni-corrected alpha of .003125 for $p = .05$, and .000625 for p = .01. * Bonferroni-corrected $p < .05$; ** Bonferroni-corrected $p < .01$.

For Peer Review

Table 5.

Point bi-serial correlations with group and area under the receiver operator characteristic curves for the PAI and IOP-29.

	ALL Feigners			SCZ Feigner			PTSD Feigner			DEP Feigner		
	r	*AUC*	*SE*	*r*	*AUC*	*SE*	*r*	*AUC*	*SE*	*r*	*AUC*	*SE*
Community sample												
PAI NIM	.78**	.97	.01	.81**	.97	.01	.80**	.96	.01	.82**	.97	.01
PAI MAL	.63**	.90	.01	.75**	.94	.02	.64**	.87	.02	.71**	.91	.02
PAI RDF	.55**	.84	.02	.64**	.89	.02	.44**	.75	.03	.67**	.90	.02
IOP-29 FDS	.78**	.95	.01	.80**	.94	.02	.77**	.93	.01	.86**	.98	.01
Forensic sample												
PAI NIM	.58**	.84	.02	.54**	.82	.03	.48**	.84	.03	.54**	.85	.03
PAI MAL	.47**	.77	.03	.49**	.78	.04	.34**	.74	.04	.47**	.79	.04
PAI RDF	.48**	.78	.03	.46**	.77	.04	.22*	.63	.06	.61**	.89	.03
IOP-29 FDS	.64**	.86	.02	.62**	.84	.03	.61**	.87	.03	.64**	.86	.04
Entire sample												
PAI NIM	.70**	.92	.01	.69**	.91	.01	.69**	.92	.01	.72**	.93	.01
PAI MAL	.58**	.85	.01	.64**	.87	.02	.55**	.82	.02	.63**	.87	.02
PAI RDF	.53**	.82	.01	.55**	.84	.02	.39**	.73	.03	.66**	.90	.01
IOP-29 FDS	.73**	.92	.01	.72**	.90	.02	.72**	.91	.01	.79**	.94	.01

Note. * $p < .05$; ** $p < .001$.

For Peer Review

Table 6.

Diagnostic accuracy of PAI and IOP-29 scores by symptom presentations: Se, Sp and OCC.

	PAI					IOP-29				
	NIM		MAL		RDF	FDS				
	≥ 81T	≥ 110T	≥ 3	≥ 4	≥ 0	≥ .15	≥ .30	≥ .50	≥ .65	≥ .70
Community sample										
Se	.66	.17	.53	.29	.69	.98	.92	.80	.70	.65
Sp	.99	1.00	.98	1.00	.89	.55	.81	.96	.99	1.00
OCC	.78	.47	.70	.55	.76	.82	.88	.86	.80	.78
SCZ										
Se	.68	.28	.65	.35	.74	.95	.89	.79	.72	.65
OCC	.90	.78	.88	.80	.84	.67	.84	.91	.91	.90
PTSD										
Se	.65	.17	.43	.25	.58	.97	.89	.73	.60	.57
OCC	.86	.69	.77	.72	.77	.71	.84	.87	.84	.84
DEP										
Se	.65	.10	.55	.28	.75	.99	.96	.86	.77	.72
OCC	.85	.63	.80	.71	.83	.73	.87	.92	.90	.88
Forensic sample										
Se	.64	.16	.44	.23	.58	.94	.81	.73	.61	.56
Sp	.87	.99	.90	.97	.88	.32	.68	.90	.93	.96
OCC	.76	.58	.67	.60	.74	.63	.74	.82	.77	.76
SCZ										
Se	.66	.28	.53	.29	.59	.91	.76	.71	.64	.59
OCC	.81	.78	.80	.78	.80	.49	.70	.85	.85	.85
PTSD										
Se	.53	.14	.31	.11	.42	1.00	.86	.72	.56	.56
OCC	.80	.82	.79	.80	.79	.46	.71	.87	.86	.88
DEP										
Se	.71	.04	.42	.23	.71	.94	.83	.77	.63	.52
OCC	.83	.75	.78	.79	.84	.47	.72	.87	.86	.85
Entire Sample										
Se	.65	.17	.50	.27	.66	.97	.89	.78	.67	.62
Sp	.94	.99	.95	.99	.89	.45	.75	.93	.96	.98
OCC	.77	.51	.69	.57	.75	.75	.83	.84	.79	.77
SCZ										
Se	.67	.28	.60	.32	.68	.94	.83	.76	.68	.63
OCC	.86	.78	.85	.79	.82	.59	.78	.88	.88	.88
PTSD										
Se	.62	.16	.40	.21	.54	.98	.89	.73	.59	.57
OCC	.84	.74	.78	.75	.78	.61	.80	.87	.85	.85
DEP										
Se	.67	.08	.51	.27	.74	.98	.93	.84	.73	.66
OCC	.84	.68	.79	.74	.84	.63	.81	.90	.88	.87

Note. Se = sensitivity; Sp = specificity; OCC = Overall Correct Classification.

For Peer Review

Table 7.

Diagnostic accuracy of PAI and IOP-29 scores by symptom presentations: PPP and NPP estimated based on a malingering base rate of 15%.

	PAI					IOP-29				
	NIM		MAL		RDF	FDS				
	≥ 81T	≥ 110T	≥ 3	≥ 4	≥ 0	≥ .15	≥ .30	≥ .50	≥ .65	≥ .70
Community sample										
PPP	.92	1.00	.82	1.00	.52	.28	.47	.77	.92	1.00
NPP	.94	.87	.92	.89	.94	.99	.98	.96	.95	.94
SCZ										
PPP	.92	1.00	.84	1.00	.54	.27	.46	.77	.92	1.00
NPP	.95	.89	.94	.90	.95	.98	.98	.96	.95	.94
PTSD										
PPP	.92	1.00	.78	1.00	.48	.28	.46	.75	.91	1.00
NPP	.94	.87	.91	.88	.92	.99	.98	.95	.93	.93
DEP										
PPP	.92	1.00	.82	1.00	.54	.28	.48	.78	.93	1.00
NPP	.94	.86	.92	.89	.95	1.00	.99	.98	.96	.95
Forensic sample										
PPP	.46	.68	.45	.59	.47	.20	.31	.57	.61	.70
NPP	.93	.87	.90	.88	.92	.97	.95	.95	.93	.92
SCZ										
PPP	.47	.78	.50	.65	.47	.19	.29	.57	.62	.72
NPP	.93	.89	.92	.89	.92	.95	.94	.95	.94	.93
PTSD										
PPP	.42	.64	.36	.42	.39	.21	.32	.57	.59	.70
NPP	.91	.87	.88	.86	.90	1.00	.97	.95	.92	.92
DEP										
PPP	.49	.35	.43	.60	.52	.20	.31	.59	.62	.69
NPP	.94	.85	.90	.88	.94	.97	.96	.96	.93	.92
Entire Sample										
PPP	.65	.83	.62	.80	.50	.24	.39	.68	.77	.86
NPP	.94	.87	.92	.88	.94	.99	.97	.96	.94	.94
SCZ										
PPP	.65	.89	.66	.83	.51	.23	.37	.67	.77	.86
NPP	.94	.89	.93	.89	.94	.98	.96	.96	.95	.94
PTSD										
PPP	.64	.83	.57	.76	.46	.24	.39	.66	.74	.85
NPP	.93	.87	.90	.88	.92	.99	.97	.95	.93	.93
DEP										
PPP	.65	.71	.63	.80	.53	.24	.40	.69	.78	.87
NPP	.94	.86	.92	.88	.95	.99	.98	.97	.95	.94

Note. PPP = positive predictive power; NPP = negative predictive power.

For Peer Review

Appendix A – instructions given to the psychosis-feigning group

Malingering refers to the simulation of psychological or physical symptoms for obtaining financial benefits or avoidance of responsibility. By following the instructions and participating in this research, you will contribute to the refinement of techniques to effectively detect true simulators.

When you are answering the questions on the tests, I would like you to pretend that you are psychotic. Some people refer to this as "crazy," "insane," or "nuts." **If you can produce test results that look like those of a psychotic person and the test will not detect you as "feigner", you will qualify for a lottery of 20€ (for example, pre-paid phone card, amazon gift card etc). The two best feigners will be awarded after the results of this study have been analyzed.** It is important that you try hard to fake the tests and that you convince the examiner that you are psychotic.

Malingering of psychiatric problems is a problem that can be costly to the public. Many individuals who do not have genuine psychiatric problems attempt to feign problems so as to receive Social Security benefits, rather than joining the work force. Criminals may attempt to appear psychotic in order to be attributed with the insanity condition that exonerates them from responsibility for their actions and associated sentences. Because tax dollars support these individuals, the public may be unjustly supporting individuals who do have the capacity to support themselves.
Finding ways to better detect malingering can help save the public millions in tax dollars.

It may be easier for you to feign a mental illness if you pretend to be in a specific situation. Therefore, imagine that you have come upon some financial difficulties. With the slowing economy, you were unjustly laid off from your job, while people who with less seniority and experience kept their jobs. Although you had collected unemployment benefits for a while, bills

For Peer Review

have been piling up, and you can't cover rent, and are kicked out of your house. After living on the streets for a few days, a homeless person told you if you act like you're "crazy," you can get food, shelter, and other benefits. You could even have an opportunity to receive job retraining, and possibly a fresh start.

To clarify, psychotic people can suffer from a wide range of symptoms, such as hearing voices, seeing things other people don't see, paranoia, oddities in thinking, strange beliefs about the world, difficulty staying focused on a single topic, and difficulty communicating effectively with others. In addition, psychotic individuals are likely to have poor functioning skills in at least one major area of their lives, such as school, work, interpersonal relationships, or self-care. Not all people who are psychotic have all the symptoms. The term psychotic may mean different things to different people. It can show up in different ways, and is sometimes mild, and sometimes severe.

Therefore, be careful not "overdo it" because otherwise, you may not come across as credible and the examiner may realize that you are not really suffering from a psychotic disorder, but are only simulating it.

Good luck.

For Peer Review

Appendix B – instructions given to the depression-feigning group

When you fill out the test, I would like you to pretend that you are depressed. Many people experience moments when they feel "sad", "down" or "down in the dumps". These feelings are normal and are experienced by most people throughout their lives. However, people with a diagnosis of Major Depressive Disorder experience these feelings in a way that interferes with their daily functioning. **If you will be able to produce test results that are consistent with those produced by people with depressive symptoms and you will not look like a feigner, you may win a prize consisting of a 25€. The two best feigners will be awarded after the results of this study have been analyzed.** It is important that you try your best to simulate and convince the examiner that you are depressed.

It may be easier for you to feign a mental illness if you pretend to be in a specific situation. Therefore, imagine yourself in the following situation.

You are an administrator at a small, well-established firm. Your boss has been trying to cut expenses by having the cleaning crew work before regular work hours are over, thus getting the job done at a cut rate. You have repeatedly informed him that this is not a safe working condition for the employees, but he has not changed the procedure. One day, near the end of the day, you are leaving to do a special errand for your boss. As you cross a freshly mopped floor, you slip and fall, landing hard on your tailbone. As a result, you have been out of work for 2 weeks on disability and continue to experience a fair amount of pain, particularly when you sit for any length of time. The workers compensation physician insists that he can find nothing to explain the pain and refuses to authorize any more time off or disability payments, stating that you are able to return to work, a job that requires long periods of time sitting at your computer. You are angry with your boss for the injury you have and frustrated at the physician's apparent collusion with your boss to unreasonably limit your recovery time (thereby cutting off his disability payments). Before terminating your case,

For Peer Review

the physician refers you to the staff psychologist for a routine evaluation. You correctly realize that this evaluation is your only opportunity to remain on disability under your employer's obligation. You have no additional coverage and need an income until you are fully recovered. You also feel that your boss is responsible, and that money should come from the company through workers compensation. You know well that workers compensation will continue providing benefits to patients who are psychologically disturbed as a result of a work-related accident. This would not be too unusual because you have tried to take measures to avoid the problem, and now are suffering as a result of your boss's negligence. So, your only choice is to present yourself as having significant depression on the tests that the psychologist is going to give you. You therefore decide to attempt to present yourself as having a major depression as the result of your accident, to remain on disability. However, you also realize that if you present your condition in an extremely dramatic way, you'll look like a fraud, and you'll lose any chance of getting compensation or disability benefits. So be careful to keep a realistic and convincing profile.

Please, **review the list of symptoms below** and spend five minutes thinking about the most effective way to convince the examiner that you do have depression.

- Depressed mood most of the day, nearly every day
- Markedly diminished interest or pleasure in all, or almost all, activities
- Significant weight loss when not dieting or weight gain, or decrease or increase in appetite
- Insomnia or hypersomnia
- Psychomotor agitation or retardation
- Fatigue or loss of energy
- Feelings of worthlessness or excessive or inappropriate guilt
- Low self-esteem
- Diminished ability to think or concentrate, or indecisiveness
- Feelings of despair

For Peer Review

- Recurrent thoughts of death

Be careful, however, not "overdo it" because otherwise, you may not come across as credible and the examiner may realize that you are not really suffering from a depressive disorder, but are only simulating it.

If you have any questions concerning this research, you will be able to ask them after you have finished taking the test and filled out some questionnaires.

Good luck and thank you for volunteering to participate in this research.

For Peer Review

Appendix C – instructions given to the PTSD-feigning group

We are now asking you to try to put yourself in the shoes of a person who answers the questions on this test while **trying to pretend to have Post-Traumatic Stress Disorder.**

INSTRUCTIONS:

You will be asked to take a test that typically serves to assess some of the changes that people can undergo when they have experienced a highly stressful event. In answering, try to pretend you have been the victim of an earthquake during which you saw your house collapse and have been living in a state of fear and stress ever since. Imagine that you have applied to the government for financial compensation reserved for people who, after being exposed to a highly stressful event, have developed a mental disorder called "Post-Traumatic Stress Disorder" (see below). **Then try to answer the test questions as you imagine a person who has been truly traumatized and who wants to convince the examiner of the validity of his or her suffering and symptomatology in order to obtain some form of financial compensation might do.** Stay in this role for the duration of the test. To help you simulate best, try to imagine that you have experienced the following situation:

"Imagine that you live in a city hit by an earthquake and have experienced great economic hardship. With the economic crisis, you have been unfairly dismissed from the company where you had been working for a long time and have been forced to take up casual work for a cleaning company. To make things worse, you have just been divorced and will have to support your children, who are still minors. One day, you find out that an acquaintance of yours is receiving financial aid from the government to deal with the psychological distress that was caused to him by a recent earthquake. Therefore, you start to think that you too could apply for the same benefit to be able to support yourself and, most of all, to pay for your children's studies. Keep in mind, however, that this benefit is given by the government only to people who have a mental disorder called "Post-Traumatic

For Peer Review

Stress Disorder". Thus, you decide to study the characteristics of this disorder (described below) and take the test pretending that you have it too."

We invite you to study the list of symptoms below and think about the most effective way to convince the examiner that you have post-traumatic stress disorder:

1. Having stressful, unwanted, repetitive memories related to the event
2. Having lost interest in activities you used to do; difficulty experiencing positive emotions
3. Actively avoid event-related thoughts and feelings and consider event-related places, activities, and situations dangerous
4. Viewing self and others negatively (e.g., "I can't trust others," "I'm fragile"); social isolation, feeling like you can't count on anyone
5. Having memory lapses related to the stressful event
6. Having nightmares related to the event; difficulty falling asleep or disturbed sleep
7. Being emotionally upset in thinking back to the event; having felt that the event was happening again
8. Experiencing feelings of guilt or blaming others for the event or what happened afterward; having frequent feelings of fear, anger, or terror; feeling irritable and/or aggressive
9. Recklessness (e.g., taking substances, driving recklessly, having unprotected sex)
10. Being on constant alarm (e.g., looking around, being alert to those around you, etc...); easily startled by anything; physical hyper-reactivity when thinking about the event (e.g., palpitations, cold sweat, etc...)
11. Difficulty focusing and doing everyday things (work, study, friends, family) because of the problems described above.

For Peer Review

Be careful, however, not "overdo it" because otherwise, you may not come across as credible and the examiner may realize that you are not really suffering from a Post-Traumatic Stress Disorder, but are only simulating it.

If you can produce test results that look like those of a person who suffer from Post-Traumatic Stress Disorder and the test will not detect you as "feigner", you will qualify for a lottery of 20€ (for example, pre-paid phone card, amazon gift card etc). The two best feigners will be awarded after the results of this study have been analyzed.

For Peer Review