IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

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

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1766449 since 2021-01-12T14:34:43Z

Published version:

DOI:10.1080/23279095.2020.1725518

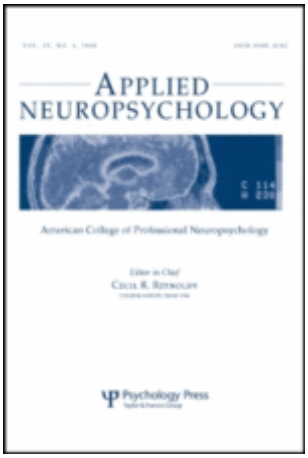
Terms of use:

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)

23 April 2024



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

Journal:	*Applied Neuropsychology: Adult*
Manuscript ID	Draft
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Ilgunaite, Guste; Mykolas Romeris University Giromini, Luciano; Department of Psychology, University of Turin, Italy, Bosi, Jessica; University of Surrey Viglione, Donald; California School of Professional Psychology, Alliant International University - San Diego Zennaro, Alessandro; Department of Psychology, University of Turin, Italy
Keywords:	malingering, feigning, depression, IOP-29, Lithuania

SCHOLARONE™
Manuscripts

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

Guste Ilgunaite[1], Luciano Giromini[2], Jessica Bosi[3], Donald J. Viglione[4], & Alessandro Zennaro[2]

[1] Department of Psychology, Mykolas Romeris University, Vilnius, Lithuania

[2] Department of Psychology, University of Turin, Italy

[3] University of Surrey, Guildford, UK

[4] Alliant International University – San Diego, California, US

Corresponding Author: Correspondence concerning this article should be addressed to Luciano Giromini, Department of Psychology, University of Turin, Via Verdi 10, 10123 Torino, TO, Italy. Tel +39 011 670 3060. E-mail luciano.giromini@unito.it

Declaration of Interest: Luciano Giromini and Donald J. Viglione declare that they own a share in the corporate (LLC) that possesses the rights to Inventory of Problems. All other three authors declare that they have no conflict of interest.

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.

For Peer Review Only

Lithuanian Validation of the IOP-29

Abstract

This article contributes to the growing research on the validity of the recently developed, Inventory of Problems – 29 (IOP-29) in the discrimination of feigned from bona fide mental or cognitive disorders. Specifically, we first developed a Lithuanian version of the IOP-29 and tested its validity on a sample of 50 depressed patients and 50 healthy volunteers instructed to feign depression. Next, we reviewed all previously published IOP-29 studies reporting on depression-related presentations (k = 5), and compared our results against previously reported findings. Statistical analyses showed that the Lithuanian IOP-29 discriminated almost perfectly between genuine and experimentally feigned major depression, with Area Under the Curve (AUC) = .98 (SE = .01) and Cohen's d = 3.31. When compared to previously published IOP-29 literature on this same topic, these findings may be characterized as similar or perhaps slightly more encouraging. Indeed, across all international, empirical studies considered in this article, Cohen's d ranged from 1.80 to 4.30, and AUC ranged from .89 to .99. Taken together, these findings contribute to supporting the strong validity and cross-cultural applicability of the IOP-29. They also provide additional support for its use in forensic evaluations.

Keywords: Malingering; Feigning; Depression; IOP-29; Lithuania.

For Peer Review Only

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

Malingering is defined as the "intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives" (American Psychiatric Association, 2013, p.726). Despite this straightforward definition, malingering can be extremely difficult to detect and the costs associated with it, both to the criminal justice system and to society, are exorbitant (Chafetz & Underhill, 2013). Therefore, accurately deciphering whether an individual's presentation of symptoms is credible or not is critical issue and is key when undertaking forensic psychological assessments (Bush, Heilbronner, & Ruff, 2014).

Major depression has a high prevalence of 7% (APA, 2013) and has associated economic consequences, including direct and indirect costs related to medications, disability compensation, insurance claims and work absence (Cuijpers & Smit, 2008). In relation to the latter, the duration of absence from work is typically longer for cases of major depression than for other medical problems such as heart disease and back pain (Druss, Rosenheck, & Sledge, 2000). Perhaps more importantly, symptoms of major depression can be feigned more easily than symptoms for other mental disorders, for several reasons. For instance, most people have at some point experienced low mood, the symptoms themselves are likely easier to empathize with (such as a lack of interest in daily life), and there is a lot of easily accessible information on the symptoms available online (Bagby, Nicholson, Buis, & Bacchiochi, 2000; Lees-Haley & Dunn, 1994; Monaro et al., 2018; Nicholson & Martelli, 2007; Steffan, Clopton, & Morgan, 2003). As a result, depression is one of the most commonly feigned mental disorders (Mittenberg, Patton,

URL: http://mc.manuscriptcentral.com/hapn Email: drmachorton@hotmail.com

For Peer Review Only

Canyock, & Condit, 2002) and consequently the veracity of reports of its symptoms must be investigated carefully in the forensic setting.

To evaluate the credibility of presented symptoms, practitioners may rely on various tools which have been developed in the absence of reliable detection based on clinical judgment alone (Young, Jacobson, Einzig, Gray, & Gudjonsson, 2016). These tools include interviews (e.g. The Structured Interview of Reported Symptoms (SIRS; Rogers, Gillis, Dickens, & Bagby, 1991; Rogers, 2010); validity indicators embedded in multiscale personality inventories (e.g., the Minnesota Multiphasic Personality Inventory (MMPI-2; Greene, 1991); MMPI-RF (Ben-Porath & Tellegen, 2008) and the Personality Assessment Inventory (PAI; Morey, 1991) and "stand-alone tests" including self-report symptom validity tests (SRVTs) (e.g. The Structed Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997)) and performance validity tests (PVTs) (e.g., the Test of Memory Malingering (TOMM; Tombaugh, 1996; Tombaugh, 1997)). Whilst all these tools have their merits, the majority rely on one detection strategy only, for example, SRVTs like the SIMS typically rely purely on the "rare symptoms strategy" (Rogers & Bender, 2018, p. 572). Although this is one of the most effective strategies for assessing feigned mental disorders (Rogers, Sewell, Martin, & Vitacco, 2003; Sharf, Rogers, Williams, & Henry, 2017) the failure to examine other strategies could lead to cases of feigning going undetected.

The newly developed Inventory of Problems-29 (IOP-29; Viglione, Giromini, & Landis, 2017), a short, paper-and-pencil or online, self-administered measure to detect *multiple* feigning strategies of both mental and cognitive disorders, aims to evaluate the credibility of presented symptoms by examining the extent to which an individual experiences and may cope with or manage their problems. Compared to available alternatives, the IOP-29 seems to offer some advantages. First, it is notably shorter than other popular stand-alone instruments such as the

For Peer Review Only

SIMS (29 versus 75 items). Emerging research suggests that it might outperform the TOMM in the detection of feigned depression (Giromini, Barbosa et al., 2019; Viglione et al., 2017), and that it has increased classification accuracy compared to the SIMS, especially with psychosis-related conditions (Giromini, Viglione, Pignolo, & Zennaro, 2018). Also of note is that it shows incremental validity when used either with the TOMM or MMPI compared to using each instrument alone (Giromini, Lettieri et al., 2019; Giromini et al., 2019). Furthermore, the IOP-29 may be applied to various forensic evaluation contexts with no need to adjust its cut scores, as it has been shown to perform similarly well with four different types of symptom presentations (those related to depression, psychosis, post-traumatic stress disorder (PTSD) and mild traumatic brain injury (mTBI)) (Giromini, Viglione, Pignolo, & Zennaro, 2019).

This Study

The IOP-29 has primarily been investigated by the two IOP-29 authors using Italian or US populations with limited studies focusing on the cross-cultural applicability of this new measure (see Giromini, Barbosa et al., 2019 for its validation with a Portuguese sample and Winters et al., 2019 for a British validation). To our knowledge, it has not yet been used or validated in Lithuania. In fact, Lithuanian research on malingering itself is incredibly scarce: on running a search in PsycINFO for the terms "malinger*" and "Lithuania*" in October 2019, no papers were found. In comparison, a search for just "malinger*" on the same date, returned over 4,000 papers. We believe that providing researchers with a validated, Lithuanian adaptation of the IOP-29 would facilitate further research in this field as it is so short (29 items) and easy-to-use.

The primary goals of this study therefore were threefold: 1) to develop a Lithuanian version of the IOP-29, 2) to test its validity in the detection of feigned depression and 3) to

For Peer Review Only

compare our findings against those published in previous IOP-29 literature on feigned depression. To do this, the original (English) version of the IOP-29 was first adapted in Lithuanian and permission was then sought to use the Center for Epidemiologic Studies Depression Scale (CES-D, Radloff, 1997) in Lithuanian. The CES-D is a widely used tool to evaluate the presence of depressive symptoms. It was used in this study to confirm the presence of depression in the patient group and to ensure that feigners followed instructions and in fact did attempt to appear depressed. Finally, 50 bona fide patients with depression and 50 feigners in Vilnius, Lithuania were recruited.

Materials and Methods

Participants

Participants for both groups had to be native Lithuanian speakers, over the age of 18 and had to know how to read and sign an informed consent form.

Patients. Fifty Lithuanian adults, 33 women (66%) and 17 men (34%), ranging in age from 18 to 67 years of age (M =35.60, SD = 12.21) with diagnosed depression were recruited for the patient group, i.e., to answer the IOP-29 and CES-D honestly. The average number of years of education was 16.08 (SD =1.90). Thirty-four (68%) had been diagnosed with major depression disorder, 11 (22%) had been diagnosed with depression and anxiety disorder and 5 (10%) had been diagnosed with adjustment disorder with depression. The majority (n = 37) were outpatients who attended regular consultations with therapists every week and 13 were inpatients who were taken to hospital following a crisis or were there for rehabilitation purposes. Diagnoses were made by psychiatrists based on clinical judgement, consulting with ICD-10-am or using a structured interview for diagnosing depression. Participants were accepted to the depression group if they had depression and were excluded if they had some other diagnosis.

URL: http://mc.manuscriptcentral.com/hapn Email: drmachorton@hotmail.com

For Peer Review Only

Experimental Simulators. Fifty Lithuanian adults, 36 women (72%) and 14 men (28%), ranging in age from 21 to 55 years of age (M = 31.82, SD = 8.41) were recruited to feign depression, i.e., to try and trick the questionnaires into believing that they had a diagnosis of depression. The average number of years of education was 16.42 (SD = 1.80). The simulator group thus did not differ from the patient group with regard to gender (Phi = .065, p = .52), age ($t(87.0) = 1.80$, p = .07)[1], or education ($t(98) = .92$, p = .36).

Materials

The Inventory of Problems-29 (IOP-29; Viglione et al., 2017). As introduced above, the IOP-29 is a brief, self-administered test, newly developed to aid practitioners in their evaluation of symptom presentation credibility related to a variety of psychiatric or cognitive disorders. It is comprised of 29 items, 27 of which are statements about cognitive, social and emotional experiences for which the test taker must choose either "true", "false" or "doesn't make sense", and 2 of which are open-ended questions requiring mathematical reasoning to be applied. The responses are analyzed using a logistic regression-derived formula to generate the main feigning index of the IOP-29, the False Disorder Probability Score (FDS), which is expressed as a probability score: the higher the FDS, the lower the credibility of the reported symptoms, with zero being the minimum and one being the maximum. Cut-off scores of FDS ≥ .50 has been shown to ensure the best balance between sensitivity and specificity (Giromini et al., 2018; Viglione et al., 2017).

For this study, a Lithuanian version of the IOP-29 was developed by following the standard, "back-translation" method (Brislin, 1980; Geisinger, 2003; Van de Vijver &

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

For Peer Review Only

Hambleton, 1996). This process involved a native Lithuanian individual, translating the original, English IOP-29 into Lithuanian followed by a Lithuanian-English bilingual individual who was blind to the original version of the IOP-29 back-translating this Lithuanian version into English. Finally, two of the developers of the IOP-29 compared the two English versions to deal with any inconsistencies and revise the translations with the Lithuanian translators, if necessary.

The Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977). The CES-D, a 20-item self-report tool, was used to assess symptoms of depression and anxiety. It includes items that evaluate affective, cognitive and somatic symptoms. Individuals must select from a 4-point Likert scale (0 = "rarely/less than 1 day" to 3 = "most of the time/5-7 days") how often the statements applied to them in the previous week. Scores range from 0 to 60 and individuals are categorized into one of four groups depending on their total score: a) not depressed (0–9 points), b) mildly depressed (10–15 points), c) moderately depressed (16–24 points), or d) severely depressed (more than 25 points). A cut-off score of ≥ 16 is a widely used indicator for likely clinically meaningful depressive symptoms (Radloff, 1977).

Procedure

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 was obtained from all individual participants included in the study.

In order to recruit participants for the patient group, contact was made with psychologists and psychotherapists working in private practice, mental health centers and hospitals in Vilnius, Lithuania from June 2018. For those practitioners who agreed to participate, a short description of the study was sent, detailing the aim of the study – to see if the IOP-29 would recognize

For Peer Review Only

people who are feigning depression and people who are bona fide patients with depression. Patients were told that if they were willing to participate, they would be asked to complete a few questionnaires by responding honestly. For those who had private consultations, the questionnaires were taken to the therapist in person and collected once completed. Participants were given contact details of the primary researcher if they wanted to be debriefed on the results of the study. To recruit participants for the feigning group, an email was sent to a group of colleagues in the Vilnius University with a brief description of the research project. Additional participants were recruited via convenience sampling and a snowball effect.

The recruitment of participants for both groups was completed after about four months. All received a participant information sheet and informed consent was given. Participants were told to try their best to feign to trick the tests and that the three best feigners (operationalized by us as the three people who got the lowest FDS score on the IOP-29) would receive a free coffee card in a local coffee shop. Participants were then given a vignette, a scenario, a description of the symptoms of depression and a cautionary statement. These were all used in an earlier study (to see them full, please see Giromini, Carfora Lettieri et al., 2019) but in brief, the vignette gave participants a description of what the study involves along with the instruction to "take the tests as you imagine someone who really is depressed would do." The scenario involved an incident in the workplace resulting in time off work due to injury with the inability to claim disability allowance for physical injury. Therefore, "your only choice is to present yourself as having significant depression on the tests." The symptoms of depression were then listed, and participants were told that "depressed patients typically have 5 or more, but most likely not all of them." Finally, participants were instructed not to "over-do it" and not to "present [their]

URL: http://mc.manuscriptcentral.com/hapn Email: drmachorton@hotmail.com

For Peer Review Only

condition in an extremely dramatic way" or else the performance would not be believable. Post-test questions were also included to check that participants had understood the task.

Data Analysis

To evaluate the validity of the Lithuanian IOP-29 in discriminating experimentally feigned from bona fide depression, we performed t-test and receiver operating characteristic (ROC) curve analyses. Additionally, we also inspected the sensitivity and specificity values yielded by the following cut scores: FDS ≥ .70; FDS ≥ .65; FDS ≥ .50; FDS ≥ .30; FDS ≥ .15. The former two cut scores may be recommended in high-stakes forensic situations in which one might need to seek for specificity values of .95 and .90 (respectively); the latter two cut scores may be recommended when the IOP-29 is used for screening purposes only, as they are supposed to yield sensitivity values of .90 and .95 respectively (Giromini et al., 2018). With regard to the cut score of FDS ≥ .50, it is the 'standard' IOP-29 cut score, which maximizes the classification accuracy of the test by offering the best balance between sensitivity and specificity (Giromini et al., 2018; Viglione et al., 2017).

Next, to contribute to evaluating the cross-cultural applicability of the IOP-29, we compared our findings against those presented in the five published articles that reported on the validity of the IOP-29 in discriminating feigned from bona fide depression-related symptom presentations. More specifically, one of Viglione et al.'s (2017) cross-validation samples included 43 American patients with a diagnosis of major depression and 42 American experimental simulators instructed to feign depression. A subsample of Giromini et al.'s (2018) study comprised 127 Italian patients genuinely suffering from a non-psychotic, anxiety, depression and/or trauma-related condition and 111 Italian experimental simulators instructed to fake similar mental health problems. Fifty of the 100 Portuguese experimental simulators

URL: http://mc.manuscriptcentral.com/hapn Email: drmachorton@hotmail.com

For Peer Review Only

included in Giromini, Barbosa, et al.'s (2019) took the IOP-29 with the instruction to fake depression (no honest controls were recruited for this study). One of the four subsamples included in Giromini, Viglione, et al. (2019) consisted of 100 healthy, Italian volunteers instructed to take the IOP-29 three times: one time responding honestly, one time faking depression, and one time responding at random. Finally, Giromini, Carfora Lettieri et al. (2019) reported on 62 Italian depressed patients – 26 came from real-life evaluations and had been deemed to be genuinely affected by depression, and 36 were psychiatric patients being treated for depression-related diagnoses – and 93 healthy, Italian volunteers instructed to feign depression. The findings from our Lithuanian study were thus compared against those from the five studies from the US, Italy and Portugal described above.

Results

Simulators scored notably higher than patients on both the IOP-29 and CES-D, with the differences between the two groups being more extreme for the IOP-29 (Table 1). More specifically, Cohen's *d* effect size was $d = 3.31$ for the IOP-29, and $d = .96$ for the CES-D. Based on Rogers et al.'s (2003) characterization of *d* values in experimental feigning studies, the IOP-29 produced a "very large" effect size (i.e., ≥ 1.75), whereas the CES-D generated a "moderate" effect size (i.e., $\geq .75$). Area under the curve (AUC) was .98 (*SE* = .01) for the IOP-29 and .77 (*SE* = .05) for the CES-D (Figure 1).

When looking at the distribution of IOP-29 (Figure 2) and CES-D (Figure 3) scores, it becomes evident that while simulators scored notably higher than patients, on average, on both tests, the degree of overlap between the two groups was dramatically reduced when considering the IOP-29. Figure 3 also shows that all participants, i.e., both bona fide patients and

URL: http://mc.manuscriptcentral.com/hapn Email: drmachorton@hotmail.com

For Peer Review Only

experimental simulators, had a CES-D score above the recommended cut score for identifying clinically meaningful depressive symptoms, i.e., CES-D Total ≥ 16 (Radloff, 1977).

With regard to diagnostic efficiency statistics, Table 2 shows that – as expected – the standard IOP-29 cut score of FDS ≥ .50 yielded the best balance between sensitivity and specificity, maximizing the overall correct classification (OCC) rate. The a-priori selected cut scores deemed to maximize sensitivity, i.e., FDS ≥ .30 and FDS ≥ .15, yielded in this study sensitivity values of .98 and 1.00 respectively. Along the same lines, the a-priori selected cut scores deemed to maximize specificity, i.e., FDS ≥ .65 and FDS ≥ .70, yielded in this study specificity values of .98 and 1.00 respectively.

Table 3 shows how our results compare to previously published, empirical data informing on the validity of the IOP-29 in discriminating feigned from bona fide depression. All in all, the results from our study are comparable, but perhaps slightly more encouraging than previous empirical data on the same topic. Indeed, across all reviewed studies, Cohen's *d* ranged from 1.80 to 4.32, and AUC ranged from .89 to .99. Giromini, Carfora Lettieri, et al. (2019) reported the least satisfactory results; the most satisfactory results were reported by Giromini, Viglione, et al. (2019). Our study produced the second most encouraging results.

Discussion

This study aimed to develop and validate a Lithuanian IOP-29. To test its ability in discriminating experimentally feigned versus bona fide depression, a clinical comparison, simulation/analogue study design was used. The results are very promising as they replicate the findings of previous research on the validity of the IOP-29 and thus contribute to establishing its cross-cultural adaptability and generalizability. They also confirm the applicability of the IOP-29 in a Lithuanian context specifically.

URL: http://mc.manuscriptcentral.com/hapn Email: drmachorton@hotmail.com

For Peer Review Only

Lithuanian Validation of the IOP-29

The Lithuanian IOP-29 created and developed in this study indeed discriminated almost perfectly between feigned and genuine depression, with a satisfactory AUC (.98), a very high Cohen's *d* (3.31) and excellent diagnostic efficiency statistics (OCC = .95, when using the standard cut score of IOP-29 FDS ≥ .50). Its specificity (.96 for IOP-29 FDS ≥ .50) was very high despite the fact that the patients with depression suffered from mild to severe depression according to the CES-D (M = 39.4, SD = 6.7). Its sensitivity (.94 for IOP-29 FDS ≥ .50) was notably high even though simulators were given a vignette to help them feign the symptoms of depression in a realistic way and were specifically warned "not to overdo it."

To our knowledge, this is the sixth article reporting on the validity of the IOP-29 in discriminating feigned from bona fide depression and so its use in applied settings can be seen to be well supported. According to Rogers et al (2003), Cohen's *d* values in simulation/analogue studies may be characterized as "moderate" when $d \geq .75$, "large" when $d \geq 1.25$, and "very large" when $d \geq 1.75$. Across the six studies reporting on feigned depression summarized in Table 3, Cohen's *d* ranged from 1.80 to 4.30, thus showing "very large" effect sizes in all cases. Along the same lines, Hosmer and Lemeshow (2000) suggested that AUC scores above 0.80 should be characterized as "excellent", and all the six studies published so far achieved an AUC ≥ .89. Taken together, these results contribute to establishing the excellent validity of the IOP-29 in discriminating feigned from bona fide depressive presentations.

It is worth mentioning, however, that there is some level of variability from one study to another (Table 3). The fact that the highest effect size was found in Giromini, Viglione, et al. (2019) is not surprising, given that their control group was comprised of healthy volunteers and it is known that using healthy volunteers as controls may boost the effect size (Rogers & Bender, 2018; van Impelen, Merckelbach, Jelicic, & Merten, 2014). The other studies produced relatively

URL: http://mc.manuscriptcentral.com/hapn Email: drmachorton@hotmail.com

For Peer Review Only

similar results, with the exception of Giromini, Carfora Lettieri et al. (2019), which showed less encouraging results. We may speculate that using a lengthy test that requires a great deal of cognitive effort (i.e., the MMPI-2 which contains 567 items) may reduce the compliance of test-takers. As a result, when they take the IOP-29, they are perhaps either fatigued or preoccupied that they will have to answer a very large number of questions, and thus may not complete it with their full level of attention. Alternatively, we may consider that when other long tests are included in the assessment, experimental simulators 'disperse' their feigning efforts across all available tests. Note that these are mere speculations and that additional studies are needed to clarify whether the IOP-29 does indeed perform better when used alone. If this is found to be the case, one might recommend using it first when undertaking a multimethod symptom validity assessment.

This study is not without its limitations. Although we did attempt to maximize ecological validity by providing participants with symptoms and a scenario and a cautionary statement "not to overdo it", there was no strong incentive to feign without being detected (the best feigners would win a free coffee card), which limits the ecological validity of this study. Additionally, experimental feigning of depression symptoms may be different from real-life feigning which could also affect ecological validity. We also cannot rule out the possibility that some of the patients exaggerated their symptoms, although this is perhaps unlikely given that there was no incentive to do so. Furthermore, the sample size is relatively small and therefore the results may not be generalizable. Finally, differently from typical real-life forensic evaluations, no other SVT's or PVT's were used alongside the IOP-29. With this regard, however, it should be noted that the lack of available measures for use with Lithuanian populations is exactly one of the reasons why we undertook this project.

For Peer Review Only

Despite these limitations, this study adds to the emerging research foundation for using the IOP-29 in the detection of noncredible psychological disorders, contributes to supporting its cross-cultural applicability, and facilitates further research on malingering in Lithuania due to the fact that a brief and easy-to-administer test has now been created and validated.

For Peer Review Only

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA, US: American Psychiatric Publishing, Inc.

Bagby, R. M., Nicholson, R. A., Buis, T., & Bacchiochi, J. R. (2000). Can the MMPI-2 validity scales detect depression feigned by experts? *Assessment, 7*(1), 55-62. doi:10.1177/107319110000700104

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI-2-RF (minnesota multiphasic personality inventory-2 restructured form) manual for administration, scoring, and interpretation*. Minneapolis: University of Minnesota Press.

Brislin, R. W. (1980). Translation and content analysis of oral and written material. In H. C. Triandis, & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (pp. 389-444). Boston, MA: Allyn & Bacon.

Bush, S., Heilbronner, R., & Ruff, R. (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. doi:10.1007/s12207-014-9198-7

Chafetz, M., & Underhill, J. (2013). Estimated costs of malingered disability. *Archives of Clinical Neuropsychology : The Official Journal of the National Academy of Neuropsychologists, 28*(7), 633-639. doi:10.1093/arclin/act038

Cuijpers, P., & Smit, F. (2008). Subclinical depression: A clinically relevant condition? *Tijdschrift Voor Psychiatrie, 50*(8), 519. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18688776

For Peer Review Only

Druss, B. G., Rosenheck, R. A., & Sledge, W. H. (2000). Health and disability costs of depressive illness in a major U.S. corporation. *American Journal of Psychiatry, 157*(8), 1274-1278. doi:10.1176/appi.ajp.157.8.1274

Geisinger, K. F. (2003). Testing and assessment in cross-cultural psychology. In J. R. Graham, & J. A. Naglieri (Eds.), *Handbook of psychology: Assessment psychology* (pp. 95-117). Hoboken, NJ, US: John Wiley & Sons Inc.

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,* , 1-13. doi:10.1080/23279095.2019.1570929

Giromini, L., Lettieri, S. C., 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,* doi:10.1007/s12207-019-09357-7

Giromini, L., Viglione, D. J., Pignolo, C., & Zennaro, A. (2019). An inventory of problems-29 sensitivity study investigating feigning of four different symptom presentations via malingering experimental paradigm. *Journal of Personality Assessment,* , 1-10. doi:10.1080/00223891.2019.1566914

Giromini, L., Viglione, D., 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. doi:10.1007/s12207-018-9314-1

Greene, R. L. (1991). *The MMPI-2/MMPI : An interpretive manual*. United States: Retrieved from http://catalog.hathitrust.org/Record/002501124

For Peer Review Only

Hosmer, D.W. and Lemeshow, S. (2000) *Applied logistic regression.* 2nd Eds, John Wiley & Sons, Inc., New York. doi:10.1002/0471722146

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. doi:10.1002/1097-4679(199403)50:23.0.CO;2-T

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24*(8), 1094. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12650234

Monaro, M., Toncini, A., Ferracuti, S., Tessari, G., Vaccaro, M. G., De Fazio, P., . . . Sartori, G. (2018). The detection of malingering: A new tool to identify made-up depression. *Frontiers in Psychiatry, 9*, 249. doi:10.3389/fpsyt.2018.00249

Morey, L. C. (1991). *Personality assessment Inventory–Professional manual*. Odessa, FL: Psychological Assessment Resources.

Nicholson, K., & Martelli, M. (2007). Malingering: Posttraumatic stress disorder and depression. *Causality of psychological injury* (pp. 501-508). Boston, MA: Springer US. doi:10.1007/978-0-387-36445-2_18

Radloff, L. S. (1977). The CES-D scale. *Applied Psychological Measurement, 1*(3), 385-401. doi:10.1177/014662167700100306

Rogers, R. (2010). *Structured Interview of Reported Symptoms (SIRS). 10.1002/9780470479216.corpsy0957.*

Rogers, R., & Bender, S. D. (2018). *Clinical assessment of malingering and deception* (Fourth ed.) The Guilford Press. Retrieved from

For Peer Review Only

http://www.vlebooks.com/vleweb/product/openreader?id=none&isbn=9781462533503&uid=none

Rogers, R., Gillis, J. R., Dickens, S. E., & Bagby, R. M. (1991). Standardized assessment of malingering. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*(1), 89-96. doi:10.1037/1040-3590.3.1.89

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*(2), 160-177. doi:10.1177/1073191103010002007

Sharf, A., Rogers, R., Williams, M., & Henry, S. (2017). The effectiveness of the MMPI-2-RF in detecting feigned mental disorders and cognitive deficits: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment, 39*(3), 441-455. doi:10.1007/s10862-017-9590-1

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the structured inventory of malingered symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law Online, 25*(2), 183. Retrieved from http://www.jaapl.org/content/25/2/183.abstract

Steffan, J. S., Clopton, J. R., & Morgan, R. D. (2003). An MMPI-2 scale to detect malingered depression (md scale). *Assessment, 10*(4), 382-392. doi:10.1177/1073191103259548

Tombaugh, T. N. (1996). *Test of memory malingering (TOMM)*. New York: Multi-Health Systems, Inc.

Tombaugh, T. N. (1997). The test of memory malingering (TOMM). *Psychological Assessment, 9*(3), 260-268. doi:10.1037/1040-3590.9.3.260

For Peer Review Only

Lithuanian Validation of the IOP-29

Van de Vijver, F., & Hambleton, R. K. (1996). Translating tests. *European Psychologist, 1*(2), 89-99. doi:10.1027/1016-9040.1.2.89

van Impelen, A., Merckelbach, H., Jelicic, M., & Merten, T. (2014). The structured inventory of malingered symptomatology (SIMS): A systematic review and meta-analysis. *The Clinical Neuropsychologist, 28*(8), 1336-1365. doi:10.1080/13854046.2014.984763

Viglione, D. J., Giromini, L., & Landis, P. (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. doi:10.1080/00223891.2016.1233882

Winters, C. (2019). *A british validation of the inventory of problems - 29 (IOP-29) investigating feigned schizophrenia and random responding;* Unpublished manuscript.

Young, S., Jacobson, R., Einzig, S., Gray, K., & Gudjonsson, G. H. (2016). Can we recognise malingerers? the association between malingering, personality traits and clinical impression among complainants in civil compensation cases. *Personality and Individual Differences, 98*, 235-238. doi:10.1016/j.paid.2016.04.052

For Peer Review Only

Lithuanian Validation of the IOP-29

Table 1. IOP-29 and CES-D Scores of Depressed Patients and Experimental Simulators

	Patients (n = 50)		Simulators (n = 50)		t	df	p	d
	M	SD	M	SD				
IOP-29 FDS	.27	.12	.77	.18	16.56	85.5[a]	< .001	3.31
CES-D Total	39.4	6.7	45.3	5.6	4.79	98.0	< .001	.96

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

For Peer Review Only

Lithuanian Validation of the IOP-29

Table 2. Classification Accuracy for Five A-Priori Selected IOP-29 Cut Scores

Cut Score	Se	Sp	OCC
IOP-29 FDS ≥ .70	.68	1.00	.84
IOP-29 FDS ≥ .65	.74	.98	.86
IOP-29 FDS ≥ .50	.94	.96	.95
IOP-29 FDS ≥ .30	.98	.72	.85
IOP-29 FDS ≥ .15	1.00	.16	.58

Se = Sensitivity; Sp = Specificity; OCC = Overall Correct Classification.

For Peer Review Only

Table 3. Validity of the IOP-29 in Discriminating Feigned from Bona Fide Depression-related Complaints across Studies

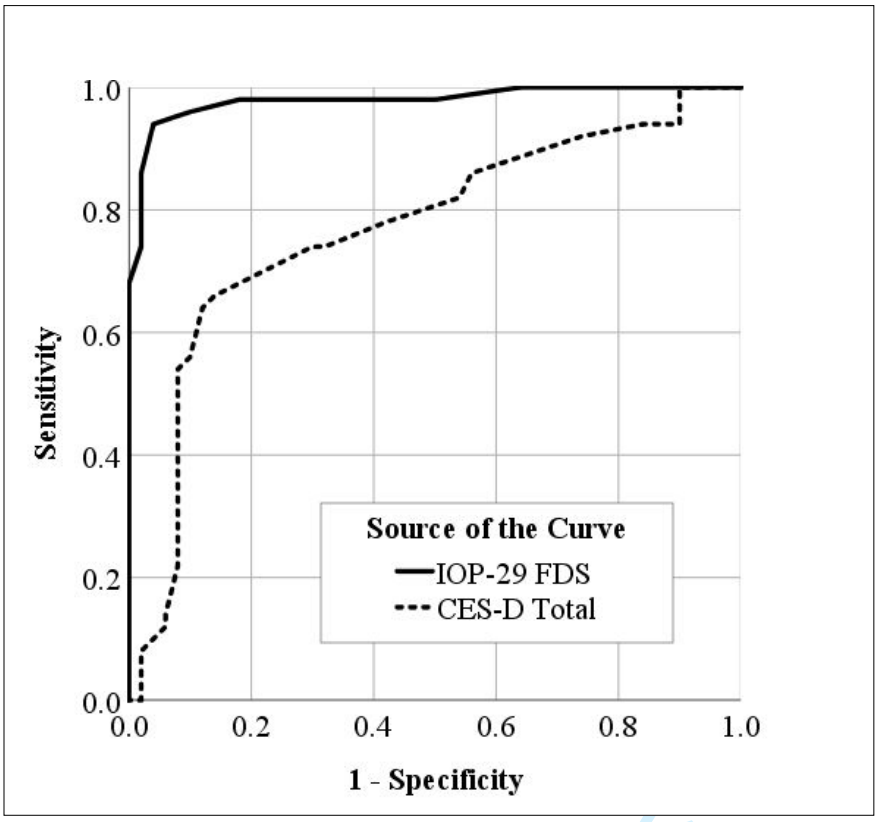
	Viglione et al. (2017)[a]	Giromini et al. (2018)[b]	Giromini, Barbosa, et al. (2019)[c]	Giromini, Viglione, et al. (2019)[d]	Giromini, Carfora Lettieri, et al. (2019)	This study
Experimental simulators	42	111	50	100	93	50
Honest controls	43	127	0	100	62	50
Controls characterization	Patients	Patients	-	Healthy volunteers	Patients	Patients
IOP-29 version	English	Italian	Portuguese	Italian	Italian	Lithuanian
Other relevant test(s)	TOMM	SIMS	TOMM	-	MMPI-2	CES-D
Se for IOP-29 FDS ≥ .50	.84	.81	.96	.95	.75	.94
Sp for IOP-29 FDS ≥ .50	.86	.83	-	.96	.87	.96
Cohen's d	1.97	2.02	-	4.32	1.80	3.31
AUC	.90	.90	-	.99	.89	.98

[a] These data refer to Abramsky's (2005) depression subsample described in Study 2 of Viglione et al.'s (2017) article. [b] These data refer to the anxiety, depression, and/or trauma-related subsample described in Giromini et al.'s (2018) article. [c] These data refer to the depression-related condition described in Giromini, Barbosa et al.'s (2019) article. [d] These data refer to the depression-related subsample of Giromini, Viglione et al.'s (2019) article: this study used a within-subject design, in which participants were asked to take the IOP-29 three times, one time answering honestly, one time faking depression, and one time responding with a random-like approach.

For Peer Review Only

Lithuanian Validation of the IOP-29

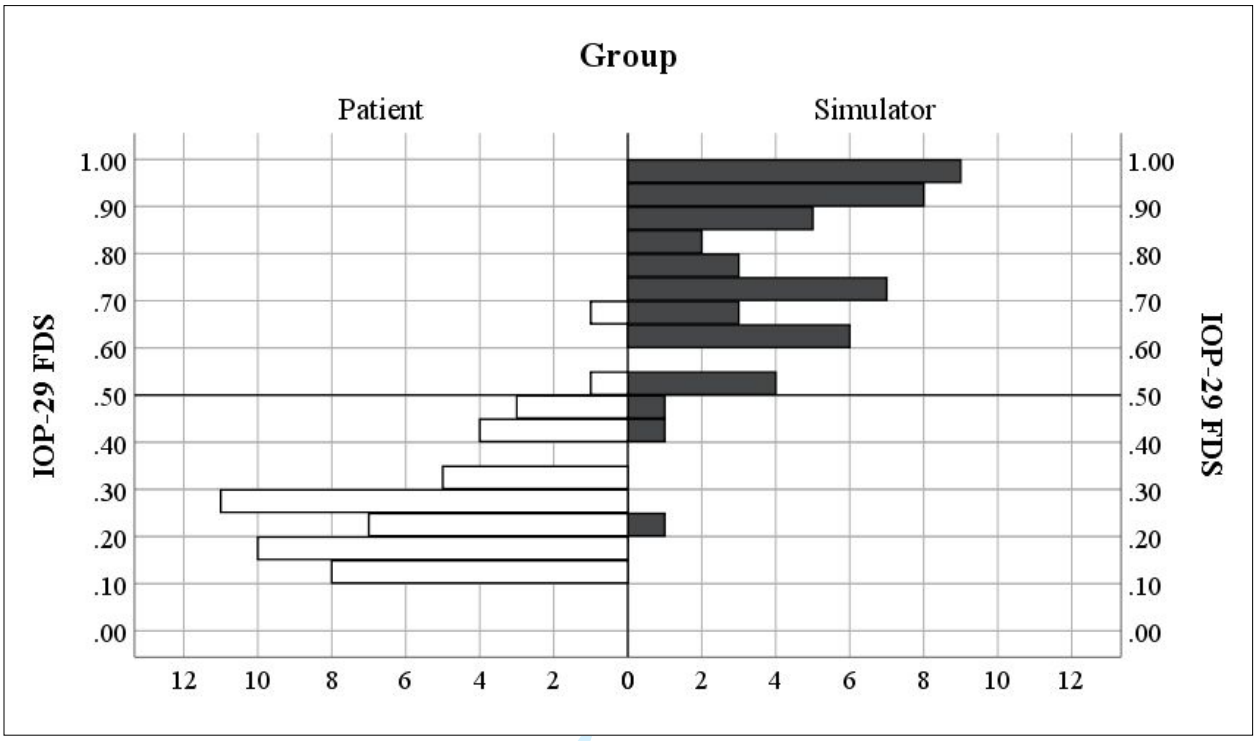
Figure 1. Graphical Representation of Receiver Operating Characteristic (ROC) Curves



For Review Only

Lithuanian Validation of the IOP-29

Figure 2. Distribution of IOP-29 FDS Scores by Group

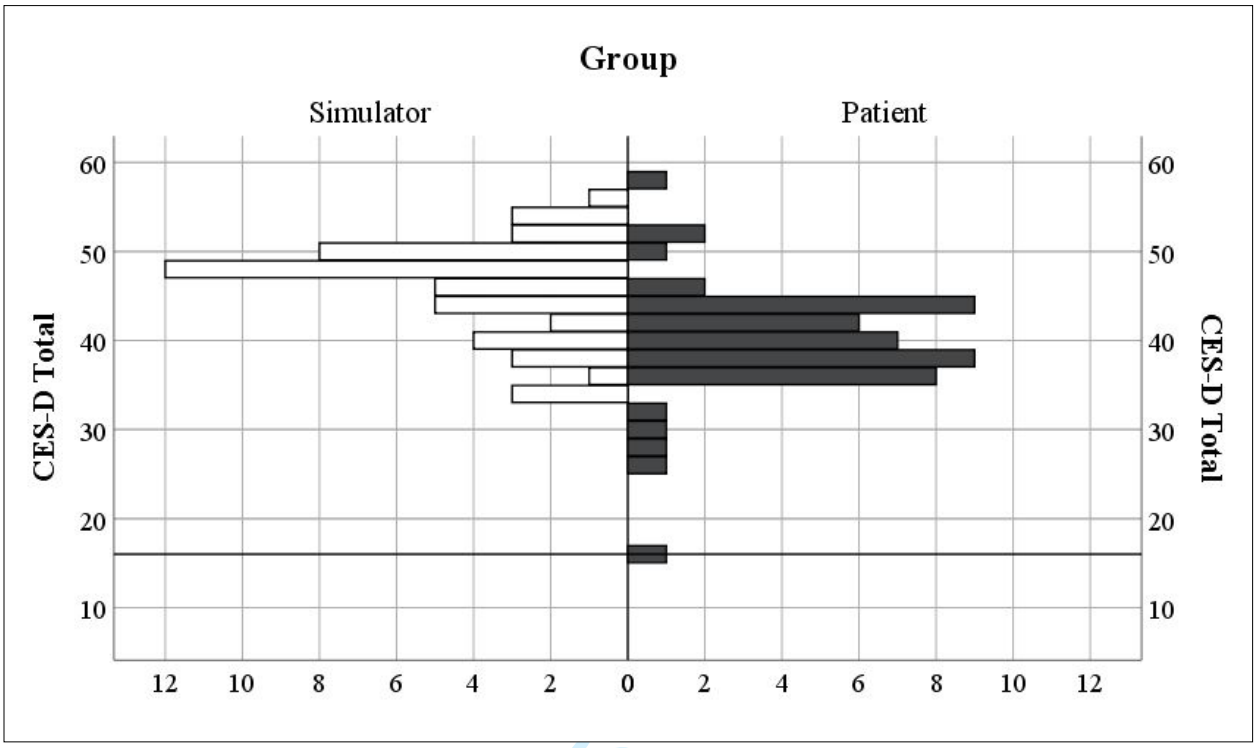


Note. The reference line in the X-axis corresponds to the IOP-29 FDS value of .50.

For Review Only

Lithuanian Validation of the IOP-29

Figure 3. Distribution of CES-D Scores by Group



Note. The reference line in the X-axis corresponds to the CES-D value of 16.

For Review Only