Abstract

Compared to other Western countries, malingering research is still relatively scarce in the United Kingdom, partly because only a few brief and easy-to-use symptom validity tests (SVTs) have been validated for use with British test-takers. This online study examined the validity of the recently introduced Inventory of Problems – 29 (IOP-29; Viglione, Giromini, & Landis, 2017) in the detection of feigned schizophrenia and random responding in 151 British volunteers. Each participant took three IOP-29 test administrations: (a) responding honestly; (b) pretending to suffer from schizophrenia; and (c) responding at random. Additionally, they also took the O-LIFE questionnaire of schizotypy under standard instructions (i.e., responding honestly). The chief feigning scale of the IOP-29 (FDS) showed excellent validity in discriminating honest responding from feigned schizophrenia (AUC = .99), and its classification accuracy was not significantly affected by the presence of schizotypal traits. Additionally, a recently introduced IOP-29 scale aimed at detecting random responding (RRS) also demonstrated very promising results.

Keywords: British; Feigning; Inventory of Problems; IOP-29; Malingering; O-LIFE; Online; Random Responding; Schizophrenia; Schizotypy; Validity.

An Inventory of Problems – 29 (IOP–29) Study Investigating Feigned Schizophrenia and Random Responding in a British Community Sample

Malingering is an intentional feigning or exaggeration of symptoms in order to gain external incentive (American Psychiatric Association, 2013). Feigning illness or disability is costly to society as resources are displaced away from people who are genuinely ill. Malingering should therefore be considered a possibility every time an individual may gain from presenting as impaired (Binder, 1993).

To evaluate the possible presence of malingering, forensic assessors typically rely on multiple sources of information (Boone, 2013). In addition to clinical interviews and collateral information, psychological tests are a rich source of information which assessors can rely on in order to derive judgment. These tests are often grouped into two major categories: symptom (SVT) and performance (PVT) validity tests. The former refers to tests aimed at evaluating the credibility of self-reported psychological difficulties or problems, the latter refers to tests aimed at evaluating the credibility of scores on cognitive tests. SVTs and PVTs, however, can only inform on the level of validity/credibility of a given presentation; they cannot tell whether an invalid/non-credible clinical presentation is feigned for an external versus internal motivation (van Impelen, Merckelbach, Jelicic, & Merten, 2014). As such, neither SVTs nor PVTs, per se, measure malingering. In line with Rogers and Bender (2013), in this article we thus refer to *malingering* to indicate the "deliberate fabrication or gross exaggeration of psychological or physical symptoms for the fulfilment of an external goal," and *feigning* to indicate the "deliberate fabrication or gross exaggeration of psychological or physical symptoms (Rogers & Vitacco, 2002) without any assumptions about its goals" (p. 518).

Most SVT research and resulting base rates of non-credible symptom have come from the United States (Martin, Schroeder, & Odland, 2015; Young, 2014, 2015). Comprehensive meta-analyses of various forensic assessment studies point to base rates of 15±15% for malingering (see Young 2015), although non-credible presentations seem to occur at higher base rates, possibly around 40% (Larrabee, 2003), in neuropsychological assessment (see Young, 2014), and at an even higher than 50% base rate in medico-legal disability claimants and forensic criminal cases, especially if validated screens such as the Miller Forensic Assessment of Symptoms Test (M-FAST; Miller, 2001) are used (Rogers & Bender, 2018). Less is known concerning base rates of malingering and non-credible symptom reporting in the UK, as there is not widespread use of the psychological tests (McCarter, Walton, Brooks, & Powell, 2009), and this is in large part due to the division of medical and legal systems of care (Halligan, Bass, & Oakley, 2003). More specifically, the UK substantially lacks symptom validity assessment research (Merten et al., 2013), which notably limits the possibility to investigate malingering-related phenomena within this cultural context.

Symptom Validity Assessment in the UK

While SVTs are very commonly used in the US, a less stringent approach is taken in the UK, in large part due to British Psychological Society (BPS) caution against using these instruments (McMillan et al., 2009), especially when there is initially no forensic context for treatment. As the UK approaches mental healthcare treatment and forensic rehabilitation in highly centralized systems, use of SVTs in clinical clients within a medico-legal context (e.g., claimants) is a grey area fraught with potential systemic complications.

In a review of symptom validity practices in European countries, Merten and colleagues (2013) note a paucity of SVT research in Great Britain, specifically a lack of litigant studies, a

lack of studies evaluating chronic pain, and a lack of studies in the context of criminal forensic neuropsychological assessments. This is reflected in UK clinical practice as "few psychologists [provide] these specialist assessments" (p. 135). In a review of both academic and government statistics, there is a scarcity of information concerning base rates of cognitive impairment (McMillan et al., 2009), suspected rates of feigning of specific disorders, and fraudulent medico-legal claims in the UK. One of the biggest issues concerning SVT research in the UK is a general lack of reported base rates for non-credible responding, however there is no data suggesting UK rates would be dramatically different from those reported in the US, and there is certainly not reason to believe rates of non-credible responding would be lower in the UK. For medico-legal disability claimants and forensic criminal cases, the BPS points to the US as a guide for base rates of malingering (McMillan et al., 2009), at approximately 54 to 72% (see Miller, Ryan, Carruthers, & Cluff, 2004; Chafetz, 2008), and approximately 54% respectively (see Ardolf, Denney, & Houston, 2007).

In a self-selected survey of 91 British neuropsychologists practicing in medico-legal clinical cases (McCarter et al., 2009), only 7% reported they viewed SVTs as mandatory, and only 13% of them reported using SVTs most of the time (>95%). Top reported reasons for not using SVTs included: invalidity is obvious in presentation (38%), invalidity is obvious in (other) test scores (38%), insufficient time (35%), and the belief that few patients exaggerate (34%). An outdated reliance on clinical intuition and the belief that most clients were genuine was largely regarded as the reason that most of these experts did not use SVTs, and the authors acknowledged that this finding, in conjunction with varying approaches, frequency of use, and measures, was likely to significantly bias attempts to report base rates of malingering. Currently, practitioner and community-based whistleblowing via the NHS Counter Fraud Authority serves

as the main system for combating patient abuse of services (Department of Health & Social Care, 2020).

Despite the paucity of information concerning base rates of non-credible responding, fraudulent medico-legal claims are becoming an increasing issue in the UK (McCarter et al., 2009). The UK Department for Work and Pensions (2019) reported overpaying £4.1 billion in welfare benefits in 2018-2019, as fraudulent overpayments have jointly been awarded at the highest estimated level (1.2%) steadily since 2016-2017. The UK Disability Unit (2020) offers several financial benefits for individuals affected by long term (i.e., if it is likely to last 12 months) mental health problems, and schizophrenia is listed one of these stated conditions. To promote research in this area, it would be beneficial to validate brief and easy-to-use SVT like the Inventory of Problems – 29 (IOP-29; Viglione & Giromini, 2020; Viglione, Giromini, & Landis, 2017) for use with a British population.

Schizophrenia and Schizotypy

Schizophrenia and its associated symptoms are among the more commonly feigned psychiatric complaints in criminal forensic contexts (see Pierre, Shnayder, Wirshing, & Wirshing, 2004). According to the World Health Organization classification (WHO, 2008) schizophrenia is one of the most severe disabilities. The DSM-5 considers schizophrenia as a spectrum disorder, which includes delusions, hallucinations, and/or disorganized speech, and can also include grossly disorganized behavior, catatonic behavior, or negative symptoms. Similarly, the ICD-11 contains a section on Schizophrenia and other primary psychotic disorders, which are characterized by significant impairments in reality testing and alterations in behavior. These symptoms manifest as positive symptoms (i.e., changes in behavior or thoughts), such as persistent delusions, persistent hallucinations, disorganized thinking (typically manifest as

disorganized speech), grossly disorganized behavior, experiences of passivity and control, and negative symptoms (i.e., withdrawal or lack of function), such as blunted or flat affect, avolition, and psychomotor disturbances (WHO, 2018).

Schizotypy is a psychological construct that is intimately connected to latent schizophrenia-related liability and symptomatology (Meehl, 1962; Meehl, 1994; Lezenweger, 2006; Rado, 1960). Over the years there has been a debate over the measure of schizotypic psychopathology, and to what extent schizotypy is helpful in determining risk of the development of schizophrenia or psychotic-related disorders (Grant et al., 2013; Lenzenweger, 2015). The DSM-5 describes Schizotypal Personality Disorder (STPD) as a "pervasive pattern of social and interpersonal deficits marked by acute discomfort with, and reduced capacity for, close relationships as well as by cognitive or perceptual distortions and eccentricities of behavior, beginning by early adulthood and present in a variety of contexts" (APA, 2013, p. 655). In contrast, the ICD-11 does not include STPD as a diagnosis, as its features are coded within Schizophrenia and other primary psychotic disorders. Instead, the ICD-11 classifies Personality Disorder severity in terms of whether the patient experiences "dissociative states or psychotic-like beliefs or perceptions" and its diagnostic approach is conceptualized by the capacity for reality testing (WHO, 2018). Schizotypy as a dynamic and latent pathological construct is more closely aligned with the ICD-11's conceptualization of its features (i.e., unusual experiences), in which they serve to moderate a given diagnosis of either Schizophrenia or a Personality Disorder, as opposed to being defined as a standalone diagnosis (STPD) in the DSM-5 (Kirchner, Roeh, Nolden, & Hasan, 2018). Broadly speaking, psychotic disorders can be conceptualized as a spectrum with schizotypy at the less severe end and schizophrenia at the more severe end (Claridge & Beech, 1995). There is a growing body of evidence that suggests

schizotypy shares a common biological basis with schizophrenia as defined by both genetic susceptibility and pathological processes related to dopamine dysregulation (Avramopoulos et al., 2002; Grant et al., 2013; Lachman et al., 1996; Smyrnis et al., 2007; Vandenbergh et al., 1992).

The Inventory of Problems – 29 (IOP-29)

To assist practitioners in evaluating the credibility of psychological and cognitive disorder presentations, a particularly promising tool was introduced in 2017. Comprised of only 29 self-administered items and named the "Inventory of Problems – 29" (IOP–29; Viglione & Giromini, 2020; Viglione, Giromini, & Landis, 2017), it differs from most SVTs in five ways:

1) It focuses on the manner in which purported symptoms are presented, as opposed to the presence or absence of atypical versus bona fide symptoms.

2) It intermixes self-report and cognitive (e.g., calculation, logic) items, so it is applicable for both psychiatric and cognitive complaints.

3) In addition to the typical "True" and "False" response options, self-report items also offer a third option: "Doesn't make sense", allowing the test-taker to indicate that the question is unanswerable or awkwardly stated. This trichotomous response choice also allows each item to be scored positively or negatively for more than one response choice.

4) The IOP-29 does not use a T-score metric based on a single set of normative reference data obtained from healthy volunteers. Instead, the IOP-29 standardized score used for interpretation, the False Disorder Probability Score (FDS) is based on the comparison of the test-taker's responses against two different sets of reference values, one coming from bona fide patients, and the other one coming from experimental simulators. A logistic regression-derived formula generates the False Disorder Probability Score which

establishes the statistical probability that a given IOP-29 comes from valid versus invalid symptom presentation. Greater FDS scores are associated with non-credible presentations and lower scores are associated with credible presentations. As a probability score it ranges from zero to one. Without a priori expectation, the FDS cut-off score is ≥ .50.

5) Indeed, the fifth critical distinction between the IOP-29 and typical SVTs is that this cut-off has been stable across schizophrenia and psychosis, depression, PTSD, and mild cognitive disorders in all the research on the text as summarized in the test manual (Viglione & Giromini, 2020).

Viglione et al.'s (2017) initial clinical comparison simulation studies conducted in the US (which compared experimental feigners to bona fide patients) showed that the IOP-29's classification accuracy was similar to that of other symptom validity measures, including the MMPI-2 and PAI validity scales, with sensitivity and specificity values of about .80 for FDS ≥ .50. Further, an Italian clinical comparison study demonstrated that the IOP-29 outperformed the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997; Widows & Smith, 2005), with the greatest effect sizes (patients versus experimental feigners) between the two tests found in psychotic spectrum disorders-related presentations (*d* (IOP-29) = 1.80 vs. *d*(SIMS) = 1.06; AUC(IOP-29) = .89 vs. AUC(SIMS) = .79). More recently, studies conducted in Portugal (Giromini, Barbosa et al., 2019) and Italy (Giromini, Carfora Lettieri et al., 2019) showed that the IOP-29 yielded incremental validity over the Test of Memory Malingering (TOMM; Tombaugh, 1996) and Minnesota Multiphasic Personality Inventory (MMPI-2; Butcher et al., 1989, 2001), respectively. Additionally, the cross-cultural adaptability of the IOP-29 has been recently demonstrated also in Lithuania (Ilgunaite, Giromini, Bosi,

Viglione, & Zennaro, 2020). To our knowledge, however, no study has yet examined the validity of the IOP-29 FDS with a British population.

Random Responding

Like other SVTs, the IOP-29 may be susceptible to the effects of random or inattentive responding, which might occur either because a test taker did not understand the meaning of the items or because they somehow did not cooperate with the testing situation. Random responding and malingering are both considered to be invalid response styles, and both response styles may at times produce overstated pathology and suboptimal performance on cognitive items and/or neuropsychological tests (Rogers, 2008). However, unlike malingering, random responding is not characterized by deception or an intention to deceive the examiner. Instead, it involves responding without paying proper attention or without really understanding the meaning of the item(s) (Nichols, Greene, & Schmolck, 1989). Alternatively, it also might arise from resistance to the testing in the form a purposeful attempt to avoid disclosing information to the examiner. Partial resistance might emerge in the form of discontinuing effort and cooperation after initially attempting to answer questions honestly. When an SVT includes items describing rare symptoms or unlikely behaviors and attitudes, a bona fide responder may inadvertently endorse these items, as pathological individuals are more likely than non-pathological individuals to endorse rare complaints on these types of tests (Greiffenstein & Baker, 2008; Rogers & Bender, 2018; Slick, Sherman, Grant, & Iverson, 1999). Consequently, if the response-pattern appears random-like, test scales which address overreporting of symptoms and problems may be artificially inflated (Burchett et al., 2016).

Both overreporting and random responding can co-occur and interact, and they can both involve “inconsistent” and “infrequent” responding (e.g., Morey, 1991). Inconsistent responding

is endorsing contradictory items whereas infrequent responding is choosing response options which are rarely selected by others. Practitioners, however, should try to discriminate whether a non-credible symptom presentation is caused by random responding versus feigning. As such, the inclusion of embedded measures of task engagement within SVTs are one way to help clinicians discriminate different types of non-credible response styles. To that goal, Giromini, Viglione et al. (2019b) have recently derived the IOP-29 Random Responding Scale (RRS), an IOP-29 index aimed at detecting random responding. In their developmental research, the IOP-29 RRS yielded promising results. However, no study has yet cross-validated Giromini, Viglione et al.'s (2019b) findings.

Current Study

The current study sought to evaluate the applicability of the IOP-29 to a British population, and to provide initial cross-validation data to evaluate the potential utility of the IOP-29 FDS and RRS. Additionally, we aimed at testing the extent to which the presence of schizotypal traits would influence IOP-29 FDS scores. More specifically, because clinical test-takers are known to score higher on SVTs than healthy test-takers do (Rogers & Bender, 2018; van Impelen et al., 2014), we intended to evaluate the extent to which individuals with higher schizotypal traits would score higher on the IOP-29 when answering honestly. Besides, we also wanted to explore whether specific schizotypal traits could influence one's ability to effectively feign schizophrenia without being detected by the IOP-29. Briefly stated, we wanted to evaluate whether greater levels of schizotypal traits could increase the likelihood of obtaining false negative and false positive classifications on the IOP-29.

We conducted an online study with a community sample, resembling the procedures followed by Giromini, Viglione, et al. (2019a) when testing the applicability of the IOP-29 FDS

to various symptom presentations. Participants took the IOP-29 three times, in three different conditions. In the honest condition (HON), they were instructed to respond honestly following standard instructions; in the simulation or feigning condition (SIM) they were coached to simulate schizophrenia; and in the random responding condition (RND) they were asked to respond randomly, with no apparent pattern to their responding. Additionally, all participants also took a brief measure of schizotypal traits under standard instructions, i.e., with the request to respond honestly.

We hypothesized that: 1) SIM condition would yield significantly greater FDS scores than HON, with large effect sizes; 2) individuals with higher schizotypal traits would score higher on FDS in condition HON in comparison to individuals with low schizotypal traits (as they are expected to show higher inconsistency in their responses) and perhaps lower on FDS in condition SIM (due to the overlapping symptomatology with schizophrenia itself); 3) RND condition would yield significantly higher RRS scores than both HON and SIM conditions, with no significant differences in RRS scores between HON and SIM.

Method

Participants

A British community sample made up of 151 adult volunteers (74.17% women), ranging in age from 18 to 59 (M = 25.79, SD = 9.33) participated in this online study. In terms of education, a little less than half of the sample (n = 67; 44.37%) completed high school (A-levels) or less, 28.48% completed an undergraduate degree (n = 43), 25.83% completed a postgraduate degree or more (n = 39), and two individuals endorsed the response option “Other.” Most of the sample (98.68%) spoke English as their native language (two were native in Russian). Inclusion criteria required literacy and the ability to provide informed consent. Participants who reported

using intoxicating substances (i.e., alcohol, drugs) that day were excluded from participating. Further criteria which prevented full participation included: experiencing visual or audio hallucinations in the last 30 days; being treated for substance abuse problems, a neurological disorder, or traumatic brain injury in the past six months; and any historical diagnosis of psychosis, schizophrenia, bipolar disorder, or schizoaffective disorder.

Prior to reaching the final sample size (N = 151), approximately 30 individuals were excluded from participating; and this was primarily due to admission of recent (<30 days) hallucinations, secondarily due to recent (<1 day) intoxicant use, and tertiarily due to incorrectly responding to control question (i.e., manipulation check; see below) which determined that the participant was not reading the administration instructions.

Materials

The Inventory of Problems-29 (IOP-29; Viglione & Giromini, 2020). The IOP-29 is a 29 item, self-administered test designed to evaluate the credibility of various clinical presentations. It includes 27 items with three response options: True/False/Doesn't make sense. Among them are 26 self -report items about emotional, ideational, social, or personal experiences and a verbal reasoning item in the form of an analogy. The test also includes two open-ended questions which require mathematical reasoning and calculations. For the purpose of this study, the test was administered three times, in which the test taker was asked to respond honestly (HON), randomly (RND), and experimentally feigning schizophrenia (SIM). FDS and RRS scores were generated for each participant in each test taking condition.

The Oxford-Liverpool Inventory of Feelings and Experiences: Short scale for measuring schizotypy (Mason & Claridge, 2006). The O-LIFE short scale is a 43-item self-administered test, which reliably measures multi-dimensional schizotypy. The measure is based

on Claridge's (1997) conceptualization of schizotypy as fully-dimensional, suggesting intra-individually static basis of personality-based traits, whereby (despite an absence of cut-off scores), high values suggest an increased risk of developing psychotic disorder. The O-LIFE short scale quantifies the endorsement (i.e., "Yes/No") of items loaded onto the following subscales (reported α in this sample): Unusual Experiences (α = .75), Cognitive Disorganization (α = .78), Impulsive Nonconformity (α = .62), and Introvertive Anhedonia (α = .68).

The O-LIFE was administered once, at the end of the study, and participants were asked to report honestly. It was included to evaluate whether individuals with higher schizotypal traits would be more likely to generate false positive and false negative classifications on the IOP-29, as noted above. In addition, because STPD and schizophrenia share a common biological basis, we were interested in exploring the associations between specific schizotypal traits and FDS response styles.

Procedure

Ethical approval was obtained from the relevant institutional review board. The first author advertised the online study via a UK university SONA research participation system and social media groups for UK-based research participation. Further, the authors encouraged snowball sampling, as they emailed the shareable study link with UK-based researchers. Participants were informed of the nature of the study before participating, and that they would be asked to take the same questionnaire three times - once, responding honestly, once responding randomly, and once responding as if they had schizophrenia. Participants were asked to not participate if they might be uncomfortable disclosing their mental health or substance use history. Participants received an information form and provided their electronic informed consent via Qualtrics.

First participants were asked demographic questions, including their highest level of acquired education (or equivalent) within the British system (GCSE's/O-level, A-levels, undergraduate degree, postgraduate degree, and other). They were then asked about their recent substance use and historical mental health issues. Participants were then instructed to take the same questionnaire three times, in three different ways - responding honestly, responding randomly, and responding as if they are faking schizophrenia. The order of the administration of the three different testing conditions was randomized across participants. In the standard instructions for responding honestly, participants were told to "respond honestly - that is, how [they felt] today - not faking, and not role playing." In the random responding instructions, participants were asked to respond "randomly, with no pattern." In the faking schizophrenia instructions, participants were asked to "respond as if you are faking schizophrenia (but without 'over-exaggerating' your presentation, not to look like a feigner)." As part of instructions for the faking condition, participants were provided with short scenario to read (see Appendix A) about a character who wishes to fake schizophrenia, and they were told to imagine themselves as that individual. The character in the story is motivated to fake a presentation of schizophrenia in order to mitigate impending financial hardship. The vignette included a link to the UK's National Health Service website on schizophrenia, and participants were encouraged to read about its symptoms. To incentivize participants to fake schizophrenia well, it was emphasized that the best three fakers who can "trick the psychologist into thinking [they] have schizophrenia," would win one of three £20 (~$30 USD) cash prizes.

After reading response instructions and immediately preceding the beginning of each IOP-29 test administration, participants were asked the following control question (manipulation check): "How should you answer the following 29 questions?" to ensure that they understood

how they were expected to respond (honestly, randomly, or faking schizophrenia). If a participant responded incorrectly to this control question in any condition, the study automatically ended, and their data were not analyzed (that is, all 151 participants included in the analyses responded accurately to the control questions).

At the end of the three IOP-29 administrations, participants were asked to honestly complete a questionnaire about their own thoughts, feelings, experiences, preferences (O-LIFE, short version). Data downloaded from Qualtrics and imported to SPSS and excel for use in R. De-identified participant IOP-29 data was imported to www.iop-test.com.

While we initially aimed to exclude participants from analysis who did not take at least 10 seconds to read the vignette, (as we expected that they would not take the study seriously), we decided to include these cases (n = 11) in analyses after examining the data for two reasons. First, many of these specific individuals were psychology students, and therefore it is possible that they may have skipped reading the vignette because they thought they had a good understanding of schizophrenia. Second, because post-hoc analyses revealed that including or excluding those 11 individuals would lead to virtually identical results.

Data Analysis

For the FDS, we only investigated the honest (HON) and simulated (SIM) conditions, as the FDS was designed to discriminate honest (credible) from feigned (non-credible) presentations. For the RRS, we inspected all three conditions, i.e., HON, SIM, and random (RND), as the RRS was designed to discriminate random responding from both honest and feigned responding. For FDS, we first report the results of a series of linear mixed effects models, conducted using the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015)[1].

[1] For access to the code, please contact author Lara Warmelink at l.warmelink@lancaster.ac.uk

Simple models were built first, starting with adding the effect of condition, then adding O-LIFE scores, and then adding the effect of education. Where a newly added variable led to a significant improvement in the model, interaction effects were also tested. All models included a random effect of participant, to account for the repeated measures nature of the data. Models were compared using the ANOVA function. Graphs were extracted using the effects package (Fox & Weisberg, 2019). For RRS, we only used a linear mixed effect model to look at the effect of condition.

For both FDS and RRS, we then report Cohen's *d*, receiver operator characteristic curve (AUC), sensitivity, and specificity. With regard to Cohen's *d* effect size, in line with Dunlap, Cortina, Vaslow, and Burke's (1996) recommendations, we calculated it using standard independent samples *d* formula (1988) rather than Morris and DeShon's (2002) corrected value, as we were interested in calculating the actual effect size as opposed to an a priori power calculation. Lastly, because the IOP-29 RRS was designed to measure random responding while remaining independent from the IOP-29 FDS, the correlation between these two IOP-29 indexes was tested too.

Results

Effectiveness of the IOP-29 FDS

Descriptive statistics for all IOP-29 and O-LIFE scores included in the analyses are presented in Table 1. With regard to the effectiveness of the IOP-29 FDS, there was a significant main effect of condition in the first model, *estimate* = 0.67, *SE = 0.02*, *t* = 38.47. The scores of the IOP-29 FDS were indeed remarkably higher in condition SIM (*M* = .82; *SD* = .18) than in condition HON (*M* = .14; *SD* = .14), with a very large Cohen's *d* of 4.20 (see also Figure 1). Adding the four subscales of the O-LIFE (Unusual Experiences, Cognitive Disorganization,

Introvertive Anhedonia and Impulsive Nonconformity) was a significant improvement in the model, Chi^2 (4)= 29.42, p = 0.000006 (Table 2). However, only Introvertive Anhedonia was a significant predictor (Table 3). A model with only Introvertive Anhedonia was not significantly different from a model with all O-LIFE subscales, Chi^2 (3) = 2.72, p = 0.44. Therefore, only Introvertive Anhedonia was taken forward into more complex models.

Adding an interaction between condition and Introvertive Anhedonia did not improve the model over just including the main effects, Chi^2 (1) = 2.58, p = 0.11 (Table 2). Since Introvertive Anhedonia is a positive predictor of FDS, this may lead to concern that the FDS is less accurate in people with high Introvertive Anhedonia. However, there is no evidence to suggest this. Figure 2 contains the non-significant interaction effect between condition and Introvertive Anhedonia. It shows that FDS scores in the HON and SIM conditions lie on different sides of the standard FDS = 0.50 cut-off regardless of the Introvertive Anhedonia score. To confirm that O-LIFE scores did not significantly affect FDS, we ran a model predicting the accuracy of each participant's classification (as HON and SIM) by using their score on each of the four O-LIFE subscales, and a random effect of participant as predictors. A model that simply predicted classification accuracy based on the random effect of participant only performed equally well, Chi^2 (4) = 6.69, p = 0.15. Similarly, none of the O-LIFE subscales were significant predictors of classification accuracy in this model (all $z < 1.90$, all $p > 0.05$).

There is no evidence that education influenced FDS scores. As noted above, two individuals endorsed "Other" for the education field. They were treated as missing in these analyses, so that a direct comparison between a model with and without education is not possible. However, a model including condition, Introvertive Anhedonia and education showed no significant effect of education, *estimate* = -0.003, *SE* = 0.01, t = -0.27.

Table 4 provides some additional information on the classification accuracy of the IOP-29 FDS by inspecting a conservative (IOP-29 FDS ≥ .65), standard (IOP-29 FDS ≥ .50), liberal (IOP-29 FDS ≥ .30) and very liberal (IOP-29 FDS ≥ .15) cut-off score (Giromini et al., 2018). Specificity ranged from 71.5 (for IOP-29 FDS ≥ .15) to 99.3 (for IOP-29 FDS ≥ .65); sensitivity ranged from 82.8 (for IOP-29 FDS ≥ .65) to 100 (for IOP-29 FDS ≥ .15).

Effectiveness of the IOP-29 RRS

We next focused on the effectiveness of the IOP-29 RRS in the detection of random responding. Again, we found a strong effect of condition in the first model (HON v. RND *estimate* = 19.05, *SE* = 0.91, *t* = 20.98; HON v. SIM *estimate* = 3.17, *SE* = 0.91, *t* = 3.50). The highest IOP-29 RRS scores were produced by RND (*M* = 68.4; *SD* = 9.2), followed by SIM (*M* = 52.5; *SD* = 8.9), and lastly by HON (*M* = 49.3; *SD* = 6.5). The size of the difference between RND and the other two conditions was *d* = 2.40 for HON and *d* = 1.76 for SIM; the difference between HON and SIM consisted of a much smaller *d* of .41. A graphical representation of the IOP-29 RRS scores obtained across the three conditions is reported in Figure 3.

By using the cut-off score of T ≥ 61 as recommended by Giromini, Viglione, et al. (2019b), we inspected the classification accuracy of the IOP-29 RRS across the three conditions. The results of these analyses, reported in Table 5, produce specificity values ranging from 84.1% (condition SIM) to 96.7% (condition HON), and sensitivity of 83.4% (condition RND). AUC values were .95 (*SE* = .02), .90 (*SE* = .02), and .92 (*SE* = .02) respectively, when comparing condition RND versus condition HON, versus condition SIM, and versus conditions HON and SIM combined (Figure 4).

The correlation of IOP-29 RRS to IOP-29 FDS was .34 in condition HON, -.27 in condition SIM, and .22 in condition RND (all of these correlations were statistically significant

at $p < .01$). Unexpectedly, thus, the scores of the IOP-29 RRS were not independent from those of the IOP-29 FDS. Interestingly, in conditions HON and RND the two indexes correlated *positively*, whereas in condition SIM they correlated *negatively*.

Discussion

The current study was designed to test the validity of the Inventory of Problems-29 (IOP-29; Viglione & Giromini, 2020; Viglione et al., 2017) with a British population. Additionally, it also provided a first independent validation of the IOP-29 Random Responding Scale (RRS; Giromini, Viglione, et al., 2019b). Examination of 453 IOP-29 protocols from 151 adult volunteers revealed that: 1) the False Disorder Probability Score (FDS) of the IOP-29 discriminated feigned schizophrenia from honest responding with excellent accuracy in this UK sample; 2) the IOP-29 RRS accurately differentiated random responding from both feigned schizophrenia and honest responding.

When comparing the IOP-29 FDS values in the honest versus feigning schizophrenia conditions our Cohen's *d* was 4.20 and AUC was .99. We thus may conclude that the performance of the IOP-29 FDS with our British sample was at least as good as it was in Giromini, Viglione, et al.'s (2019a) study conducted with Italian healthy volunteers (where Cohen's *d* was 3.16 and AUC was .96 in the schizophrenia-related condition). However, simulation/analogue studies yield larger effect sizes when comparing experimental feigners against nonclinical controls rather than against bona fide patients (Rogers & Bender, 2018; van Impelen et al., 2014). Indeed, Viglione et al.'s (2017) studies conducted in the US included a subsample of 45 bona fide psychosis patients and 45 healthy schizophrenia feigners and found a lower Cohen's *d* value of 1.95, and a smaller AUC of .92. Along the same lines, Giromini et al. (2018) found a Cohen's *d* of 1.80 and a AUC of .89 in a large Italian sample ($N = 452$) when

comparing a subsample of 89 bona fide patients affected by psychosis to 125 schizophrenia feigners. Thus, future replications with clinical control samples are sorely needed. Nevertheless, it is noteworthy that the average IOP-29 FDS values found in the SIM condition of our study (*M* = .82; *SD* = .18) closely resemble those observed in other International experimental simulator samples. For instance, in Giromini, Viglione, et al.'s (2019a) Italian study (N = 400) the average IOP-29 FDS value for simulators was .82 (*SD* = .20); in Giromini, Barbosa, et al.'s (2019) Portuguese study (N = 100) simulators scored on average .82 (*SD* = .20); in Ilgunaite et al.'s (2020) Lithuanian study simulators produced an average IOP-29 FDS of .77 (*SD* = .18).

Our analyses also indicated that the presence of schizotypal traits did not notably influence the IOP-29's accuracy. Indeed, although participants with high Introvertive Anhedonia tended to generate slightly inflated FDS scores in both HON and SIM conditions, using the standard cut-off score of IOP-29 FDS ≥ .50 ensured the same classification accuracy regardless of what their score on Introvertive Anhedonia was. This finding and the consistency noted of the FDS noted in the above paragraph supports using the standard, IOP-29 FDS cut-score of .50 whenever applicable. It is important to acknowledge, however, that assessors using more liberal FDS cut-off scores such as IOP-29 FDS ≥ .30 might run a slightly increased risk of falsely classifying individuals with high Introvertive Anhedonia as feigners. As those lower cut-off scores are typically used for screening purposes, we recommend that when setting these cut-offs, subsequent, follow-up testing would include a measure of schizotypy-related traits, if possible.

Another encouraging finding is that the IOP-29 RRS discriminated IOP-29s in the RND condition from those in the HON and SIM conditions (AUC ≥ .90). Giromini, Viglione et al. (2019b) developed the IOP-29 RRS with the purpose of identifying content unrelated distortions, associated for example with impaired cognitive or reading abilities or uncooperative responding.

The ultimate goal was to identify possible IOP-29 FDS false positive outcomes originated by poor comprehension of the items, distraction, resistance to the testing, and random responding rather than active feigning. Because we instructed our participants to respond completely at random, however, additional research is needed to address whether the RRS also performs adequately with partial random responding or reduced attention/concentration.

One last consideration deserves mentioning, with regard to the relationship between the FDS and RRS. When Giromini, Viglione et al. (2019b) developed the RRS, their goal was to pull apart content-related (e.g., voluntary exaggeration or malingering) from content-unrelated (e.g., random responding, unable to understand the questions, resistance) distortion sources. As such, we did not expect the RRS values to be directly associated with those of the FDS. Instead, in our study the RRS and FDS correlated positively in the HON condition and negatively in the SIM condition. A possible explanation for this finding is that the RRS might moderate the validity of the FDS. Indeed, in the HON condition, where the FDS was supposed to be low, the presence of some randomness in the responses tended to artificially inflate the FDS values. Conversely, in the SIM condition, where the FDS was supposed to be high, the presence of some randomness in the responses tended to artificially diminish the FDS values. That is, the more the participant understood and put an effort into completing the IOP-29 with the needed attention per administration instruction, the more accurate their FDS. If these speculations were true, the RRS might prove particularly helpful in those cases in which the FDS is close to .50, i.e., the 'too-close-to-classify' cases (Rogers & Bender, 2018). In those situations, if the FDS is moderately high, there are two likely possibilities: The person is exaggerating/feigning, or they misunderstood/lacked attention to the test items. To determine whether a marginally high FDS scorer misunderstood/lacked attention to the test items, we should examine the RRS. If the

corresponding RRS is high, it is possible that the person did not understand the test or did not cooperate with it, whereas if RRS is low, it is more likely that the person simply exaggerated their symptoms. When both FDS and RRS are low, this increases the likelihood that a test taker is responding in a valid manner, as scoring low on both of these scales require meticulous attention to the test content. Additional research – particularly with clinical samples – is needed, however, so these recommendations should be considered to be largely speculative, at this point.

Our study has some important practical limitations. First, the ecological conditions and motivations that prevail in real-life malingering or random responding cannot be imposed in an online study. Thus, the external validity of our study is limited, especially when one visualizes that many participants likely completed their testing at their leisure in the comfort of their home. Secondly, our quasi-experimental design was vulnerable to confounds as participant compliance with instruction at each administration was unknown. Respondent noncompliance is an ongoing threat to validity in malingering studies (Rai, An, Charles, Ali, & Erdodi, 2019; Walls, Wallace, Brothers, & Berry, 2017), as these studies operate on the assumption that test outcomes are primarily linked to the absence or presence of motivation, as opposed to commitment to instructions (see An, Charles, Ali, Enache, Dhuga, & Erdodi, 2019). Although our study used a manipulation check to exclude respondents who were not attentive to instructions, future studies should employ a more rigorous manipulation check. On the other hand, such internal validity problems would reduce rather than inflate effect sizes. Third, our use of one vignette in which the character is motivated to feign psychosis for external gain may have limited our ability to evaluate the impact of vignette-specific characteristics. Future studies should examine different scenarios within vignettes (Giromini, Viglione et al., 2019a), and additionally use vignettes in which characters are motivated to feign in order to mitigate criminal culpability and punishment.

Fourth, while random responding in the study was instructed as "respond randomly - that is, with no apparent pattern," we did not enquire about the strategies individuals used to respond randomly and thus, we could not determine whether specific strategies impacted the RRS. Future studies should identify and investigate specific styles of random responding (e.g., reading difficulties, distractibility, and resistance) as they relate to honest and feigned responding. Fifth, our administration did not allow test takers to skip responses, which may be common in both feigning and random responding. Sixth, although we excluded individuals with certain psychiatric problems from participating (<6 months: neurological, TBI, substance use; lifetime history: psychosis-related disorders) it is unclear how or whether our inclusion of individuals with psychiatric diagnoses such as depression, anxiety, and learning disorders may have affected the findings. Despite this, the inclusion of individuals with less severe pathological issues may have served to represent an accurate real-world sample, in that feigners may be suffering psychological problems which are not related to their presenting complaint. Finally, generalizability of our findings is limited due to absence of a clinical comparison group. Future clinical comparison samples might differentiate FDS and RRS values in non-pathological controls versus individuals who already suffer one or more specific psychiatric complaints, or possibly test the ecological validity of the IOP-29 with a 'real-life' forensic sample as was recently done by Roma et al. (2020).

Despite the limitations, this study adds to the growing literature on the applicability and utility of the IOP-29. This is the first use of the IOP-29 with a British sample, further strengthening its cross-cultural generalizability. Our study also supports the notion that even if schizotypic test takers might score slightly higher than controls on the FDS scale, they should not do so at a level that might notably interfere with the IOP-29's classification accuracy. Finally,

this is the first IOP-29 study to independently cross-validate the psychometric properties of the IOP-29 RRS, providing preliminary but encouraging evidence in support of its possible applicability to real-life contexts.

References

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

An, K.Y., Charles, J., Ali, S., Enache, A., Dhuga, J., & Erdodi, L.A. (2019). Reexamining performance validity cutoffs within the Complex Ideational Material and the Boston Naming Test-Short Form using an experimental malingering paradigm. *Journal of Clinical and Experimental Neuropsychology, 41,* 1, 15-25. https://doi.org/10.1080/13803395.2018.1483488

Ardolf, B.R., Denney, R.L. & Houston, C.M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist, 21,* 6, 899–916. https://doi.org/10.1080/13825580600966391

Avramopoulos, D., Stefanis, N.C., Hantoumi, I., Smyrnis N., Evdokimidis I., & Stefanis C.N. (2002). Higher scores of self reported schizotypy in healthy young males carrying the COMT high activity allele. *Molecular Psychiatry 7,* 706–711. https://doi.org/10.1038/sj.mp.4001070

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, *67,* 1, 1–48. https://doi.org/10.18637/jss.v067.i01

Binder, L.M. (1993). Assessment of malingering after mild head trauma with the Portland Digit Recognition Test. *Journal of Clinical and Experimental Neuropsychology, 15,* 170-182. https://doi.org/10.1080/01688639308402555

Boone, K. B. (2013). *Clinical Practice of Forensic Neuropsychology.* New York, NY: Guilford.

Burchett, D., Dragon, W.R., Smith Holbert, A.M., Tarescavage, A.M., Mattson, C.A., Handel, R.W., & Ben-Porath, Y.S. (2016). "False Feigners": Examining the impact of non-content-based invalid responding on the Minnesota Multiphasic Personality Inventory-2 Restructured Form content-based invalid responding indicators. *Psychological Assessment, 28*, 5, 458-470. https://doi.org/10.1037/pas0000205

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

Butcher, J.N., Graham, J.R., Ben-Porath, Y.S., Tellegen, A.M., & Dahlstrom, W.G. (2001). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring* (rev. ed.). Minneapolis, MN: University of Minneapolis Press.

Chafez, M.D. (2008). Malingering on the social security disability consultative exam: Predictors and base rates. *The Clinical Neuropsychologist, 22,* 3, 529–546. https://doi.org/10.1080/13854040701346104

Claridge, G. (Ed.). (1997). Schizotypy: Implications for illness and health. New York, NY: Oxford University Press. https://doi.org/10.1093/med:psych/9780198523536.001.0001

Claridge, G., & Beech, T. (1995). Fully and quasi-dimensional constructions of schizotypy. In A. Raine, T. Lencz, & S. A. Mednick (Eds.), *Schizotypal personality* (pp. 192-216). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511759031.010

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Dunlap, W.P., Cortina, J.M., Vaslow, J.B., & Burke, M.J. (1996). Meta-Analysis of experiments with matched groups of repeated measures designs. *Psychological Methods, 1,* 2*,* 170-177. https://doi.org/10.1037/1082-989X.1.2.170

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed.). Thousand Oaks, CA: Sage. http://tinyurl.com/carbook.

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

Giromini, L., Carfora Lettieri, S., Zizolfi, S., Zizolfi, D., Viglione, D.J., Brusadelli, E.,... 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

Giromni, L., Pignolo, C., Zennaro, A., & Viglione, D. (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. https://doi.org/10.1007/s12207-018-9314-1

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

Giromni, L., Viglione, D.J., Pignolo, C., & Zennaro, A. (2019b). An Inventory of Problems - 29 (IOP-29) study on random responding using experimental feigners, honest controls, and computer-generated data. *Journal of Personality Assessment, 68*, 1, 1-12. https://doi.org/10.1080/00223891.2019.1639188

Grant, P., Kuepper, Y., Mueller, E.A., Wielpuetz, C., Mason, O., & Hennig, J. (2013). Dopaminergic foundations of schizotypy as measured by the German version of the Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE)-a suitable endophenotype of schizophrenia. *Frontiers in Human Neuroscience*, *7*, 1. https://doi.org/10.3389/fnhum.2013.00001

Greiffenstein, M.F., & Baker, W.J. (2008). Validity testing in dually diagnosed post-traumatic stress disorder and mild closed head injury. *The Clinical Neuropsychologist, 22,* 565–582. https://doi.org/10.1080/13854040701377810

Halligan, P.W., Bass, C., & Oakley, D.A. (Eds.) (2003). *Malingering and illness deception*. Oxford: Oxford University Press.

Ilgunaite, G., Giromini, L., Bosi, J., Viglione, D. J., & Zennaro, A. (2020). A clinical comparison simulation study using the Inventory of Problems-29 (IOP-29) with the Center for Epidemiologic Studies Depression Scale (CES-D) in Lithuania. *Applied Neuropsychology: Adult, [Epub ahead of print]*, https://doi.org/10.1080/23279095.2020.1725518

Kirchner, S.K., Roeh, A., Nolden, J., & Hasan, A. (2018). Diagnosis and treatment of schizotypal personality disorder: Evidence from a systematic review. *NPJ Schizophrenia*, *4*, 1, 20. https://doi.org/10.1038/s41537-018-0062-8

Lachman, H.M., Papolos, D.F., Saito, T., Yu, Y.M., Szumlanski, C.L., & Weinshilboum, R.M. (1996). Human catechol-O-methyltransferase pharmacogenetics: Description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics 6,* 243–250. https://doi.org/10.1097/00008571-199606000-00007

Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist, 17*, 54–68. https://doi.org/10.1076/clin.17.3.410.18089

Lenzenweger, M.F. (2015). Thinking clearly about schizotypy: Hewing to the schizophrenia liability core, considering interesting tangents, and avoiding conceptual quicksand. *Schizophrenia Bulletin, 41*, 2, 483-S491. https://doi.org/10.1093/schbul/sbu184

Martin, P.K., Schroeder, R.W., & Odland, A.P. (2015). Neuropsychologists' validity testing beliefs and practices: A survey on North American professionals. *The Clinical Neuropsychologist, 29*, 741–776. https://doi.org/10.1080/13854046.2015.1087597

Mason, O., & Claridge, G. (2006). The Oxford-Liverpool Inventory of Feelings and Experiences: Further description and extended norms. *Schizophrenia Research, 82,* 203-211. https://doi.org/10.1016/j.schres.2005.12.845

McCarter, R.J., Walton, N.H., Brooks, D.N., & Powell, G.E. (2009). Effort testing in contemporary UK neuropsychological practice. *The Clinical Neuropsychologist, 23,* 1050-1066. https://doi.org/10.1080/13854040802665790

McMillan, T.M., Anderson, S., Baker, G., Berger, M., Powell, G.E., & Knight, R. (2009) *Assessment of effort in clinical testing of cognitive functioning for adults.* Leicester, UK: British Psychological Society.

Meehl, P.E. (1962). Schizotaxia, schizotypy, schizophrenia. *American Psychologist, 17,* 827–838. https://doi.org/10.1037/h0041029

Meehl, P.E. (1994). Toward an integrated theory of schizotaxia, schizotypy, and schizophrenia. *Journal of Personality Disorders, 4,* 1–99. https://doi.org/10.1521/pedi.1990.4.1.1

Merten, T., Dandachi-FitzGerald, B., Hall, V., Schmand, B.A., Santamaría, P., & González-Ordif, H. (2013). Symptom validity assessment in European countries: Development and state of the art. *Clínica y Salud, 24,* 129-138. https://doi.org/10.1016/S1130-5274(13)70014-8

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

Miller, L.J., Ryan, J.J., Carruthers, C.A., & Cluff, R.B. (2004). Brief screening indexes for malingering: A confirmation of Vocabulary minus Digit Span from the WAIS-III and the Rarely Missed Index from the WMS-III. *The Clinical Neuropsychologist, 18,* 2, 327–333. https://doi.org/10.1080/13854040490501592

Morris, S.B., & DeShon, R.P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*, 105-125. https://doi.org/10.1037/1082-989X.7.1.105

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

Nichols, D., Greene, R., & Schmolck, P. (1989). Criteria for assessing inconsistent patterns of item endorsement on the MMPI: Rationale, development, and empirical trials. *Journal of Clinical Psychology, 45*, 2, 239-250. https://doi.org/10.1002/1097-4679(198903)45:2<239::AID-JCLP2270450210>3.0.CO;2-1

Pierre, J.M., Shnayder, I., Wirshing, D.A., & Wirshing, W.C. (2004). Intranasal quetiapine abuse. *American Journal of Psychiatry*, *161,* 1718 (letter to the editor).

Rado, S. (1960). Theory and therapy: The theory of schizotypal organization and its application to the treatment of decompensated schizotypal behavior. In: Scher, S.C., Davis, H.R., eds. *The Outpatient Treatment of Schizophrenia*. New York, NY: Grune & Stratton, 87–101.

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

Rogers, R. (Ed.). (2008). Clinical assessment of malingering and deception (3rd ed.). New York, NY: Guilford Press.

Rogers, R., & Bender, S.D. (2013). Evaluation of malingering and related response styles. In R. K. Otto & I. B. Weiner (Eds.), *Handbook of psychology: Forensic psychology* (p. 517–540). Hoboken, NJ: John Wiley & Sons Inc.

Rogers, R., & Bender, S.D. (Eds.). (2018). Clinical assessment of malingering and deception (4th ed.). New York, NY: The Guilford Press.

Rogers R., & Vitacco, M.J. (2002). Forensic assessment of malingering and related response styles. In B. Van Dorsten (Eds.), *Forensic Psychology.* Boston, MA: Springer.

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, 13*, 57-65. https://doi.org/10.1007/s12207-019-09368-4

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. https://doi.org/10.1037/t04573-000

Slick, D.J., Sherman, E.M.S., Grant, L., & Iverson, G.L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist, 13, 4*, 545–561. https://doi.org/10.1076/1385-4046(199911)13:04;1-Y;FT545

Smyrnis, N., Avramopoulos, D., Evdokimidis, I., Stefanis, C.N., Tsekou, H., & Stefanis, N.C. (2007). Effect of schizotypy on cognitive performance and its tuning by COMT val(158) met genotype variations in a large population of young men. *Biological Psychiatry 61,* 845–853. https://doi.org/10.1016/j.biopsych.2006.07.019

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

UK Department of Health & Social Care. (2020). NHS Counter Fraud Authority. https://reportfraud.cfa.nhs.uk Accessed on 2 April 2020.

UK Department for Work & Pensions. (2019). Fraud and Error in the Benefit System: Financial year 2018 to 2019 estimates. National Statistics. Published on 9 May 2019. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/801594/fraud-and-error-stats-release-2018-2019-estimates.pdf Accessed on 2 April 2020.

UK Disability Unit. (2020). Disability Rights: When a mental health condition becomes a disability. Published on 2 April 2020. https://www.gov.uk/when-mental-health-condition-becomes-disability Accessed on 20 April 2020.

van Impelen, A., Merckelbach, H., Jelicic, M., & Merten, T. (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. *The Clinical Neuropsychologist, 28,* 8, 1336-1365. https://doi.org/10.1080/13854046.2014.984763

Vandenbergh, D.J., Persico, A.M., Hawkins, A.L., Griffin, C.A., Li, X., Jabs, E.W., & Uhl, G.R. (1992). Human dopamine transporter gene (DAT1) maps to chromosome-5p15.3 and displays a VNTR. *Genomics 14,* 1104–1106. https://doi.org/10.1016/S0888-7543(05)80138-7

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*, 534-44. https://doi.org/10.1080/00223891.2016.1233882

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

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

World Health Organization. (2018). *International classification of diseases for mortality and morbidity statistics* (11th rev.). Retrieved from https://icd.who.int/browse11/l-m/en

World Health Organization. (2008). The Global Burden of Disease - 2004 Update. Geneva, Switzerland: World Health Organization.

Young, G. (2014). *Malingering, feigning, and response bias in psychiatric/psychological injury: Implications for practice and court.* Dordrecht: Springer Science + Business Media.

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

Table 1. Descriptive Statistics for IOP-29 and O-LIFE Scores Included in the Analyses

	M	SD
IOP-29 FDS		
HON	0.14	0.14
SIM	0.82	0.18
IOP-29 RDS		
HON	49.3	6.5
SIM	52.5	8.9
RND	68.4	9.2
O-LIFE		
Unusual Experiences	2.75	2.48
Cognitive Disorganization	5.66	3.01
Introvertive Anhedonia	2.70	2.24
Impulsive Nonconformity	3.44	2.15

Table 2. Linear Mixed Effects Models Testing the Effectiveness of the IOP-29 FDS

Model	NPAR	AIC	BIC	Loglikelihood	Deviance
Condition only	4	-243.99	-229.15	125.99	-251.99
Condition + all O-LIFE subscales	8	-265.40	-235.72	140.70	-281.40
Condition + Introvertive Anhedonia	5	-268.68	-250.13	139.34	-278.68
Condition * Introvertive Anhedonia	6	-269.26	-247.00	140.63	-281.26

Table 3. Predictors Included in the Linear Mixed Effects Model Testing the Effect of O-LIFE subscales on the IOP-29 FDS

	Estimate	*SE*	*t*
(Intercept)	0.11	0.02	4.80
Condition SIM	0.67	0.02	38.47
Cognitive Disorganization	0.00	0.00	-0.59
Introvertive Anhedonia	0.02	0.00	5.25
Impulsive Nonconformity	-0.01	0.01	-1.23
Unusual Experiences	0.00	0.00	0.52

Table 4. Sensitivity and specificity of the IOP-29 FDS (HON versus SIM).

	HON		SIM	
	n	%	n	%
Conservative Cut Score				
IOP-29 FDS ≥ .65	1	0.7	125	82.8 [b]
IOP-29 FDS < .65	150	99.3 [a]	26	17.2
Standard Cut Score				
IOP-29 FDS ≥ .50	5	3.3	139	92.1 [b]
IOP-29 FDS < .50	146	96.7 [a]	12	7.9
Liberal Cut Score				
IOP-29 FDS ≥ .30	17	11.3	150	99.3 [b]
IOP-29 FDS < .30	134	88.7 [a]	1	0.7
Very Liberal Cut Score				
IOP-29 FDS ≥ .15	43	28.5	151	100.0 [b]
IOP-29 FDS < .15	108	71.5 [a]	0	0.0

[a] Specificity; [b] Sensitivity.

Table 5. Sensitivity and specificity of the IOP-29 RRS.

	HON		SIM		HON & SIM		RND	
	n	%	n	%	n	%	n	%
IOP-29 RRS ≥ 61	5	3.3	24	15.9	29	9.6	126	83.4 [b]
IOP-29 RRS < 61	146	96.7 [a]	127	84.1 [a]	259	90.4 [a]	25	16.6

[a] Specificity; [b] Sensitivity.

Figure 1. Graphical Representation of IOP-29 FDS Scores in HON and SIM Conditions.

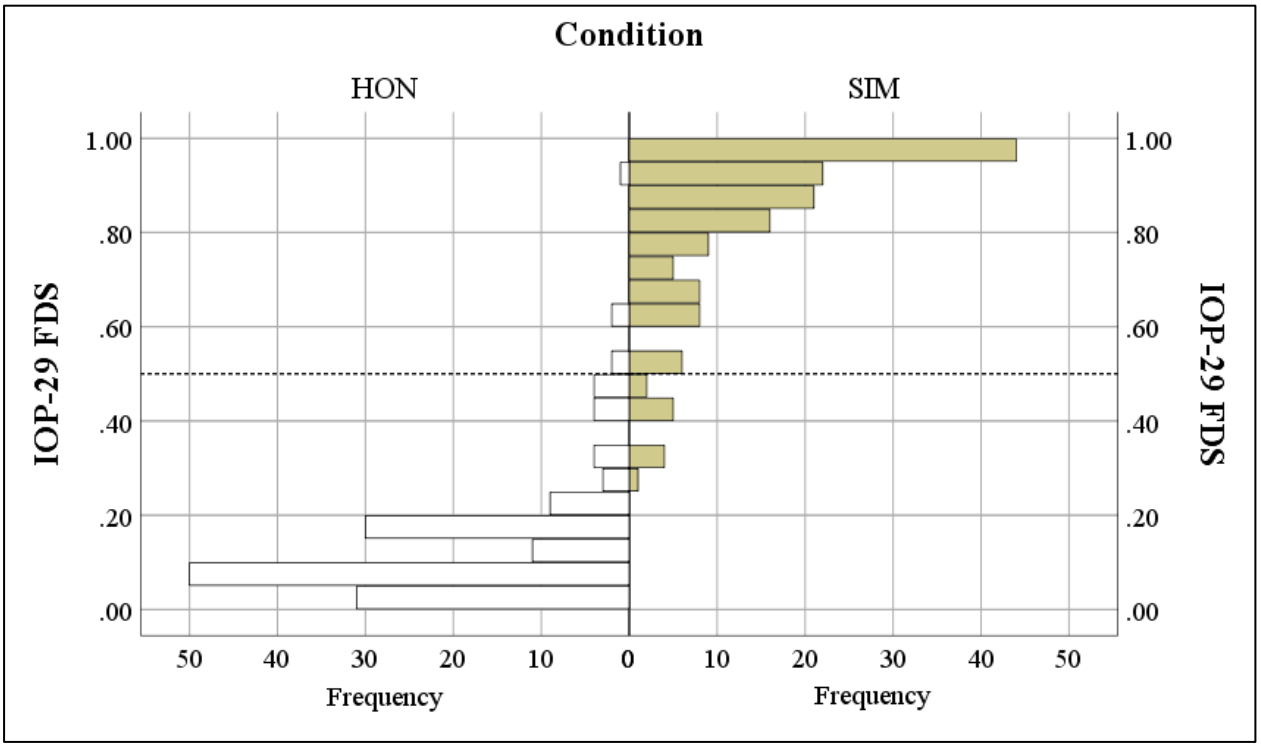


Figure 2. Interaction between IOP-29 FDS and Introvertive Anhedonia, derived from the interaction model (see Table 2)

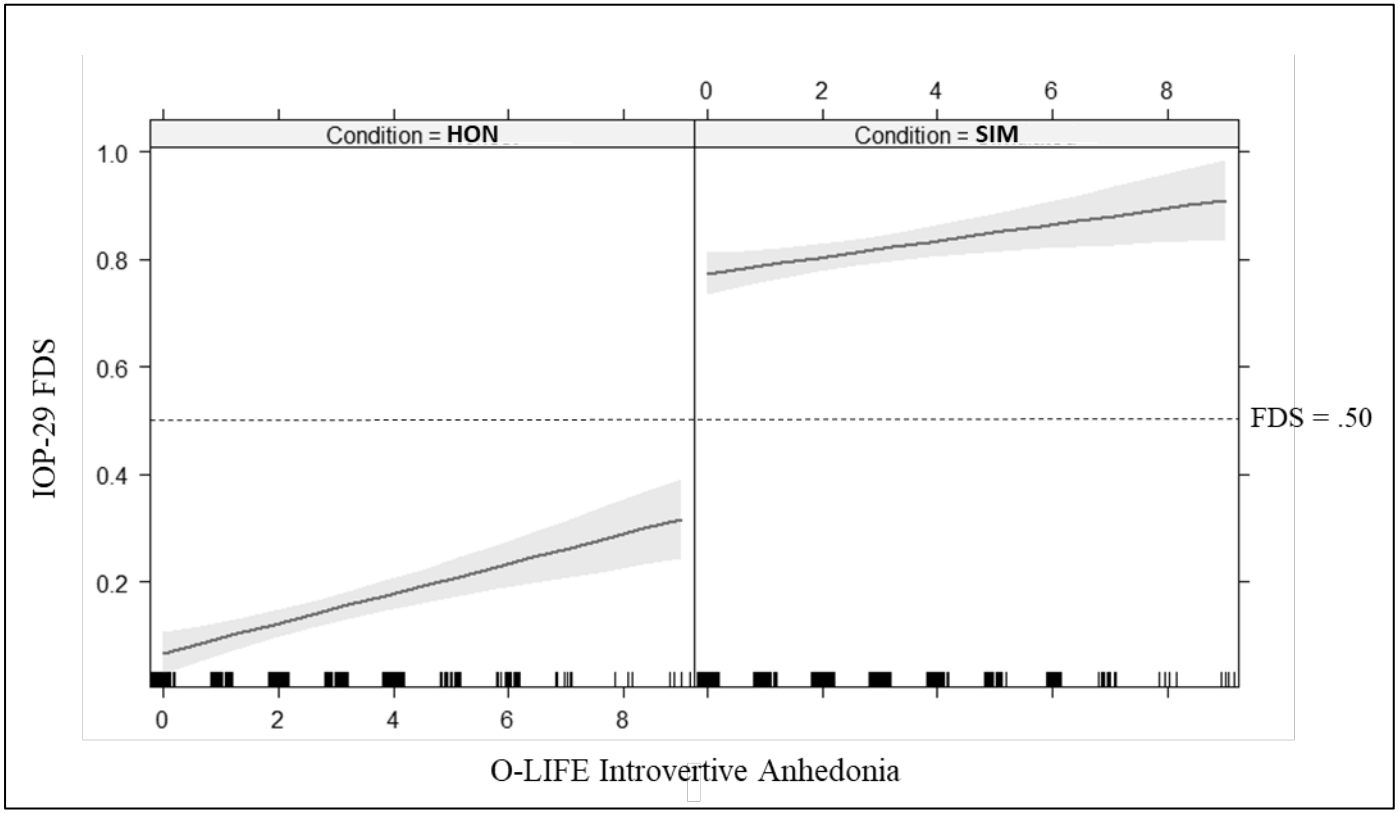


Figure 3. IOP-29 RRSD Scores by Condition

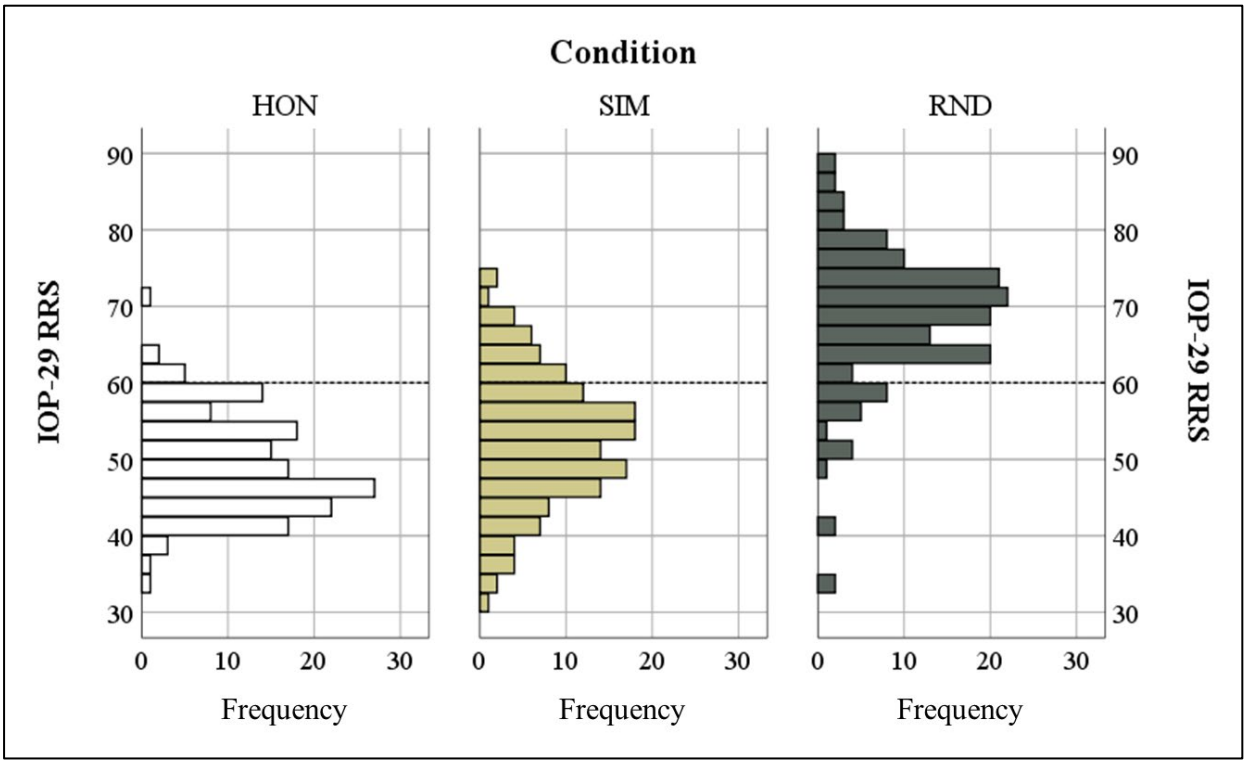
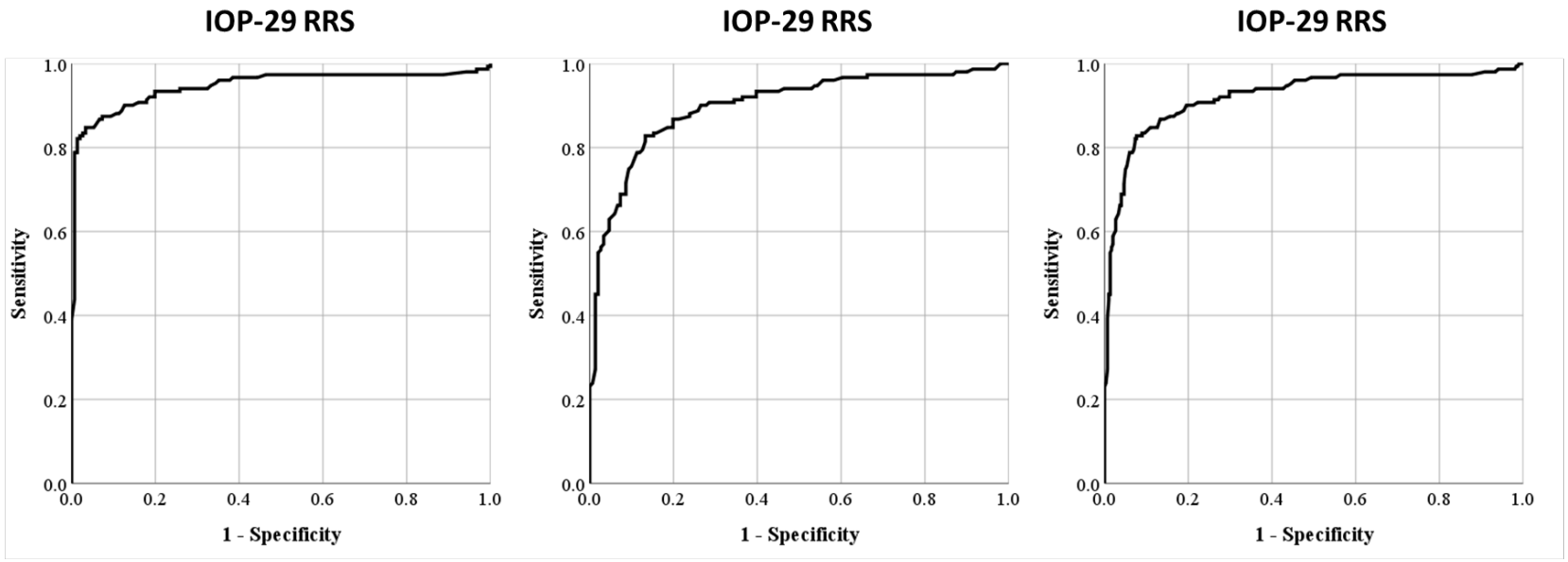


Figure 4. Receiver Operator Characteristic curve contrasting RND condition against HON and SIM conditions.



Appendix A

Now you will be asked to fake schizophrenia. To help you assume the role of a person with schizophrenia, please read the following short story and link to NHS website. You will be asked to respond to questions as if you are the main character.

Please read the story and link carefully. You cannot return to this page.

Recently, you've fallen on hard times. Your long-time partner who you hoped to marry dumped you last week. You have not been performing well at work and you're nervous about your upcoming performance review. Your boss has noticed your poor performance and has hinted at the possibility of letting go of some employees. You have student loans to pay off and extra rent to pay now that your partner has moved out. On top of all of this, you hate your co-workers and your job. The situation is causing you immense stress.

Two years ago, your partner encouraged you to see a psychologist because you seemed unhappy all the time. You went for a couple of visits. After sharing your childhood history and current problems, the psychologist diagnosed you with depression. The psychologist recommended Cognitive Behavioural Therapy and talking to your GP about going on medication. This made you feel small and like she was not listening to you. You refused to see the psychologist any longer because you disagreed with the diagnosis. At the time, you felt your problems were attributed to your partner's lack of support and being unemployed.

To calm your nerves, you use marijuana recreationally. After work, you went to a friend's house and smoked so much weed that for ten minutes, you thought you heard the devil inside your heart whispering to you. After you sober up, you remember all your problems that you are currently facing and you feel powerless. Then, out of nowhere, you think back to the time you saw the psychologist. You remember very clearly that when you shared your family history, the psychologist mentioned that you may be at risk of developing a psychotic disorder. At the time, you disagreed, as you know you are not "crazy" like your Aunt Suzie. You affirm to yourself that this strange incident was due to your marijuana use.

You go home and search the internet for symptoms of psychotic disorders just to make sure you aren't crazy. While you're doing this, you realise that one psychotic disorder – schizophrenia, is a condition that deems you eligible for disability benefits, such as Employment and Support Allowance, and possible protection from losing your job. You decide a schizophrenia diagnosis may be the answer to all your problems. You learn about the symptoms of schizophrenia on the NHS website. Take a few minutes to read the page of schizophrenia symptoms (remember, **the top 3 fakers win £20 each!**).