Detecting malingering mental illness in forensics: Known-Group Comparison and Simulation Design with MMPI-2, SIMS and NIM

Barbara De Marchi[1] and Giulia Balboni[2]

[1] Centro Ferrarese di Neuropsichiatria, Neuropsicologia e Riabilitazione, Ferrara, Italy
[2] Department of Philosophy, Social and Human Sciences and Education, University of Perugia, Perugia, Italy

Corresponding author
Barbara De Marchi,
bademarchi@yahoo.com

ABSTRACT

Background. Criminal defendants may often exaggerate psychiatric symptoms either to appear non-accountable for their actions or to mitigate their imprisonment. Several psychometric tests have been proposed to detect malingering. These instruments are often validated by Simulation Design (SD) protocols, where normal participants are explicitly requested to either simulate a mental disorder or respond honestly. However, the real scenarios (clinical or forensic) are often very challenging because of the presence of genuine patients, so that tests accuracy frequently differs from that one obtained in well-controlled experimental settings. Here we assessed the effectiveness in criminal defendants of three well-known malingering-detecting tests (MMPI-2, SIMS and NIM) by using both Known-Group Comparison (KGC) and Simulation Design (SD) protocols.

Methods. The study involved 151 male inmates. Participants to the KGC protocol were all characterized by a positive psychiatric history. They were considered as genuine patients (KGC_Controls) if they had some psychiatric disorders already before imprisonment and scored above the cutoff of SCL-90-R, a commonly used test for mental illness, and as suspected malingerers (KGC_SM) if they were diagnosed as psychiatric patients only after imprisonment and scored below the SCL-90-R cutoff. Participants to SD protocol had no history of psychiatric disease and scored below the SCL-90-R cutoff. They were randomly assigned to either group: Controls (requested to answer honestly, SD_Controls) and simulated malingerers (requested to feign a psychiatric disease, SD_SM). All participants were then submitted to MMPI-2, NIM and SIMS.

Results. Results showed that while MMPI-2, SIMS and NIM were all effective in discriminating malingerers in the SD, SIMS only significantly discriminated between KGC_Controls and KGC_SM in the Known-Group Comparison. Receiver Operating Characteristic (ROC) curves analysis confirmed the better sensitivity of SIMS with respect to the other tests but raised some issues on SIMS specificity.

Discussion. Results support the sensitivity of SIMS for the detection of malingering in forensic populations. However, some specificity issues emerged suggesting that further research and a good forensic practice should keep into account multiple measures of malingering, including psychometric data, clinical and social history and current clinical situation. These methodological constraints must be kept in mind during detection of malingering in criminal defendants reporting psychiatric symptoms.

INTRODUCTION

The detection of malingering is an important topic in psychology (for a comprehensive review, see *Seron, 2014*) raising practical and ethical issues (*Bass & Halligan, 2014*). A significant percentage of individuals undergoing psychological evaluation may feign psychopathological symptoms especially when the context is perceived as challenging, such as in the forensic (*Rogers, Sewell & Goldstein, 1994*) and clinical (*Noeker & Petermann, 2011*) settings. The forensic framework and in particular criminal defendants represent a challenging situation characterized by a high frequency of malingering behaviors (*Merten & Rogers, 2017*) although with different base rate (*Mittenberg et al., 2002*; *Young, 2014*). Criminal defendants simulate psychopathology to avoid or to delay punishment or to obtain more favorable conditions (*Rogers, 2008*). Because of these reasons, over the past two decades, symptom validity research in this field has significantly intensified (*Otto & Heilbrun, 2002*; *Douglas, Otto & Borum, 2003*; *McLaughlin & Kan, 2014*).

Various studies have demonstrated that several psychopathologies can be simulated by malingering individuals (*Lees-Haley & Dunn, 1994*; *Bowen & Bryant, 2006*; *Baity et al., 2007*). They include major depression, generalized anxiety disorder and post-traumatic stress disorder (PTSD). Several psychometric instruments have been proposed to detect malingering, from multidimensional personality inventories (*Berry et al., 2001*; *Edens, Cruise & Buffington-Vollum, 2001*) to specific tests for malingering (for a review, see *Sellbom & Bagby, 2008*).

In the present study we selected three tests used for malingering detection: The Minnesota Multiphasic Personality Inventory (MMPI)-2 (*Butcher et al., 1989*), the Structured Inventory of Malingered Symptomatology (SIMS) (*Smith & Burger, 1997*; *Widows & Smith, 2005*) and the Negative Impression Management (NIM) scale of the Personality Assessment Inventory (PAI) (*Morey, 1991*) based on the existence of Italian versions for MMPI-2 (*Pancheri & Sirigatti, 1995*) and SIMS (*La Marca et al., 2012*) and of an Italian translation of the NIM specifically done for the present study (see 'Methods' Section).

These three instruments have been selected because of two main reasons. The first one was their diffusion. Indeed, according to an analysis of the literature on malingering detection in forensics, MMPI-2, SIMS and NIM were among the most used instruments at the moment of data collection. The second reason was the existence of an Italian version of the instrument.

MMPI-2 is a personality questionnaire that enables the detection of several psychopathological dimensions. Additionally, by means of specific validity scales, MMPI-2 can be used to detect inconsistent responding, over-reporting and under-reporting psychological symptoms, even in healthy participants. These reasons have expanded the use of the MMPI-2 outside the clinical environment, from the selection of workers and employees to the detection of fraudulent behaviors in people asking financial compensations and to prisoners and criminal defendants (*Pope, Butcher & Seelen, 2006*). In particular, as

far as criminal defendants are concerned, the family of F (Infrequency) scales provides most information about simulation behaviors (*Meehl & Hathway, 1946*) being made up of items very rarely endorsed by the MMPI normative group (for a discussion, see *Rogers, 1997a*; *Rogers, 1997b*). This strategy has been extended to the MMPI-2 with the creation of the Fb (Infrequency Back), the Fp (Infrequency Psychopathology; *Arbisi & Ben-Porath, 1995*), and the Fc (Criminal Offender Infrequency; *Mergargee, 2004*; *Mergargee, 2006*) scales. More recently, other scales have been developed to detect malingering behaviors in non-criminal settings (Fake Bad Scale, FBS; Response Bias Scale, RBS; for a recent review, see *Merten & Rogers, 2017*).

MMPI-2 is, however, very long to administer. For this reason, clinicians frequently prefer faster and easier screening measures, particularly useful when the assessment has to be conducted in large participants cohorts in particular for preliminary screenings (*Edens, Poytress & Watkins-Clay, 2007*). The NIM scale of the PAI (*Morey, 1991*), has been considered to provide an adequate level of malingering detection (*Morey & Lanier, 1998*). Indeed, NIM scale has been included in the PAI as a measure of exaggeration or malingering and consists of items rarely endorsed in the case of significant diseases and items associated with a very negative description of participants themselves. Studies of the NIM scale have shown different effectiveness cutoff values. Some authors (*Kucharski et al., 2007*) have reported that NIM scale was very effective in discriminating between malingering and genuine patients inmates, previously classified by the Structured Interview of Reported Symptoms (SIRS; *Rogers, Bagby & Dickens, 1992*).

The Structured Inventory of Malingered Symptomatology (SIMS) (*Smith & Burger, 1997*) is another test which has been demonstrated to achieve an acceptable level of accuracy in discriminating between malingering and honestly responding individuals (*Smith, 1993*). However, the SIMS test seems less effective when applied to "sophisticated" simulators (e.g., graduate students in psychology) than in the case of malingerers lacking a high level of education (*Vitacco et al., 2007*). Some authors, however, support the effectiveness of the SIMS across genders, cultures and languages and its low reading level, that make it available to a wide range of individuals (*Alwes et al., 2008*; *Smith, 2008*).

The majority of the works aiming at validating malingering detecting instruments in forensics adopt the so-called Simulation Design (SD) paradigm (see *Sellbom et al., 2010*). In SD, participants assigned to the experimental group are explicitly requested to answer by simulating a psychiatric illness while control participants are asked to answer honestly.

A different approach to evaluate the validity of malingering detection instruments is represented by the Known-Group Comparison (KGC) paradigm, where participants are previously subdivided into cathegories (suspected simulators and supposed sincere answerers) by using some supposedly independent criterion such as the Structured Interview of Reported Symptoms (SIRS), a largely used test for detecting simulation behaviors (*Rogers, 1986*; *Rogers, Bagby & Dickens, 1992*) or indications of the psychiatric staff (*Edens, Poytress & Watkins-Clay, 2007*). The KGC paradigm has been used to evaluate MMPI-2 (*Bocaccini, Murrie & Duncan, 2006*; *Barber-Rioja et al., 2009*; *Toomey, Kucharski & Duncan, 2009*), NIM (*Rogers et al., 1998*; *Mogge & LePage, 2004*; *Bocaccini, Murrie & Duncan, 2006*; *Gaines et al., 2007*) and SIMS (*Lewis, Simcox & Berry, 2002*; *Edens, Poytress*

& Watkins-Clay, 2007; *Vitacco et al., 2007*; *Clegg, Fremouw & Mogge, 2009*). However, SIRS has been more recently shown to be a classification criterion less independent than expected because its outcome significantly correlates with the above mentioned tests (*Laffon, 2009*; *Sellbom et al., 2010*).

In the present study we investigated the effectiveness of MMPI-2, NIM and SIMS in detecting malingering behaviors using both KGC and SD paradigms. Participants to KGC were all criminal defendants with a diagnosis of psychiatric disease after imprisonment. They were considered as genuine patients (KGC_Controls) if they did show psychiatric symptoms already before imprisonment and scored above the cutoff of the Symptom Check List-90-Revised (SCL-90-R, a commonly used test for mental illness; *Derogatis, Lipman & Covi, 1973*). As suspected malingerers (KGC_SM) we considered those inmates lacking any psychiatric history before imprisonment, diagnosed as psychiatric patients only after it and scoring below the SCL-90-R cutoff. Participants to SD protocol had no history of psychiatric disease and scored below the SCL-90-R cutoff. They were randomly assigned to either group: Controls (requested to answer honestly, SD_Controls) and simulated malingerers (requested to feign a psychiatric disease, SD_SM). All participants were then submitted to MMPI-2, NIM and SIMS.

METHODS

Participants

Participants ($n = 151$, mean age 39.3 ± 11 (SD), range 22–72) were male inmates of Penitentiary Institutes of the North of Italy, voluntarily participating to the study after providing their written informed consent. The experimental procedure was approved by the Institutional Review Board of the Department of Penitentiary Administration of the Piedmont and Valle D'Aosta Region (approval n. 49720/09). All the procedures complied with the Helsinki Declaration of Ethical Principles for Medical Research. Participants, anonymity was fully preserved.

All participants were submitted to SCL-90-R, a self-report questionnaire designed to assess the presence and the severity of psychological symptoms during the last week before test administration (*Derogatis, Lipman & Covi, 1973*; Italian adaptation by *Sarno et al., 2011*).

There were two different experimental designs: KGC and SD (see Table 1). The KGC design involved defendants with a diagnosis of psychiatric illness after imprisonment (DSM IV-TR Axis I disorders, see Table 1 and Table 2). They were assigned to either two categories on the basis of the following criteria: inmates diagnosed only after imprisonment and scoring below the SCL-90-R cutoff were classified as Suspected Malingerers (KGC_SM). Inmates having a positive psychiatric history already before imprisonment and scoring above the SCL-90-R cutoff (T value ≥ 55 for at least two primary symptom dimensions or for at least one global distress index) were considered as honest responders (KGC_Controls).

It should be clarified that being diagnosed after imprisonment and not scoring high on a self-measure of psychiatric symptomatology does not indicate *per se* a situation of

Table 1 Criteria used to assign participants to the four groups of the present study.

	Known-Group Comparison (KGC)	Simulation Design (SD)
Controls	$n = 35$ –Mental disorders (Axis I) diagnosed *before and after* imprisonment –*Above* SCL-90-R cutoff –Requested to respond honestly	$n = 36$ –No history of mental disorders, free from psychoactive drugs –Below SCL-90-R cutoff –Requested to respond honestly
Malingerers (SM)	$n = 22$ –Mental disorders (Axis I) diagnosed *only after* imprisonment –Below SCL-90-R cutoff –Requested to respond honestly	$n = 35$ –No history of mental disorders, free from psychoactive drugs –Below SCL-90-R cutoff –Requested to simulate mental illness

Notes.

Axis I refers to DSM IV-TR classification (*American Psychiatric Association, 2000*). The names of the groups of participants adopted in the text arise from the combination of those used in this Table: KGC_SM, suspected malingerers; SD_SM, simulated malingerers; KGC_Controls, genuine patients; SD_Controls, controls of the SD requested to respond honestly. For clarifications, see text.

malingering. However, psychiatric diagnoses in correctional institutes are often obtained in a quite easy way, particularly when the inmates report symptoms concerning anxiety or depression which usually lead to the prescription of largely distributed drugs (i.e., antidepressants). In our work we decided to create the "suspected malingerers" group by selecting a very specific subgroup of inmates characterized by the absence of any psychiatric illness before imprisonment, negative at the SCL-90-R test and therefore considered as psychiatric patients only because of a positive diagnosis obtained while in prison. Conversely, honest responders (KGC_Controls) were participants with a personal history of psychiatric illness already assessed before imprisonment and positive at SCL-90-R test. It may be argued that suspected malingerers had no reasons to feign at the tests even if they did it after imprisonment to gain some benefit. In our view, when an individual starts a malingering behavior it becomes an adaptive model of behavior. Moreover, all participants were unaware about the purpose of the study.

Participants to the SD paradigm were inmates with no history of psychiatric symptoms, scoring below the cutoff on the SCL-90-R. They were randomly assigned to either two groups. Participants of the first group (Simulated Malingerers, SD_SM) were asked to simulate a psychiatric illness in answering the tests. Those of the second group (SD_Controls) were requested to answer honestly. For a summary of the experimental/control groups see Table 1.

Instruments
SCL-90-R

The scale Symptom Check List-90 Revised (*Derogatis, Lipman & Covi, 1973*) is a self-report questionnaire designed to assess the presence (and the severity) of psychiatric symptoms during the last week before the test. The test consists of 90 items, five-step responses from 0 (not at all) to 4 (very much), that identify nine primary symptom dimensions (somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation and psychoticism) and three global distress indexes (GSI, PSDI and PST). Participants were considered positive at the test if they scored ≥ 55 (cutoff T value) for at least two dimensions or for at least one global

Table 2 Socio-demographic data of the participants.

	Known-Group Comparison		Simulation Design	
	KGC_SM ($n = 22$)	KGC_Controls ($n = 35$)	SD_SM ($n = 35$)	SD_Controls ($n = 36$)
Age: mean±SD	36 ± 10	38 ± 8	44 ± 10	41 ± 11
Education level: *n* (%)				
> = High school	2 (9)	1 (3)	3 (9)	5 (14)
Middle school	14 (64)	30 (86)	20 (57)	23 (64)
Elementary school	6 (27)	4 (11)	12 (34)	8 (22)
Nationality: *n* (%)				
Italy	10 (43)	20 (57)	*29 (80)*	*15 (42)*
Europe	4 (17)	1 (3)	*3 (11)*	*7 (19)*
Extra-Europe	8 (39)	14 (40)	*3 (9)*	*14 (39)*
Type of Crime: *n* (%)				
Personal injuries	5 (23)	5 (14)	*7 (20)*	*6 (17)*
Drugs	10 (45)	18 (52)	*22 (63)*	*13 (36)*
Property Crimes[*]	7 (32)	12 (34)	*6 (17)*	*17 (47)*
Judicial status: *n* (%)				
Condemned	2 (9)	1 (3)	3 (9)	5 (14)
Waiting for trial	20 (91)	34 (97)	32 (91)	31 (86)
Psychiatric disease: *n* (%)				
OCD	3 (14)	1 (3)	–	–
Anxiety Disorders	12 (55)	27 (77)	–	–
Sleep Disorders	3 (14)	2 (6)	–	–
Adaptation Disorders[**]	3 (14)	3 (9)	–	–
Psychotic Disorder	1 (5)	2 (6)	–	–

Notes.

Differences between simulators and controls for both KGC and SD paradigms were tested by One-way ANOVA (Age) and Fisher's Exact Test (the remaining data). Significance threshold was set at $p < 0.05$. Numbers in italics identify significant differences between the two groups within the same experimental design. Significant differences for Nationality and Type of Crime are present in the SD group only (see text for a discussion). All participants were fluent in Italian and fully understood the questions of the questionnaires. This ruled out the possibility of a significant bias due to cultural aspects (*Correa, 2010*). OCD, Obsessive Compulsive Disorder; *, fraud, theft, robbery; **, Adaptation Disorders include clinically significant emotional or behavioral symptoms that develop in response to one or more identifiable psychosocial stressors; KGC_SM, suspected malingerers; SD_SM, simulated malingerers.

distress index. The coefficients of internal consistency for the nine clinical scales were satisfactory (*Derogatis, Rickels & Rock, 1976*; *Horowitz et al., 1998*) as well as the test-retest reliability (*Derogatis, Rickels & Rock, 1976*).

MMPI-2

The MMPI-2 (*Butcher et al., 1989*; Italian version by *Pancheri & Sirigatti, 1995*) is a multidimensional test of psychological assessment that identifies the pathological personality traits. It consists of 567 dichotomous items (true/false) divided into a series of clinical and validity scales. The Fp (Infrequency Psychopathology) scale consists of 27 items endorsed by less than 20% of both hospitalized psychiatric patients and MMPI-2 normative group. The Fc (Criminal Offender Infrequency) scale is composed of 51 items endorsed by less than 15.5% individuals in *Mergargee*'s (*2004*; *2006*) correctional sample and by less than

15% of the MMPI-2 normative sample. More recently, after verifying that non-malingering inmates often elevated both the F and Fp scales likely because of their peculiar condition, *Mergargee* (*2004*; *2006*) developed the Criminal Offenders Infrequency (Fc) scale which included items rarely endorsed by criminal offenders (*Barber-Rioja et al., 2009*). In the present work, we evaluated Fp and Fc validity scales only, being the most sensitive scales of MMPI-2 for malingering detection in criminal defendants.

NIM

The Negative Impression Management scale of the Personality Assessment Inventory (*Morey, 1991*) is composed of 9 items that are frequently endorsed by participants having a very negative image of themselves. NIM is characterized by high test-retest reliability (*Morey, 1991*; *Boyle & Lennon, 1994*; *Rogers et al., 1995*) and by an adequate construct validity (*Morey, 1996*). In this work, we administered an Italian translation of the NIM authorized by the Author.

SIMS

The Structured Inventory of Malingered Symptomatology (*Smith & Burger, 1997*, Italian version by *La Marca et al., 2012*) is a psychometric instrument for malingering detection in people with at least a primary school education level. Several studies have demonstrated a high-degree of construct validity (*Smith & Burger, 1997*; *Poytress, Edens & Watkins, 2001*; *Merckelbach & Smith, 2003*).

Procedure

Before the evaluation, a numbered form with socio-demographic data (nationality, education level, type of crime, judicial status, psychiatric history, eventual medications, proficiency in Italian language) was filled for each participant. The participant's name was then written on a separate sheet of paper stapled on each numbered form. They were then asked to answer to the questions of the SCL-90-R test. Few days after, on the basis of their psychiatric history and of the results of the SCL-90-R, participants were assigned to the categories of the two experimental designs: SD (participants with a negative psychiatric history, negative at the SCL-90-R) and KGC (participants with a positive psychiatric history, either positive or negative at the SCL-90-R). SD participants were randomly assigned to either the SD_SM or the SD_Control groups. KGC participants were subdivided into two subgroups on the basis of the combination between psychiatric history and SCL-90-R results (see Table 1). At this stage, however, participants were not told about their assignment to avoid any exchange of information between them. The day of tests administration, each participant was given an envelope containing a copy of MMPI-2, SIMS and NIM tests and a pencil. A piece of paper with the same identification number of the socio-demographic form was stapled to the envelope. The sequence of the three tests was counterbalanced among participants to mitigate order effects. Participants assigned to the group of SD_Simulated Malingerers were individually told (separately from the others) to fill the tests as if they were a psychiatric patient. Participants assigned to the remaining three groups (SD_Controls, KGC_Controls and KGC_Suspected Malingerers) were individually told to always answer honestly. The time to fill the questionnaires was

approximately 2 h, without any time constraint imposed on participants. At the end of the procedure, the piece of papers with the name previously stapled to the socio-demographic form was destroyed and every test was associated to the specific socio-demographic form according to the numerical code only.

Twenty-three participants were excluded from the analysis because they did not fit the grouping criteria: for the KGC design, three participants getting a psychiatric diagnosis only after imprisonment were positive at SCL-90-R and 16 participants having a double psychiatric diagnosis (before and after imprisonment) were negative at SCL-90-R. For the SD, four participants were positive at SCL-90-R. The final number of participants was therefore 128, divided into the four groups according to the criteria shown in Table 1.

Statistical analysis

Statistical analyses were performed by Statistica 8.0 (Stat Soft Inc., USA). Demographic data were submitted to both one-way ANOVA (for age) and Fischer's exact test (remaining data). Correlations between malingering measures (SIMS, NIM, Fp and Fc scales of MMPI-2) and SCL-90-R were calculated by Pearson's r. Group differences were assessed by ANOVA performed on tests scores. Post hoc analyses (Newman Keuls) were performed only when factors/interactions were significant at the ANOVA. Cohen's d effect size (*Cohen, 1988*) was determined by using the tool at: https://www.uccs.edu/lbecker/ Receiver Operating Characteristics (ROC) analysis was performed with SPSS 20 (IBM Corp., USA).

RESULTS

Table 2 shows the main socio-demographic characteristics of the participants to the study. Note that the most represented psychiatric symptoms among KGC participants were anxiety-related disorders. Severe diseases (i.e., psychosis) are indeed pretty rare in normal correctional institutes, as severe psychiatric patients are often found non-guilty because of their insanity and are detained in special psychiatric correctional institutes. For the SD only, some significant difference emerged between SD_SM and SD_Controls for Nationality and Type of Crime.

Table 3 displays the correlations among malingering detecting measures and between these measures and the SCL-90-R. Correlations were computed on all participants in order to take into consideration the whole spectrum of possible behaviors. At this stage, in fact, a subdivision into experimental groups would have reduced the statistical power and would have been outside the aim of this analysis. From the table, it emerged a considerable variability of the correlation values among the scales (values ranging from $r = 0.38$ to $r = 0.89$). More in detail, SIMS better correlated with NIM and, among MMPI-2 indexes, the Fc scale was the one with higher correlation with both NIM and SIMS. The value of correlation between Fp and Fc was very high ($r = 0.89$) as well as the correlation between SIMS and NIM ($r = 0.73$). It is important to note the poor correlation between the instruments used for malingering detection and the SCL-90-R, which (together with the psychiatric diagnosis criterion) was used to categorize genuine patients and suspected malingerers in the KGC paradigm. Even though other authors propose the use of SCL-90-R as an indicator of malingering (*Sullivan & King, 2010*), the lack of correlations shown by

Table 3 Scales intercorrelation for the whole sample ($n = 128$).

	NIM	Fp	Fc	SCL_90R
SIMS	0.73[*]	0.51[*]	0.62[*]	−0.14
NIM		0.38[*]	0.51[*]	−0.12
Fp			0.89[*]	−0.13
Fc				−0.12

Notes.

Raw scores were used for each scale.

Fp, Infrequency Psychopathology scale; Fc, Infrequency Criminal Offender scale; SIMS, Structured Inventory of Malingered Symptomatology scale; NIM, Negative Impression Management scale.

Asterisks indicate statistically significant correlations ($p < 0.05$).

Table 4 Mean scores ± SD of the four scales assessed in the two experimental designs.

	Known-Group Comparison			Simulation Design		
	Suspected malingerers	Controls	Cohen's d	Simulated malingerers	Controls	Cohen's d
SIMS	20.4 ± 11.3[*]	15.1 ± 11.9[*]	0.43	37.5 ± 8.8[*]	9.2 ± 5.8[*]	3.80
NIM	5.2 ± 4.0	5.0 ± 4.9	0.04	13.8 ± 4.6[*]	2.9 ± 2.9[*]	2.83
Fp	5.2 ± 4.5	5.1 ± 3.7	0.02	9.7 ± 5.4[*]	4.6 ± 2.1[*]	1.24
Fc	8.2 ± 8.7	8.7 ± 6.7	0.06	18.5 ± 8.9[*]	6.3 ± 4.4[*]	1.74

Notes.

Asterisks indicate the presence of a statistically significant difference ($p < 0.05$) at the post-hoc analysis (Newman Keuls). For other abbreviations, see text. Cohen's d measures the size of the effect. Values are considered as negligible (<0.20), small (0.20–0.49), medium (0.50–0.79), or large (≥ 0.80).

our results supports the use of SCL-90-R as possible external, independent criterion for grouping subjects in the KGC design.

Table 4 and Fig. 1 show the main results for both KGC and SD protocols. From figure inspection it emerges the different sensitivity of the instruments in discriminating malingerers from controls in the two investigated protocols, with clear differences for all the tests in the SD but with evident differences for the SIMS only in the KGC. An ANOVA performed on individual scales (SIMS, NIM, Fp and Fc) using Design (KGC, SD) and Malingering (SM, Controls) as factors showed the significance of both of them (Design, $F(4, 121) = 6.2$, $p = 0.0001$; Malingering, $F(4, 121) = 26.4$), $p = 0.0001$) and the significance of the interaction between the two (Design × Malingering, $F(4, 121) = 17.8$, $p = 0.0001$). Post-hoc analysis (Newman Keuls, $p < 0.05$) revealed that while all the four instruments were highly sensitive in discriminating simulated malingerers (SD_SM) from controls (SD_Controls) in the SD protocol, only SIMS reached the statistical significance in discriminating between suspected malingerers (KGC_SM) from controls (KGC_Controls) ($p = 0.01$) in the KGC protocol.

The sensitivity of SIMS in detecting suspected malingerers in the KGC design with respect to the other instruments was also shown by the Receiver Operating Characteristics (ROC) curves calculated for the four scales (SIMS, NIM, Fp and Fc) and presented in Table 5. Areas Under the Curve (AUC) values of KGC were in general much smaller than those of SD. However, SIMS AUC was the largest one among the four scales in KGC
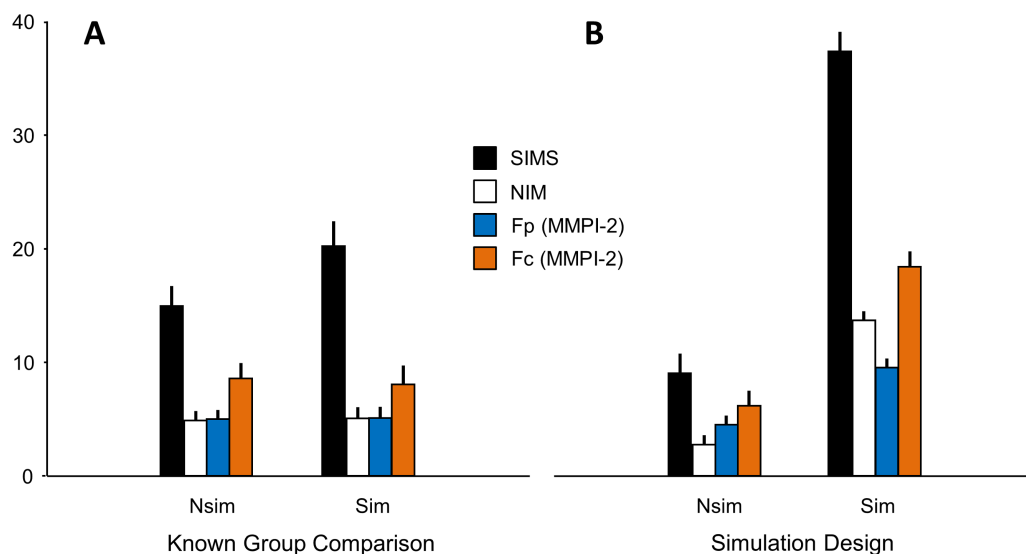
Figure 1 **Mean scores ± SEM for SIMS, NIM, Fp and Fc psychometric tests.** Nsim and Sim refer to the two groups of the Know Group Comparison (A) and Simulation Design (B) paradigms (see text). Ordinate: tests scores.

Table 5 **Areas under the Receiver Operating Characteristics Curve (AUC).**

	KGC (*n* = 57)		SD (*n* = 71)	
	AUC	SEM	AUC	SEM
SIMS	0.553	0.055	0.940	0.022
NIM	0.433	0.058	0.926	0.031
Fp	0.404	0.061	0.771	0.056
Fc	0.392	0.060	0.836	0.045

Notes.

Values ± Standard Error of Mean (S.E.M.) are shown for the instruments tested in the two experimental designs, KGC and SD. AUC, area under curve; SEM, standard error of mean. For other abbreviations, see text.

suggesting a better sensitivity of this measure with respect to the others. ROC curves for the KGC protocol are shown in Fig. 2.

Altogether, these findings show that SIMS was the only instrument among the evaluated ones characterized by a sufficient level of sensitivity to detect suspected malingerers in a KGC design and drove us to investigate the results of cut-off data (see Table 6).

As shown by Table 6, cut-off data were in line with the results of the ANOVA, qualitatively showing an excellent capability of SIMS in discriminating simulated malingerers from controls in the SD (100%). When looking at the KGC, however, the picture became much less clear. In fact, while 77% of suspected malingerers were detected as such, about one half of control participants (51%) were detected as malingerers as well. This observation testifies a quite high proportion of false positives in the population of psychiatric patients, already diagnosed before imprisonment and scoring above cutoff on the SCL-90-R. An alternative interpretation could be that the supposed sensitivity of SCL-90-R as
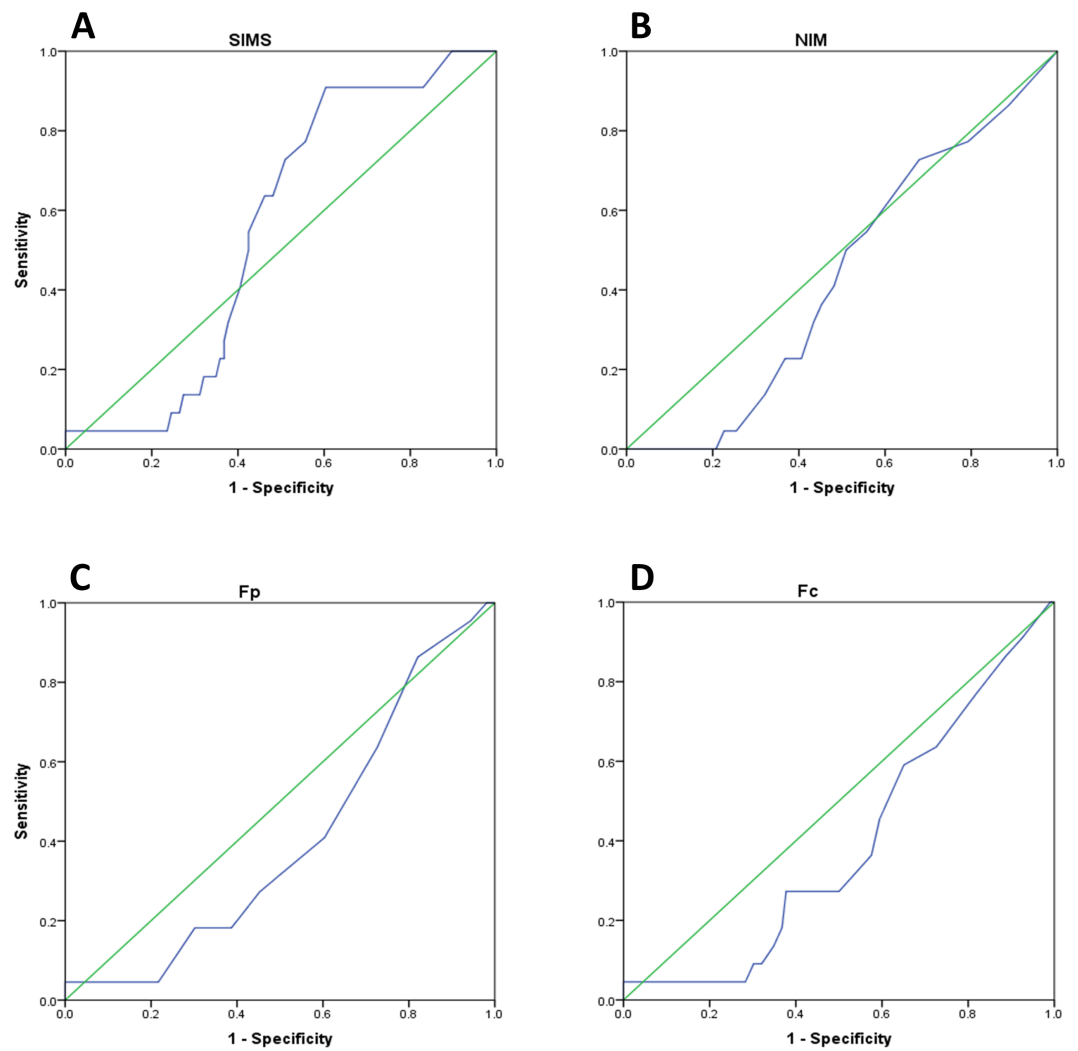
Figure 2 **ROC curves for the Known-Group Comparison paradigm.** (A) SIMS, (B) NIM, (C) Fp and (D) Fc Receiver Operating Characteristic curves (blue). The green diagonal in each panel indicates the reference value of 0.5 AUC (area under curve). For other abbreviations see text.

Table 6 **Percentage of participants detected by SIMS as malingerers and non-malingerers in the two experimental designs (cut-off >= 14).**

KGC				SD			
Suspected malingerers		Controls		Simulated malingerers		Controls	
Nsim	Sim	Nsim	Sim	Nsim	Sim	Nsim	Sim
23%	77%	49%	51%	0%	100%	83%	17%

Notes.
Abbreviations as in Fig. 1.

an indicator of malingering (*Sullivan & King, 2010*), although not corroborated by the correlation study we showed in the present work, could have revealed a subpopulation of malingering individuals among those assumed to be genuine patients. Although unlikely (KGC_Controls were all diagnosed as psychiatric patients already before imprisonment), some exaggeration of symptoms remains possible making evident the necessity to proceed with caution when using a single instrument to detect psychiatric malingering in inmates populations.

DISCUSSION

The goal of the present study was to determine the effectiveness of three instruments used to detect malingering of psychiatric illness (Fp and Fc validity scales of the MMPI-2, NIM and SIMS) in a group of criminal defendants suspected of malingering because they manifested psychiatric symptoms only after imprisonment but scored negatively at the SCL-90-R (suspected malingerers, KGC_SM). As KGC_Controls we considered those inmates with a positive psychiatric diagnosis formulated both before and after imprisonment and positive at SCL-90-R (double convergence). All participants were requested to answer honestly to the tests. Moreover, because most of the tests have been validated by Simulation Design paradigms, we evaluated the same instruments by SD paradigm as well. To this purpose, we recruited inmates with negative psychiatric history and negative at the SCL-90-R. They were randomly assigned to two groups. The first group (simulated malingerers, SD_SM) included inmates asked to simulate psychiatric illness. The second group (SD_Controls) was formed by inmates requested to answer honestly to the tests. Results show that for the SD all the measures obtained high rates of accuracy in detecting malingering and were rather similar in terms of global performance indicators (ANOVA and AUC values). This finding is in agreement with what already reported in literature (for recent reviews on the use of SIMS in SD/KGC studies see *Wisdom, Callahan & Shaw, 2010*; *Van Impelen et al., 2014*). Conversely, for the KGC, three out of the four instruments showed poor performance in detecting malingering. ANOVA post-hoc analysis showed that only SIMS reached the significance threshold in discriminating KGC_SM from KGC_Controls.

While the effectiveness of the SIMS in detecting malingering in SD paradigms is largely confirmed by the literature (for an overview, see *Edens, Otto & Dwyer, 1999*), the case of KGC paradigms is much more complicated. On one side the number of studies using SIMS in KGC designs is much less than those based on SD (*Wisdom, Callahan & Shaw, 2010*; *Van Impelen et al., 2014*). On the other side, the criteria used to group individuals in KGC protocols often raise serious concerns. The most used criterion is the performance on the Structured Interview of Reported Symptoms (SIRS, *Rogers, Bagby & Dickens, 1992*; *Edens, Poytress & Watkins-Clay, 2007*; *Sellbom et al., 2010*). The SIRS uses multiple strategies to detect feigned psychopathology, such as absurd symptoms, unlikely combinations of symptoms, discrepancies between reported and observed symptoms and referred abnormal severity of symptoms. The SIRS has been well studied; a meta-analysis (*Green & Rosenfeld, 2011*) yielded a sensitivity (i.e., the likelihood of a positive symptom validity test, SVT result in feigners) of 0.49 and a specificity (i.e., likelihood of a negative SVT result in honest responders) of 0.95. The efficacy of SIRS in classifying KGC groups for the validation

of the three tests we used in the present study has been shown for NIM (*Rogers et al., 1998*; *Mogge & LePage, 2004*; *Bocaccini, Murrie & Duncan, 2006*; *Gaines et al., 2007*), SIMS (*Lewis, Simcox & Berry, 2002*; *Edens, Poytress & Watkins-Clay, 2007*; *Vitacco et al., 2007*; *Clegg, Fremouw & Mogge, 2009*) and MMPI-2 (*Bocaccini, Murrie & Duncan, 2006*; *Toomey, Kucharski & Duncan, 2009*; *Barber-Rioja et al., 2009*). However, a strong concern on the use of SIRS to categorize participants in malingering-detecting KGC designs arises from the fact that it correlates with the majority of the other tests. *Sellbom et al. (2010)* reported a significant intercorrelation between SIRS and SIMS and between SIRS and NIM. Similar conclusions were reached by *Laffon (2009)* who also showed a correlation between SIRS and both, SIMS and NIM. Thus, the good sensitivity of SIRS in detecting malingering could arise from the fact that it was validated by using instruments strongly correlating with it. Moreover, according to Calhoun and coworkers (*2000*) SIRS seems to misclassify true patients as malingerers.

For these reasons, in our KGC paradigm we decided to use clinical criteria to create categories for the KGC design (psychiatric diagnosis and SCL-90-R) instead.

By doing this, we have been able to show that only the SIMS was able to discriminate suspected malingerers from controls. It should be stressed, however, that *Edens, Otto & Dwyer (1999)*, found a positive correlation between SIMS and the GSI scale of the SCL-90-R and this argument could be used to generate a criticism similar to that we used for SIRS. However, the KGC group of suspected malingerers of our work was formed by individuals all scoring below the cutoff for SCL-90-R test. This should protects us from the criticisms that would arise from a situation similar to that of SIRS as underlined by *Sellbom et al. (2010)* and *Laffon (2009)*. On the contrary, the absence of correlation between SCL-90-R and SIMS, NIM, Fp and Fc (r ranging from -0.12 to -0.14) seems to corroborate our approach.

In other words, we are aware of the difficulty in creating reliable known groups in a KGC design involving psychiatric inmates but in our study we used two criteria to extract a very specific subgroup of participants: inmates lacking any history of psychiatric diseases before incarceration and negative at the SCL-90-R test. According to these criteria, they were considered as 'psychiatric patients' only because of a diagnosis obtained after imprisonment. Such diagnoses are pretty easy to obtain because their main outcome, in general, is the administration of some anxiolytic or antidepressant medication (largely distributed in the prison environment and ''appreciated'' by both doctors and inmates for obvious practical reasons). Indeed, depression and mood-related disorders were the most represented in our population. One may argue that a further validation of our approach would have been provided by testing a further category in the KGC paradigm: inmates diagnosed by a psychiatrist only after imprisonment and scoring above threshold at the SCL-90-R test. In this case, no malingering behavior should be expected. In our view this would have only been a further confirmation of what we already report here. Moreover, the dimension of our sample did not allow the investigation of groups others than those described in the present study.

Rogers (*1997a*, *1997b*) argued that the best approach to validate malingering-detecting psychometric tools is to compare validity measures obtained by both SD and KGC

paradigms. Nevertheless, only few studies have been conducted using both SD and KGC on the same population (see *Wisdom, Callahan & Shaw, 2010*; *Van Impelen et al., 2014*). As far as we know, only one paper evaluated the efficacy of SIMS in detecting malingering in a group of defendants classified as supposed simulators by a psychological screening (*Edens, Poytress & Watkins-Clay, 2007*). According to this study, SIMS failed in detecting malingering. Here we found an opposite, positive result likely because of the different criteria used to categorize participants. At the same time, however, despite SIMS sensitivity, we showed a relatively poor specificity of this instrument (see the ROC analysis and the cutoff evaluation in our Results Section). We therefore agree with *Edens, Poytress & Watkins-Clay (2007)*, when they say that SIMS specificity is pretty poor when administered to honestly responding, symptomatic individuals.

CONCLUSIONS

Research to validate malingering-detecting tools is inherently problematic. Both over-identification and under-identification of feigning or exaggeration are reasons of concern. Despite our finding that SIMS significantly discriminates malingering defendants from controls in the KGC design, the poor specificity of this instrument lead us to conclude that one should be very cautious in using SIMS as a stand-alone measure. As far as we know this is one among the few studies where psychometric tools for malingering detection have been validated by both, Simulation Design and Known-Group Comparison. The main problem of Known-Group Comparison approaches is represented by the method adopted to create the experimental groups. Here we used both clinical and psychometric criteria and we are aware that the grouping criteria we decided *a priori* here are somewhat arbitrary. However, we are convinced that they could provide some useful cues to evaluate the solidity of malingering testing instruments outside the "protected environment" of the Simulation Design. Other works, on the contrary, have adopted psychometric instruments only (mainly the SIRS). Alternative grouping criteria as well as malingering behaviors in socio-demographic groups others than the European males evaluated by the present work should be explored by additional investigations. Finally, our results cannot be generalized to other populations often associated with a significant degree of malingering (e.g., personal injury litigants) different from criminal defendants. In conclusion, understanding what malingering looks like, how to design and use assessment measures and how to appropriately manage malingering behaviors remain therefore challenging questions which require responsible empirical approaches.

ACKNOWLEDGEMENTS

We thank the Direction of the Correctional Institutes for providing support to the study, LC Morey for his permission to use the NIM test, F Del Corno for his preliminary suggestions, R Carbone for his clinical support, C Bonifazzi for his suggestions on statistical analysis, and L Fadiga, T Merten, X Seron and G Zara for their comments on earlier versions of this manuscript.

Funding

The authors received no funding for this work.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Barbara De Marchi conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Giulia Balboni conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The Institutional Review Board of the Provveditorato Regionale dell'Amministrazione Penitenziaria di Piemonte e Valle D'Aosta approved this research (approval n. 49720/09).

Data Availability

The raw data are provided in a Data S1.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.5259#supplemental-information.

REFERENCES

Alwes YR, Clark JA, Berry DTR, Granacher RP. 2008. Screening for feigning in a civil forensic setting. *Journal of Clinical and Experimental Neuropsychology* **30(2)**:133–140 DOI 10.1080/13803390701260363.

American Psychiatric Association. 2000. *DSM-IV-TR Manuale diagnostico e statistico dei disturbi mentali.* Milano: Masson.

Arbisi PA, Ben-Porath YS. 1995. An MMPI-2 infrequent response scale for use with psychopathological populations: the Infrequency Psychopathology Scale, F(p). *Psychological Assessment* **7(4)**:424–431 DOI 10.1037/1040-3590.7.4.424.

Baity MR, Siefert JF, Chambers A, Blais M. 2007. Deceptiveness on the PAI: a study of naïve faking with psychiatric inpatients. *Journal of Personality Assessment* **88**:16–24 DOI 10.1080/00223890709336830.

Barber-Rioja V, Zottoli TM, Kucharski LT, Duncan S. 2009. The utility of the MMPI-2 criminal offender infrequency (Fc) scale in the detection of malingering in criminal defendants. *International Journal of Forensic Mental Health* **8**:16–24 DOI 10.1080/14999010903014689.

Bass C, Halligan P. 2014. Factitious disorders and malingering: challenges for clinical assessment and management. *The Lancet* **38**:1422–1432 DOI 10.1016/S0140-6736(13)62186-8.

Berry DTR, Cimino CR, Chong NK, LaVelle SN, Ho IK, Morse TL, Thacker SR. 2001. MMPI-2 fake-bad scales: an attempted cross-validation of proposed cutting scores for outpatients. *Journal of Personality Assessment* **76(2)**:295–314 DOI 10.1207/S15327752JPA7602_11.

Bocaccini MT, Murrie DC, Duncan SA. 2006. Screening for malingering in a criminal-forensic sample with the Personality Assessment Inventory. *Psychological Assessmen* **18**:415–423 DOI 10.1037/1040-3590.18.4.415.

Bowen C, Bryant RA. 2006. Malingering posttraumatic stress on the Personality Assessment Inventory. *International Journal of Forensic Psychology* **1**:22–28 DOI 10.1207/s15327752jpa7103_3.

Boyle GJ, Lennon TJ. 1994. Examination of the reliability and validity of the Personality Assessment Inventory. *Journal of Psychopathology and Behavioral Assessment* **6(3)**:173–187 DOI 10.1007/BF02229206.

Butcher JN, Dahlstrom WG, Graham JR, Tellegen A, Kaemmer A. 1989. *Manual for the administration and scoring of the MMPI-2.* Minneapolis: University of Minnesota Press.

Calhoun PS, Earnst KS, Tucker DD, Kirby AC, Beckham JC. 2000. Feigning combat related posttraumatic stress disorder on the Personality Assessment Inventory. *Journal of Personality Assessment* **75**:338–350 DOI 10.1207/S15327752JPA7502_11.

Clegg C, Fremouw W, Mogge N. 2009. Utility of the Structured Inventory of Malingered Symptomatology (SIMS) and the Assessment of Depression Inventory (ADI) in screening for malingering among outpatients seeking to claim disability. *The Journal of Forensic Psychiatry & Psychology* **20**:239–254 DOI 10.1080/14789940802267760.

Cohen J. 1988. *Statistical power analysis for the behavioral sciences.* Second Edition. Hillsdale: Lawrence Erlbaum Associates.

Correa AA. 2010. Validation of the Spanish SIRS: beyond linguistic equivalence in the assessment of malingering among Spanish speaking clinical population. Master's thesis, University of North Texas.

Derogatis LR, Lipman RS, Covi L. 1973. SCL-90: an outpatient psychiatric rating scale-Preliminary report. *Psychopharmacology Bulletin* **9(1)**:13–27.

Derogatis LR, Rickels K, Rock A. 1976. The SCL-90 and the MMPI-A step in the validation of a new self-report scale. *British Journal of Psychiatry* **128**:280–289 DOI 10.1192/bjp.128.3.280.

Douglas KS, Otto RK, Borum R. 2003. *Handbook of psychology, Vol. 2.* Hoboken: Wiley, 190–211.

Edens JF, Cruise KR, Buffington-Vollum JK. 2001. Forensic and correctional application of the Personality Assessment Inventory. *Behavioral Sciences and the Law* **19**:519–543 DOI 10.1002/bsl.457.

Edens JF, Otto RK, Dwyer T. 1999. Utility of the structured inventory of malingered symptomatology in identifying persons motivated to malinger psychopathology. *Journal of the American Academy of Psychiatry and the Law* **27(3)**:387–396.

Edens JF, Poytress NG, Watkins-Clay MM. 2007. Detection of malingering in psychiatric unit and general population prison inmates: a comparison of the PAI, SIMS, and SIRS. *Journal of Personality Assessment* **88(1)**:33–42 DOI 10.1080/00223890709336832.

Gaines MV, Giles CL, Morgan RD, Steffan JS. 2007. An examination of the combined use of the PAI and the M-FAST in detecting malingering among inmates. Poster presented at the annual meeting of the American Psychological Association, San Francisco, CA, USA.

Green D, Rosenfeld B. 2011. Evaluating the gold standard: a review and meta-analysis of the structured interview of reported symptoms. *Psychological Assessment* **23**:95–107 DOI 10.1037/a0021149.

Horowitz LM, Rosenberg SE, Baer BA, Ureno G, Villasenor VS. 1998. Inventory of interpersonal problems: psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology* **56**:885–892.

Kucharski LT, Toomey JP, Fila K, Duncan S. 2007. Detection of malingering of psychiatric disorder with the personality assessment inventory: an investigation of criminal defendants. *Journal of Personality Assessment* **88(1)**:25–32 DOI 10.1080/00223890709336831.

La Marca S, Rigoni D, Sartori G, Lo Priore C. 2012. *Structured inventory of malingered symptomatology*. Firenze: Organizzazioni Speciali.

Laffon LL. 2009. Detecting feigning a correctional setting: a comparison of multiple measures. D. Phil. Thesis, Pacific University.

Lees-Haley PR, Dunn J. 1994. The ability of naïve subjects to report symptoms of mild brain injury, post-traumatic stress disorder, major depression, and generalized anxiety disorder. *Journal of Clinical Psychology* **50**:252–256 DOI 10.1002/1097-4679(199403)50:2<252::AID-JCLP2270500217>3.0.CO;2-T.

Lewis JL, Simcox AM, Berry DTR. 2002. Screening for feigned psychiatric symptoms in a forensic sample by using the MMPI-2 and the Structured Inventory of Malingered Symptomatology. *Psychological Assessment* **14**:170–176 DOI 10.1037/1040-3590.14.2.170.

McLaughlin J, Kan LY. 2014. Test usage in four common types of forensic mental health assessment. *Professional Psychology: Research and Practice* **45(2)**:128–135 DOI 10.1037/a0036318.

Meehl PE, Hathway SR. 1946. The K factor as a suppressor variable in the MMPI. *Journal of Applied Psychology* **30**:525–564 DOI 10.1037/h0053634.

Merckelbach H, Smith GP. 2003. Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology* **18(2)**:145–152 DOI 10.1093/arclin/18.2.145.

Mergargee EI. 2004. Development and initial validation of an MMPI-2 infrequency scale (Fc) for use with criminal offenders. Paper presented at the 39th annual symposium on recent developments on the MMPI-2 / MMPI-A. Minneapolis, MN, USA.

Mergargee EI. 2006. Using the MMPI-2 in correctional settings. In: *MMPI-2 a pratictioner's guide*. Washington D.C.: American Psychological Association.

Merten T, Rogers R. 2017. An international perspective on feigned mental disabilities: conceptual issues and continuing controversies. *Behavioral Sciences and Law* **35(2)**:97–112 DOI 10.1002/bsl.2274.

Mittenberg W, Patton C, Canyock EM, Condit DC. 2002. Base rate of malingering and sympom exaggeration. *Journal of Clinical and Experimental Neuropsychology* **24**:1094–1102 DOI 10.1076/jcen.24.8.1094.8379.

Mogge NL, LePage JP. 2004. The Assessment of Depression Inventory (ADI): a new instrument used to measure depression and to detect honesty of response. *Depression and Anxiety* **20**:107–113 DOI 10.1002/da.20033.

Morey LC. 1991. *Personality assessment inventory professional manual.* Odessa: Psychological Assessment Resources.

Morey LC. 1996. *An interpretative guide to the Personality Assessment Inventory.* Odessa: Psychological Assessment Resources.

Morey LC, Lanier WV. 1998. Operating characteristics of six response distortion indicators for the Personality Assessment Inventory. *Assessment* **5**:203–214 DOI 10.1177/107319119800500301.

Noeker M, Petermann F. 2011. Simulation neurologischer versus psychischer Beschwerden. Notwendigkeit unterschiedlicher Validierungsstrategien [Malingering of neurological vs. mental complaints. Necessity of different validation strategies]. *Psychotherapeut* **56**:449–454 DOI 10.1007/s00278-011-0851-2.

Otto RK, Heilbrun K. 2002. The practice of forensic psychology. *American Psychologist* **57**:5–18 DOI 10.1037/0003-066X.57.1.5.

Pancheri P, Sirigatti S. 1995. *Minnesota multiphasic personality inventory-2.* Firenze: Organizzazioni Speciali.

Pope KS, Butcher JN, Seelen J. 2006. *MMPI, MMPI-2 e MMPI-A IN COURT: a pratical guide for expert witnesses and attorneys.* Third Edition. Washington D.C.: American Psychological Association.

Poytress NG, Edens J, Watkins MM. 2001. The relationship between psychopathic personality features and malingering symptoms of major mental illness. *Law and Human Behavior* **25**:567–582 DOI 10.1023/A:1012702223004.

Rogers R. 1986. *Conducting insanity evaluations.* New York: Van Nostrand Reinhold.

Rogers R. 1997a. Introduction. In: Rogers R, ed. *Clinical assessment of malingering and deception.* Second Edition. New York: Guilford Press, 1–19.

Rogers R. 1997b. Researching dissimulation. In: Rogers R, ed. *Clinical assessment of malingering and deception.* Second Edition. New York: Guilford Press, 398–426.

Rogers R. 2008. *Clinical assessment of malingering and deception.* Third Edition. New York: Guildford Press.

Rogers R, Bagby RM, Dickens SE. 1992. *Structured interview of reported symptoms professional manual.* Odessa: Psychological Assessment Resources.

Rogers R, Flores J, Ustad K, Sewell KW. 1995. Initial validation of the Personality Assessment Inventory Spanish version with clients from Mexican American communities. *Journal of Personality Assessment* **64**:340–348 DOI 10.1207/s15327752jpa6402_12.

Rogers R, Sewell KW, Cruise KR, Wang EW, Ustad KL. 1998. The PAI and feigning: a cautionary note on its use in forensic-correctional settings. *Assessment* **5**:399–405 DOI 10.1177/107319119800500409.

Rogers R, Sewell KW, Goldstein AM. 1994. Explanatory models of malingering: a prototypical analysis. *Law and Human Behavior* **18**:543–552 DOI 10.1007/BF01499173.

Sarno I, Preti E, Prunas A, Madeddu F. 2011. *SCL-90-R (Symptom Checklist-90-R) Adattamento italiano*. Firenze: Giunti, Organizzazioni Speciali.

Sellbom M, Bagby RM. 2008. Response styles on multiscale inventories. In: Rogers R, ed. *Clinical assessment of malingering and reception*. Third Edition. New York: Guilford Press, 182–206.

Sellbom M, Toomey JA, Wygant DB, Kucharski LT, Duncan S. 2010. Utility of the MMPI–2-RF (restructured form) validity scales in detecting malingering in a criminal forensic setting: a known-groups design. *Psychological Assessment* **22(1)**:22–31 DOI 10.1037/a0018222.

Seron X. 2014. Lying in neuropsychology. *Clinical Neurophysiology* **44**:389–403 DOI 10.1016/j.neucli.2014.04.002.

Smith GP. 1993. Detection of malingering: a validation study of the SLAM test. D. Phil. Thesis, The University of Missouri-St. Louis Dissertation Abstracts International, 53, 3795B.

Smith GP. 2008. Brief screening measures for the detection of feigned psychopathology. In: Rogers R, ed. *Clinical assessment of malingering and deception*. Third Edition. New York: Guilford, 323–342.

Smith GP, Burger GK. 1997. Detection of malingering: validation of the structured inventory of malingered symptomatology. *Journal of the American Academy of Psychiatry and Law* **25**:183–189.

Sullivan K, King J. 2010. Detecting faked psychopathology: a comparison of two tests to detect malingered psychopathology using a simulation design. *Psychiatry Research* **176(1)**:75–81 DOI 10.1016/j.psychres.2008.07.013.

Toomey JA, Kucharski LT, Duncan S. 2009. The utility of the MMPI-2 malingering discriminant function index in the detection of malingering: a study of criminal defendants. *Assessment* **16**:115–121 DOI 10.1177/1073191108319713.

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.

Vitacco MJ, Rogers R, Gabel J, Munizza J. 2007. An evaluation of malingering screens with competency to stand trial patients: a known-groups comparison. *Law Human Behavior* **31**:249–260 DOI 10.1007/s10979-006-9062-8.

Widows MR, Smith GP. 2005. *Structured inventory of malingered simptomatology*. Odessa: Psicological Assessment Resources.

Wisdom NM, Callahan JL, Shaw TG. 2010. Diagnostic utility of the structured inventory of malingered symptomatology to detect malingering in a forensic sample. *Archives of Clinical Neuropsychology* **25**:118–12 DOI 10.1093/arclin/acp110.

Young G. 2014. *Malingering, feigning, and response bias in psychiatric/ psychological injury*. Berlin: Springer Verlag.