

Assessment of the Standard Forensic Procedure for the Evaluation of Psychological Injury in Intimate-Partner Violence

Francisca Fariña¹, Ramón Arce², Manuel Vilariño² and Mercedes Novo²

¹ Universidad de Vigo (Spain)

² Universidad de Santiago de Compostela (Spain)

Abstract. In judicial terms, a victim refers to any person who has suffered injury arising from an action or omission of an action that constitutes an offence, and the burden of proof lies with the prosecution. A review of Spanish judicial judgements underscored that the lack of evidence of psychological injury in cases of intimate-partner violence (IPV) accounted for approximately 40% of acquittals. Thus, the Spanish standard of proof for the forensic evaluation of psychological injury i.e., the MMPI-2 and the unstructured interview were assessed in order to determine if they met the statutory requirement for the assessment of psychological injury and the differential diagnosis of feigning. The results of the comparison of 51 women victims of IPV with firm convictions against their aggressors, and 54 women mock victims of IPV showed that the F, K, Fb, Fp and Ds scales, and the F-K index discriminated significantly and with medium and large effect sizes, between adjudicated and mock victims. However, the results did not provide a valid decision criterion for forensic settings i.e., false negatives (identifying feigner as honest protocols) were not classified correctly. In conclusion, the standard forensic procedure for the evaluation of psychological injury in cases of IPV did not constitute valid proof for judges who acquitted defendants on the grounds of not proven due to the lack of evidence of psychological injury.

Received 8 June 2012; Revised 29 April 2013; Accepted 12 July 2013

Keywords: intimate-partner violence, psychological injury, judicial proof, feigning, MMPI-2.

The approval of the Spanish Law on Preventive Measures against Gender Violence (Ley Orgánica 1/2004 de 28 de diciembre, de Medidas de Protección Integral contra la Violencia de Género) poses new challenges for forensic psychological assessment. A crime victim is legally defined as persons who, individually or collectively, has suffered harm, including physical or psychological injury, emotional suffering, economic loss or substantial impairment of their fundamental rights, through acts or omissions that are in violation of criminal laws operative within Member States, including those laws proscribing criminal abuse of power (United Nations, 1998). The burden of proof lies with the prosecution who must prove a victim has suffered harm such as psychological injury and emotional suffering. The psychological effects of intimate-partner violence (IPV) are measured in terms of Posttraumatic Stress Disorder (PTSD),

that is comorbid with depression, social maladjustment, anxiety, and sexual dysfunctions (Arce, Fariña, Carballal, & Novo, 2009; Bargai, Ben-Shakhar, & Shalev, 2007; Kessler, Sonnega, Hughes, & Nelson, 1995), with an estimated comorbidity rate of 92% (Brown, Campbell, Lehman, Grisham, & Mancill, 2001). Notwithstanding, the presence of comorbidity without the simultaneous existence of PTSD does not constitute proof of injury sustained in a traumatic event (O'Donnell, Creamer, Bryant, Schnyder, & Shalev, 2006). The forensic evaluation of PTSD is often problematic since it relies extensively or exclusively on the implicit assumption of the accurate self-reporting of psychological symptoms, making it vulnerable to the feigning or exaggeration of symptoms in order to obtain financial compensation and disability status or to elude criminal responsibility for offences and/or mitigate the associated penalties. Thus, forensic assessment must *always* suspect feigning (American Psychiatric Association, 2000).

Even laypeople in PTSD psychopathological criteria were able to qualify for the diagnosis on a checklist, with feigning rates ranging from 86 to 94% (Resnick, West, & Payne, 2008). In cases of IPV against Spanish laywomen in psychopathology, Arce et al. (2009) found a rate of 84.2%, and Vilariño, Fariña, and Arce (2009)

Correspondence concerning this article should be addressed to Francisca Fariña. Faculdade de Ciências da Educação e do Desporto. Campus A Xunqueira, s/n. 36005. Pontevedra (Spain).

E-mail: francisca@uvigo.es

This research has been sponsored by a grant of the Secretaría de Estado de Educación (Ministerio de Educación y Ciencia) to the project "Reeducación de penados por violencia de género: Implementación y evaluación de programas de tratamiento" (Ref: EDU2011-24561).

a rate of 100%. This highlights the need for psychometric instruments that measure the validity of the protocols, the MMPI-2 being the psychometric instrument most widely used in the Spanish forensic context (Arce, Fariña, & Buela-Casal, 2008; Arch, Jarne, Peró, & Guàrdia, 2011; Fariña, Arce, & Sotelo, 2010; Jiménez, Sánchez, & Tobón, 2009; Pérez-Pareja, Sesé, González-Ordi, & Palmer, 2010), and worldwide (Archer, Buffington-Vollum, Stredny, & Handel, 2006; Bow & Quinnett, 2001; Rogers, Sewell, Martin, & Vitacco, 2003), particularly for the assessment of psychological injury (Arce, Fariña, Carballal, & Novo, 2006; Resnick et al., 2008). Nevertheless, the MMPI-2, as any other psychometric instrument, has its shortcomings in that it does not constitute sufficient standard of proof: a) the diagnosis of feigning is compatible with other alternative hypothesis (e.g., a high F score can be indicative of feigning, random responses or a psychotic disorder); b) it does not correctly classify all of the feigners (Rogers et al., 2003); c) it provides diagnostic impressions not accurate diagnosis given that high scores on the depression scale of the MMPI-2 are not exclusive to depression, but to a wide range of disorders on the DSM-IV (Rogers, 2008a); and d) it fails to establish a causal relationship (legal requirement for forensic psychologists in the inquisitorial justice system). Consequently, decision-making exclusively grounded on these instruments entails systematically incurring in two types of errors: false positives (identifying honest individuals as feigners), and false negatives (identifying feigner as honest). Moreover, the nature of the psychometric instrument itself (a recognition task) does not permit the forensic psychologist to establish, as legally required, a causal relationship between the alleged events and psychological injury. In other words, the psychometric instrument deals with symptoms and disorders, but not with the underlying causal relationships. In order to fulfill the statutory requirement of establishing a causal relationship between the alleged events and psychological injury resulting for IPV, the victim must establish a nexus between the symptoms of psychological injury and the alleged events. However, in cases of IPV, PTSD-like symptoms and other symptoms of anxiety may arise from an Adjustment Disorder linked to the couples separation or divorce that often coincides with allegations of IPV. The requisite of establishing a causal effect can be achieved through either standard or non-standard victim interviews. The latter and its subsequent non-standard diagnosis entail a high rate of diagnostic errors that have been estimated to be about 50% of diagnosis of severe depression (Rogers & Shuman, 2005). Standard interviews are of two types i.e., intrinsically clinical, and interviews designed for psychological assessment in

forensic contexts (Vilariño, Arce, & Fariña, 2013). Clinical Interviews, which constitute the “gold standard” for PTSD assessment, such as the SCID-IV (Spitzer, Williams, Gibbon, & First, 1995), are by definition clinical in nature, but are not concerned with feigning and do not incorporate standard procedures for the assessment of common feigning strategies. These strategies, however, are the precise object of standard procedures for the analysis of interviews in forensic settings, in particular the extensively used Structured Interview of Reported Symptoms, SIRS (Rogers, Bagby, & Dickens, 1992). Notwithstanding, the SIRS has not been validated for the assessment of psychological injury of PTSD in IPV. Moreover, the SIRS is not sensitive to the forensic psychologist’s statutory requirement of establishing a causal relationship given that this nexus is never questioned, but is assumed to be accurately ascertained from the victims allegations.

A systematic review of Spanish judicial judgments in cases of IPV, Arce, Alonso and Novo (2010) found that lack of evidence of psychological injury accounted for approximately 40% of acquittals on the grounds that the prosecution’s evidence of psychological injury did not constitute valid and sufficient proof. The MMPI-2 is the most extensively used and widely researched psychometric instrument for assessing adult psychopathology in forensic psychology, followed by the Unstructured Clinical Interviews; and the combination of both (Greene, 2008; Rogers et al., 2003). However, Unstructured Clinical Interviews do not contain any systematic procedures for the control of feigning, the differential diagnosis of which should always be considered prior to the diagnosis of PTSD, thus it fails to meet the Daubert Standard (Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993) which is a rule of evidence regarding the admissibility of expert witnesses’ testimony (i.e., the unstructured interview is not widely recognized by the scientific community as a valid instrument for forensic psychological assessment). Thus, as assessment of feigned symptoms is substantially informed by the MMPI-2, this study aimed to assess the validity of the standard forensic procedure followed by the Spanish forensic psychologists on this instrument in discriminating between adjudicated and mock victims of IPV.

Method

Participants

A total of 105 Caucasian Spanish women (age ranged from 19 to 73 years, $M = 33.56$, $SEM = 1.09$) participated in the study. Of these, 51 women (age ranged 19 to 64 years, $M = 37.61$, $SEM = 1.48$) were victims of IPV who had reported the offence, and had secured

a conviction for IPV against their aggressors. All of the cases involved both physical (it implies physical injury; none was classified by the prosecution as attempted homicide) and psychological violence, and 23 of these cases had been sentenced as continuous violence. The remaining 54 women (age ranged from 21 to 73 years, $M = 29.74$, $SEM = 1.41$) were living with a partner and had no history of IPV.

Design

A quasi-experimental research methodology for known-groups comparison (adjudicated vs. mock victims) (Rogers, 2008b) was used with archive data and data from the normal population. Psychological injury arising from alleged IPV was measured using a psychometric instrument that consisted of a symptoms recognition task. As for the design sensitivity analysis in a sample of 105 participants, the probability of detecting $(1 - \beta)$ significant differences ($\alpha < .05$) for a medium effect size for the different measures between two groups and for a MANOVA with 10 and 4 variables, was $> 80\%$.

Procedure

The evaluations of adjudicated victims were drawn from the archives of the Forensic Psychology Institute of the Universidad de Santiago de Compostela (Spain). The inclusion criteria for the adjudicated victim group (ground truth) were women who had reported the offence and secured a firm conviction against their aggressor; the accused had pleaded guilty i.e., had admitted the offence; and the burden of proof was beyond reasonable doubt i.e., the documented evidence, testimonies, violation of restraining orders, etc, had led to the aggressor's conviction. There was no evidence in judicial file for adjudicated victims of previous psychological distress. In line with the actual Spanish protocol for IPV charges, the complainant had passed by the Centros de Información a la Mujer (Women Information Centres), where they received judicial information, and personal and legal assistance. None of the psychological evaluations undertaken in this study were used as evidence in court. The women feigning allegations of IPV were living with their partners, had no previous history of IPV, and had negatively responded to a questionnaire on instances of IPV. Mock victims were contacted and assessed individually considering a sociodemographic (e.g., age, social status, number of sons) correspondence with the adjudicated victim group. Feigner women were informed about the purpose of the study and freely volunteered to participate in this study. Feigners received feigning instructions in line with the recommendations of Rogers (2008c) for implementing this type of design,

and to ensure the instructions were easily understandable as reported in previous studies on IPV (Arce et al., 2006, 2009). Each feigner was asked to imagine she had made false allegations of IPV and was going to be evaluated by a forensic psychologist. The making of false allegations was justified on the ground of obtaining benefits such as child custody, revenge, or financial compensation. Moreover, feigners were informed about the importance of the results of this study in detecting feigners (i.e., the indirect harm and suffering to children, and wrongful conviction). Feigners received no training in feigning, but were encouraged to self-train, be credible, and be fully committed to the task (if participants refused to comply with the instructions were asked to withdraw from the study, nevertheless all agreed to participate). Participants were given a week to plan their feigning strategies for the psychological evaluation. Finally, participants were debriefed by informing them how well they had performed the task (recall and comprehension of instructions), to determine their levels of task engagement and motivation, and to ensure participants had understood and completed the task correctly. The results confirmed task comprehension and engagement (e.g., around 75% of feigners simulate the primary and indirect psychological injury). Furthermore, the mock victim women's ability to discriminate between expected and unexpected symptoms consequence of IPV is general (e.g., none simulate a disorder not related to the psychological injury, the hypomania).

Furthermore, the MMPI-2 protocols were screened in search of highly inconsistent response profiles in the evaluations either due to extreme acquiescence (TRIN raw score > 18); random responses (VRIN raw score > 18 ; $F T \geq 120$; $|F - Fb| > 19$); a large number of unanswered items or double responses, which would indicate a lack of cooperation in the evaluations; and outliers (L raw score > 10 , K raw score > 26), in order to eliminate them from the study (Greene, 2008), but none of these contingencies were observed. All of the women freely volunteered to participate and informed consent was obtained.

Instruments

The psychometric measurement instrument used in this study was the Spanish version of the MMPI-2 (Hathaway & McKinley, 1999). As the aim of this study was to measure the mental health of the participants, the standard clinical scales were used (with the Masculinity-Femininity scale omitted because it would contravene the Article 14 of the Constitución Española¹

¹Constitución Española. (1978). *Boletín Oficial del Estado*, 311, 29313–29424. Retrieved from <http://www.boe.es/legislacion/enlaces/documentos/ConstitucionINGLES.pdf>

[Spanish Constitution], and the Ley Orgánica 3/2007, de 22 de marzo, para la Igualdad Efectiva de Mujeres y Hombres [Spanish Law for Effective Equality of Women and Men or Cohabiting Marriage Act]). In addition, the psychological injury of IPV as measured in terms of PTSD was assessed using the Pk scale (Keane, Malloy, & Fairbank, 1984). It should be borne in mind that the scale was designed for Vietnam War veterans, but has not been sufficiently contrasted in other contexts, and there are discrepancies regarding its ability to accurately classify in other contexts (Bowler, Hartney, & Ngo, 1998). Moreover, as the scale is related to distress, depressive or somatising patients score high on these scales (Lyons & Wheeler-Cox, 1999). Thus, the efficacy of the scale for the diagnosis of PTSD, in particular for cases of IPV, has not been firmly established (Friedman, Lewak, Nichols, & Webb, 2001).

The scales, control indexes, and decision-making criteria were selected from the reference literature for Spanish forensic psychologists i.e., those followed by Spanish forensic psychologists in their ordinary practice. As the criteria outlined in the different scales vary inter-studies, the manual of the Spanish version of the MMPI-2 (Hathaway & McKinley, 1999) was used. When no specific criteria for scales or indexes were available, from the different criteria outlined in the literature, the criteria in line with those informed by Jiménez and Sánchez's (2003) in their forensic psychology assessment manual (the authors of the Spanish version of the MMPI-2) were used.

The MMPI-2 has four standard validity scales for the control of the protocols: the Cannot Say Scale (?), the K Scale, the F Scale, and the L Scale. The Cannot Say score is the sum of the number of items left unanswered or with double true-false responses, which may be indicative of feigning i.e., the criterion of refusing to cooperate in the assessment (American Psychiatric Association, 2000). The protocols with 30 or more unanswered or double responses are invalid, though Graham (2006) has highlighted that this cut-off point is much too lax, and reduced the figure to 10. The K Scale measures the subject's attitude towards the assessment; a high score ($T \geq 65$) indicates defensiveness whereas a low score ($T < 40$) detects subjects who feign emotional disorders (Graham, 2006). The L Scale (Lie), was designed to evaluate faking good, a high score ($T \geq 65$) indicates faking good (Bagby, Nicholson, Buis, Radovanovic, & Fidler, 1999). The F Scale (Frequency) is related to feigning in that high scores ($T \geq 65$, Graham, 2006) indicate feigning. In addition, the combination of indexes of the standard validity scales which have been found to be effective in detecting the feigning of disorders were calculated (Friedman et al., 2001; Jiménez & Sánchez, 2003): the F-K index and the "inverted V" profile. The F-K

index was able to detect attempts to give a bad image of oneself, the cut-off point being the difference in raw scores > 12 (Rogers et al., 2003). As for the "inverted V" profile, T scores on L and K < 50 and on F > 80 are indicative of symptom exaggeration (Jiménez & Sánchez, 2003).

In addition, the efficacy and correspondence between the standard validity scales and other additional validity indicators related to the literature on feigning (Berry, Baer, & Harris, 1991; Graham, 2006; Greene, 2008; Rogers et al., 2003; Rogers, Sewell, & Salekin, 1994) were assessed using: the Back F Scale (Fb), Gough's Dissimulation Scale (Ds), and the Infrequent Psychopathology Scale (Fp) (Arbisi & Ben-Porath, 1995). Raw scores > 8 on the Fp Scale (Rogers et al., 2003); and ≥ 26 on the Ds Scale (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989); and a $T \geq 110$, since it is presumed to be a clinical context (Butcher et al., 2001), on the Fb Scale are indicative of feigning.

Finally, the scales measuring response consistency, the TRIN (True Response Inconsistency) and VRIN (Variable Response Inconsistency) scales, and $|F-Fb|$ (Greene, 2008), taking as an inconsistency cut-off point a raw score ≥ 13 on the VRIN and, TRIN, and for the $|F-Fb|$ a raw score above 6.77 ($T \geq 70$) using data from the normative population (Greene, 1997).

Data analysis

As for the mean comparisons between the mock and adjudicated victim groups, MANOVAs and ANOVAs were performed after the verification of the assumptions of variance homogeneity (Levene test), normality (Kolmogorov-Smirnov test), and independence (see the design).

Results

Analysis of the standard validity control scales of the MMPI-2

A multivariate analysis of the response validity control scales showed they were sensitive to the population factor (adjudicated vs. mock victims), $F(4, 100) = 8.07$, $p < .001$, $\eta^2 = .244$, $1-\beta = .998$. This factor explained 24.4% of the variance, a large effect size.

Univariate analysis (see Table 1) showed the number of unanswered responses was significantly higher in adjudicated victims, and significantly higher scores on the K Scale and F Scale in mock victims. Thus, the results reveal the Cannot Say, the F Scale and K Scale validity may be useful for detecting feigning in forensic settings.

The study of cases showed the Cannot Say Scale (raw score ≥ 10) classified all of the protocols of both adjudicated and mock victims as valid. In the

Table 1. ANOVAs on MMPI-2 standard validity scales by the population factor

| Scale | <i>F</i> | <i>p</i> | η^2 | $M_{\text{adjudicated}}$ | M_{mock} | 1- β |
|----------------------|----------|----------|----------|--------------------------|-------------------|------------|
| Cannot Say Scale (?) | 8.80 | .004 | .079 | 1.33 | 0.11 | .836 |
| Lie Scale (L) | 0.41 | .536 | .004 | 56.24 | 62.41 | .097 |
| Frequency Scale (F) | 26.91 | .001 | .207 | 63.59 | 83.52 | .999 |
| K Scale | 5.57 | .020 | .051 | 46.16 | 42.00 | .647 |

Note: $df(1, 103)$. $M_{\text{adjudicated}}$ = Mean of the adjudicated victim group; M_{mock} = Mean of the mock victim group.

classification of L ($T \geq 65$), 9 adjudicated victims (17.6%) were erroneously classified as good feigners; on the K Scale ($T < 40$), 10 adjudicated victims (19.6%) were erroneously classified as feigners, and on the F Scale ($T \geq 65$) 22 adjudicated victims (43.1%) were erroneously classified as feigners. In comparison, the L Scale erroneously classified as 8 mock victims (14.8%) as good feigners; the K Scale correctly classified 28 feigners (51.9%); and on the F Scale correctly classified 37 mock victims as feigners (68.5%). Thus, the L Scale, $\chi^2(1) = 0.02$, *ns*, $\phi = .038$, failed to discriminate between adjudicated and mock victims, the K Scale, $\chi^2(1) = 10.45$, $p < .001$, $\phi = .335$, and F Scale, $\chi^2(1) = 5.87$, $p < .05$, $\phi = .256$, discriminated between adjudicated and mock victims i.e., both scales were significant predictors of feigning. Moreover, the results corroborate that potential feigners collaborated with the assessment.

Indexes and configurations of the standard validity scales

Significant differences were observed between adjudicated ($M = 0.63$) and mock victims ($M = 11.15$) in the *F-K index*, $F(1, 432) = 24.02$, $p < .001$, $\eta^2 = .189$, $1-\beta = .998$. That is, the *F-K index* was higher for mock than for adjudicated victims. The study of cases ($F-K > 12$) found this occurred in 23 mock (42.6%), and in 4 adjudicated victims (7.8%), revealing a significant, $\chi^2(1) = 14.81$, $p < .001$, $\phi = -.397$, net feigning rate (true positives – false positives = 34.8%). The *inverted V profile* (L and K < 50 and F < 80) detected 12 mock victims (22.2%) as feigners and 2 adjudicated victims (3.9%) i.e., a significant net feigning rate (18.3%), discriminating between adjudicated and mock victims, $\chi^2(1) = 6.01$, $p < .05$, $\phi = -.269$.

Analysis of the additional validity scales of the MMPI-2

The Fb Scale showed significant differences mediated by the population factor (Table 2) i.e., Fb scores were higher for mock than for adjudicated victims. The study of cases ($T \geq 110$) detected 10 mock victims as feigners (18.5%), but none of the adjudicated victims. That is, the Fb Scale correctly classified 18.5% of feigners without any false positives (adjudicated victims identified as feigners).

The Fp Scale was sensitive to the population factor (Table 2), and the Fp scores were significantly greater for feigners. The study of cases (raw score > 8) revealed the correct classification of 35.2% mock victims as feigners ($n = 19$), with 2% ($n = 1$) of false positives (adjudicated victims informed as feigners), a net feigning rate (true positives – false positives) of 33.2%, $\chi^2(1) = 16.68$, $p < .001$, $\phi = -.423$.

The population factor exhibited significant differences (Table 2) on the Gough Dissimulation Scale (Ds). On the Ds Scale the scores were higher for mock than for adjudicated victims. Likewise, the study of cases (raw score ≥ 26) showed the Ds Scale classified as true positives (correct identification of feigners) 74.1% ($n = 40$) of mock victims, with 27.5% ($n = 14$) of false positives, a net feigning rate of 49.3%, $\chi^2(1) = 21$, $p < .001$, $\phi = -.466$.

Inconsistency measurement

The TRIN and VRIN validity scales, and the F-Fb index (Table 3) were not sensitive to the population factor, nor did they invalidate any of the protocols of adjudicated or mock victims i.e., adjudicated and feigned protocols were consistent.

Table 2. ANOVAs on the MMPI-2 additional validity scales by the population factor

| Scale | <i>F</i> | <i>p</i> | η^2 | $M_{\text{adjudicated}}$ | M_{mock} | 1- β |
|----------------------------------|----------|----------|----------|--------------------------|-------------------|------------|
| Back F (Fb) | 33.14 | .001 | .243 | 63.61 | 86.78 | 1 |
| Infrequency psychopathology (Fp) | 26.81 | .001 | .207 | 3.53 | 7.15 | .999 |
| Gough Dissimulation (Ds) | 51.22 | .001 | .332 | 20.16 | 33.04 | 1 |

Note: $df(1, 103)$. $M_{\text{adjudicated}}$ = Mean of the adjudicated victim group; M_{mock} = Mean of the mock victim group.

Table 3. ANOVAs on the inconsistency scales by the population factor

| Scale | F | p | η^2 | $M_{\text{adjudicated}}$ | M_{mock} | 1- β |
|-------|------|------|----------|--------------------------|-------------------|------------|
| TRIN | 0.01 | .913 | .000 | 9.65 | 9.61 | .051 |
| VRIN | 1.62 | .206 | .015 | 7.94 | 8.55 | .243 |
| F-Fb | 0.40 | .527 | .004 | 4.33 | 3.12 | .096 |

Note: $df(1, 103)$. $M_{\text{adjudicated}}$ = Mean of the adjudicated victim group; M_{mock} = Mean of the mock victim group.

Global analysis of the detection of feigning using the validity indicators of the MMPI-2

The consistency of the results is a prerequisite in Spanish judicial context (e.g., the Judgement of the Supreme Court (Sentencia del TS de 8 de febrero de 1995, RJ 808; sentencia del TS de 18 de noviembre, RJ 7987), as is inter-measure consistency a scientific requisite attesting to the reliability of measure. Thus, for a profile to be invalidated on the grounds of feigning, the protocols should have invalidating scores, the *minimum-decision criterion*, in at least two standard validity scales relevant to feigning i.e., the F and K, and their combinations (the *F-K index* and the *inverted V profile*) or the additional (Fb, Fp and Ds) scales. Due to this and other decision criteria the data/evidence must be analysed to confirm if two invalidating scores discriminate effectively between adjudicated victims and feigners. Thus, two models may be considered for this purpose, one based on standard scales and indexes, the Standard Forensic Model used by Spanish forensic psychologists given that the other scales are not available in the Spanish commercial version that consists of the F, K, Fb, F-K, and the inverted V scales and indexes. A second model referred to as the Extended Model consists of all the scales and indexes that significantly discriminate between adjudicated and mock victim protocols (i.e., F, K, Fb, F-K, Fp, Ds, and the inverted V profile).

The results revealed the presence of higher feigning indicators among mock than adjudicated victims in the Model Forensic (38 vs. 110 indicators of feigning for adjudicated and mock victims, respectively), $\chi^2(1) = 35.03$, $p < .001$, and for the Extended Model (53 vs. 169), $\chi^2(1) = 60.61$, $p < .001$. In other words, feigning indicators on the MMPI-2 statistically discriminated between adjudicated and mock victims. Notwithstanding, this result is not valid for forensic contexts as it does not provide a decision criterion for decision-making for each specific case (the forensic task). The decision criterion in forensic settings must be strict: false negatives are not admissible i.e., the forensic psychologist cannot classify a feigned protocol as honest given the implications this would have for the accused (i.e., in dubio pro-reo or the principle of reasonable doubt). Likewise, the DSM-IV (American Psychiatric Association, 2000) requires in forensic settings a differential diagnosis of feigning prior to the diagnosis of PTSD. This strict criterion is defined by analyzing the accumulated effects of feigning indicators. The results of these accumulated effects in the Forensic Model and Extended Model (Table 4) showed that the former required more than 4 criteria to discriminate between the protocols of adjudicated and mock victims, and the latter more than 5 criteria. The application of this decision criterion relies on the correct identification of 9.3% and 18.5% (for both the Forensic and Extended Model, respectively) of true positives (mock victims informed as feigners), and 100% of true negatives (adjudicated victims identified as such), and false positives (adjudicated victims identified as feigners) in both models, but the classification would be incorrect in 90.7% and 81.5% of feigners classified as adjudicated victims (false negatives). Finally, if the minimum-decision criterion were applied i.e., two invalidating criteria for the feigning protocols, it correctly classified 55.7% and 72.3% of true positives (for the Forensic and Extended models, respectively), 82.4% and 70.5% of

Table 4. Feigning indexes by population and model. Accumulative effects

| No. of indexes/population | Forensic model | | Extended model | |
|---------------------------|----------------|-----------|----------------|----------|
| | Adjudicated | Mock | Adjudicated | Mock |
| 0 | 28(54.9%) | 10(18.5%) | 27(52.9%) | 8(14.8%) |
| 1 | 14(27.5%) | 14(25.9%) | 9(17.6%) | 7(13%) |
| 2 | 5(9.8%) | 11(20.4%) | 6(11.8%) | 9(16.7%) |
| 3 | 2(3.9%) | 7(13%) | 6(11.8%) | 7(13%) |
| 4 | 2(3.9%) | 7(13%) | 1(2%) | 6(11.1%) |
| 5 | | 5(9.3%) | 2(3.9%) | 7(13%) |
| 6 | | | — | 6(11.1%) |
| 7 | | | — | 4(7.4%) |

true negatives, 17.6 and 29.5% of false positives, and 44.4 and 27.8% of false negatives. As forensic contexts cannot admit false negatives (a feigner identified as an adjudicated victim), none of the decision criteria obtained from the control scales and indexes on the MMPI-2 would be valid and sufficient for the application in forensic context.

Comparison of the mental health of adjudicated and mock victims of IPV on the standard clinical scales of the MMPI-2

Multivariate differences were found in the standard clinical scales of the MMPI-2 mediated by the population factor (adjudicated vs. mock victims), $F(9, 95) = 5.52$, $p < .001$, $\eta^2 = .343$, $1-\beta = 1$. Moreover, the effect size attributed to the population was high, explaining 34.3% of the variance.

The univariate effects (Table 5) show significant differences in the 9 standard clinical scales of the MMPI-2. Succinctly, feigners reported more clinical disorders: hypochondriasis, depression, hysteria, psychopathic disorders, psychasthenia, paranoia, schizophrenia, hypomania, and social introversion. As the effects went in the expected direction i.e., higher scores for feigners, the results corroborate the general ability to feign. This ability is also an effective indirect measure par excellence of psychological injury in cases of IPV, depression and social introversion (i.e., Bargai et al., 2007; Kessler et al., 1995). Furthermore, the results showed that feigners informed of other pathologies in dimensions unrelated to IPV (e.g., psychopathic disorders, hysteria, hypochondriasis, paranoia, schizophrenia, psychasthenia), within the region for pathology ($T > 70$) whereas in hypomania it did not enter this region. Hence, they fail to perform poorly on pathologies that are not directly associated to psychological injury in cases of IPV. In other words, feigners fail to effectively discriminate between expected depressive symptoms and social introversion and unexpected

psychotic ones, psychopathic disorders, hysteria, hypochondriasis, and hypomania. Thus, feigners follow two strategies: indiscriminant symptom endorsement (reporting a large number of psychopathology than adjudicated victims), and symptom severity or exaggeration of symptom (reporting more symptoms than adjudicated victims).

One of the salient errors of feigners is the self-reporting of symptoms rarely exhibited by adjudicated victims of IPV (Greene, 2008). The cases studies revealed that schizophrenia, which in non-psychiatric populations is related to social alienation, was the most common psychopathology ($T > 70$) unrelated to psychological injury feigned by mock victims, 75.9% as a pathology, and 46.3% as the primary disorder; whilst in adjudicated cases of IPV it was 19.6%, but none as a primary disorder. Thus, the appearance of schizophrenia as a primary disorder is a potential indicator of feigning. It should also be noted that the testimonies of adjudicated victims of IPV suffering from schizophrenia are not admitted as evidence as this disorder affects testimonial credibility (Arce, Seijo, & Novo, 2010).

As for the PK scale measuring PTSD, the primary disorder for the assessment of psychological injury arising from IPV (Bargai et al., 2007; Kessler et al., 1995; Vilariño et al., 2009), a significant modulating effect of the population factor, $F(1, 432) = 34.49$, $p < .001$, $\eta^2 = .245$, $1-\beta = 1$, was observed. Thus, feigners ($M = 76.20$) inform more PTSD symptoms than adjudicated victim ($M = 63.28$).

The study of cases ($T > 70$) of indirect psychological injury from IPV revealed a depressive disorder in 40 feigners (74.1%), and 25 adjudicated victims (49%) of IPV, $\chi^2(1) = 5.96$, $p < .05$, $\phi = -.258$, and Social Introversion in 35 feigners (64.8%), and 11 adjudicated victims (21.6%) of IPV, $\chi^2(1) = 18.21$, $p < .001$, $\phi = -.436$. The measure of primary psychological injury on the Pk scale showed a higher rate of PTSD among feigners

Table 5. Univariate effects on the MMPI-2 standard clinical scales by the population factor

| Scale | <i>F</i> | <i>p</i> | η^2 | <i>M</i> _{adjudicated} | <i>M</i> _{mock} | 1- β |
|----------------------|----------|----------|----------|---------------------------------|--------------------------|------------|
| Hypochondriasis | 10.73 | .001 | .094 | 66.43 | 76.63 | .901 |
| Depression | 8.98 | .003 | .080 | 68.43 | 76.74 | .843 |
| Hysteria | 7.28 | .008 | .066 | 63.51 | 71.18 | .762 |
| Psychopathic deviate | 9.97 | .002 | .088 | 62.88 | 70.39 | .878 |
| Psychasthenia | 23.64 | .001 | .187 | 62.84 | 74.11 | .838 |
| Paranoia | 8.84 | .004 | .079 | 69.74 | 78.31 | .998 |
| Schizophrenia | 30.42 | .001 | .228 | 65.80 | 82.70 | 1 |
| Hypomania | 6.99 | .009 | .064 | 54.57 | 59.74 | .746 |
| Social introversion | 26.76 | .001 | .206 | 60.73 | 72.87 | .999 |

Note: $df(1, 103)$. *M*_{adjudicated} = Mean of the adjudicated victim group; *M*_{mock} = Mean of the mock victim group.

($n = 41$, 75.9%) than among adjudicated victims ($n = 28$, 54.9%) of IPV, $\chi^2(1) = 4.83$, $p < .05$, $\phi = -.221$. Thus feigners were not only able to feign both primary and secondary psychological injury of IPV, but also the probability of registering psychological injury was greater among mock than adjudicated victims.

Discussion

The following limitations of this study should be borne in mind. First, it is assumed that all victims are adjudicated cases of IPV as confirmed by a court of law, but this guarantee is not absolute. Second, it is assumed feigning under real conditions parallels feigning in experimental high fidelity conditions, but the motivations of the actors do not overlap (Fariña, Arce, & Real, 1994; Konecni & Ebbesen, 1992). Third, the results should not be directly generalized to other populations and case types other than IPV. Fourth, since the mock victims were not drawn from archives, the legal process cannot be entirely replicated with subsequent limitations in the ecological validity of the design. Fifth, as results are limited to a psychometric measure, may not be generalized to the whole forensic task that requires a multi-method approach (e.g., interview, file study). Sixth, there is no a foolproof method to detect feigning. Seventh, as the cases come from psychological records, it is not available reliable information in all files about the physical expert evidence that may influence on psychological injury. Taking into account these observations, the following conclusions for forensic applications may be reached:

- a) Contrary to predictions of the criminological model of the DSM-IV (American Psychiatric Association, 2000), feigners of IPV do collaborate in the assessment (? < 10). Neither did adjudicated IPV victims refuse to cooperate in the assessment (? < 10), but the judicial archive data from The Spanish Observatory against Domestic and Gender Violence showed that in 11.6% of cases the victim withdrew its allegations (Observatorio contra la Violencia Doméstica y de Género, 2011). Thus, the higher scores observed among adjudicated victims on the Cannot Say Scale should be considered in relation to alternative hypothesis to the lack of cooperation with the evaluation, as confusion or indecision. In any case, it may not be related with malingering.
- b) The psychopathology of feigners on the MMPI-2 revealed psychological injury of IPV i.e., the general population of women has the sufficient ability to feign psychological injury in IPV, the figure being estimated at around 75%. Feigners used two strategies: indiscriminant symptom endorsement, and symptoms severity i.e., reporting a large number

and high intensity of symptoms and psychopathology. These results have three implications for forensic psychology. First, the utility of the discriminant or divergent validity for the assessment of feigning i.e., feigners are expected to exhibit psychopathology not related to IPV. Second, the indiscriminate endorsement of symptoms means feigners report rare symptoms that are rarely reported by genuine patients (in our case adjudicated victims of IPV). The control of quasi-rare symptoms is assessed by the F, Fb scales and the F-K index, and rare symptoms by the Fp Scale that have proven to be effective in detecting this strategy. Third, feigners assume the erroneous stereotypes that genuine victims suffer from severe symptoms as detected by the Ds Scale.

- c) No inconsistency in the responses was observed in the protocols of both adjudicated and mock victims.
- d) Given that in forensic contexts a strict decision criterion is a requisite, false negatives are not admissible, the psychometric models for detecting feigning based on the validity scales, indexes, and configurations of the MMPI-2 are not valid for forensic applications since the decision criterion can be obtained without incurring in false negatives. The deficiency of both psychometric models is due to two main factors. First, in line with the findings of Greene (2008) in patients with mental disorders, and the meta-analysis of Rogers et al. (2003), the scales, indexes and profiles of the MMPI-2 also informed adjudicated victim as feigners i.e., false positives. Secondly, the validity controls of the MMPI-2 may fail to detect all feigners.
- e) Given that Spanish forensic psychologist uses a forensic psychometric model and the Unstructured Clinical Interview, they fail to discriminate between adjudicated victims and feigners of IPV without false negatives i.e., they fail to adequately fulfill the forensic task.

In conclusion, the standard Spanish forensic procedure for the evaluation of psychological injury in cases of IPV did not constitute valid proof for judges who acquitted defendants on the grounds of not proven owing to the lack of evidence of psychological injury.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th Ed.). Washington, DC: Author.
- Arbisi P. A., & Ben-Porath Y. S. (1995). On the MMPI-2 infrequent response scale for use with psychopathological populations: The Infrequency Psychopathology Scale F(p). *Psychological Assessment*, 7, 424–431. <http://dx.doi.org/10.1037//1040-3590.7.4.424>

- Arce R., Alonso M. A., & Novo M. (2010). Estudio de sentencias de violencia de género con menores implicados: Carga de la prueba y decisiones sobre los menores [A study of judicial judgments on gender violence with children involved: The standard of proof and decision on children]. In F. Fariña, R. Arce, M. Novo, & D. Seijo (Eds.), *Separación y divorcio: Interferencias parentales* [Separation and divorce: Parental interference] (pp. 241–252). Santiago de Compostela, Spain: Nino.
- Arce R., Fariña F., & Buela-Casal G. (2008). Assessing and detecting the ability to faking psychological injury as a consequence of a motor vehicle accident on the MMPI-2 using mock victims. *Revista Latinoamericana de Psicología*, 40, 485–496.
- Arce R., Fariña F., Carballal A., & Novo M. (2006). Evaluación del daño moral en accidentes de tráfico: Desarrollo y validación de un protocolo para la detección de la simulación [Evaluating psychological injury in motor vehicle accidents (MVA): Development and validation of a protocol for detecting simulation]. *Psicothema*, 18, 278–283.
- Arce R., Fariña F., Carballal A., & Novo M. (2009). Creación y validación de un protocolo de evaluación forense de las secuelas psicológicas de la violencia de género [Creation and validation of a forensic protocol to assess psychological harm in battered women]. *Psicothema*, 21, 241–247.
- Arce R., Seijo A., & Novo M. (2010). Testimony validity: A comparative study of legal and empirical criteria. *Psychology in Spain*, 14, 1–7.
- Arch M., Jarne J., Peró M., & Guàrdia J. (2011). Child custody assessment: A field survey of Spanish forensic psychologists' practices. *The European Journal of Psychology Applied to Legal Context*, 3, 107–128.
- Archer R. P., Buffington-Vollum J. K., Stredny R. V., & Handel R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment*, 87, 84–94. http://dx.doi.org/10.1207/s15327752jpa8701_07
- Bagby R. M., Nicholson R. A., Buis T. Radovanovic H., & Fidler B. J. (1999). Defensive responding on the MMPI-2 in family custody and access evaluations. *Psychological Assessment*, 11, 24–28. <http://dx.doi.org/10.1037/1040-3590.11.1.24>
- Bargai N., Ben-Shakhar G., & Shalev A. Y. (2007). Posttraumatic stress disorder and depression in battered women: The mediating role of learned helplessness. *Journal of Family Violence*, 22, 267–275. <http://dx.doi.org/10.1007/s10896-007-9078-y>
- Berry D. T. R., Baer R. A., & Harris M. J. (1991). Detection of malingering on the MMPI: A meta-analysis. *Clinical Psychology Review*, 11, 585–598. [http://dx.doi.org/10.1016/0272-7358\(91\)90005-F](http://dx.doi.org/10.1016/0272-7358(91)90005-F)
- Bow J. N., & Quinnell F. A. (2001). Psychologists' current practices and procedures in child custody evaluations: Five years after American Psychological Association Guidelines. *Professional Psychology: Research and Practice*, 32, 261–268. <http://dx.doi.org/10.1037/0735-7028.32.3.261>
- Bowler R. M., Hartney C., & Ngo L. H. (1998). Amnesic disturbance and posttraumatic stress disorder in the aftermath of a chemical release. *Archives of Clinical Neuropsychology*, 13, 455–471. <http://dx.doi.org/10.1093/arclin/13.5.455>
- Brown T. A., Campbell L. A., Lehman C. L., Grisham J. R., & Mancill R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology*, 110, 585–599. <http://dx.doi.org/10.1037/0021-843X.110.4.585>
- Butcher J. N., Dahlstrom W. G., Graham J. R., Tellegen A., & Kaemmer B. (1989). *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., Dahlstrom W. G., & Kaemmer B. (2001). *MMPI-2 (Minnesota Multiphasic Personality Inventory-2): Manual for administration, scoring, and interpretation, revised edition*. Minneapolis, MN: University of Minnesota Press.
- Daubert v. Merrell Dow Pharmaceuticals, 113 S. Ct. 2786 125 L.E.2d 469 (1993).
- Fariña F., Arce R., & Real S. (1994). Ruedas de identificación: De la simulación y la realidad [Line-ups: A comparison of high fidelity research and research in a real context]. *Psicothema*, 7, 395–402.
- Fariña F., Arce R., & Sotelo A. (2010). Es efectivo el estudio psicométrico estándar del peritaje del estado clínico y de la disimulación en progenitores en litigio por la guarda y custodia de menores? [Is effective the standard psychometric forensic evaluation of the mental health and faking good of the partners litigating by the child custody?]. *Revista Iberoamericana de Psicología y Salud*, 1, 65–79.
- Friedman A. F., Lewak R., Nichols D. S., & Webb J. T. (2001). *Psychological assessment with the MMPI-2*. Mahwah, NJ: LEA.
- Graham J. R. (2006). *MMPI-2: Assessing personality and psychopathology* (4th Ed.). New York, NY: Oxford University Press.
- Greene R. L. (1997). Assessment of malingering and defensiveness by multiscale inventories. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 169–207). New York, NY: The Guilford Press.
- Greene R. L. (2008). Malingering and defensiveness on the MMPI-2. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd Ed., pp. 159–181). New York, NY: The Guilford Press.
- Hathaway S. R., & McKinley J. C. (1999). *Inventario Multifásico de Personalidad de Minnesota-2. Manual*. [The Minnesota Multiphasic Personality Inventory-2. Manual]. Madrid, Spain: TEA Ediciones.
- Jiménez F., & Sánchez G. (2003). *Evaluación psicológica forense. Contribución de las técnicas de Minnesota y Millon* [Forensic psychological evaluation. Contributions of the Minnesota and Millon's techniques]. Salamanca, Spain: Amarú Ediciones.
- Jiménez F., Sánchez G., & Tobón C. (2009). A social desirability scale for the MMPI-2. Which of the two: Wiggins (WSD) or Edwards (ESD)? *The European Journal of Psychology Applied to Legal Context*, 1, 147–163.
- Keane T. M., Malloy P. F., & Fairbank J. A. (1984). Empirical development of an MMPI subscale for the assessment of combat-related posttraumatic stress disorder. *Journal of*

- Consulting and Clinical Psychology*, 52, 888–891. <http://dx.doi.org/10.1037/0022-006X.52.5.888>
- Kessler R. C., Sonnega A., Hughes M., & Nelson C. B.** (1995). Posttraumatic stress disorder in the national comorbidity survey. *Archives of General Psychiatry*, 52, 1048–1060. <http://dx.doi.org/10.1001/archpsyc.1995.03950240066012>
- Konecni V. J., & Ebbesen E. B.** (1992). Methodological issues on legal decision-making, with special reference to experimental simulations. In F. Lösel, D. Bender, & T. Bliesener (Eds.), *Psychology and law. International perspectives* (pp. 413–423). Berlin, Germany: Walter de Gruyter.
- Ley Orgánica 3/2007, de 22 de marzo, para la Igualdad Efectiva de Mujeres y Hombres. (2007). *Boletín Oficial del Estado*, 71, 12611–12645. Retrieved from <http://www.boe.es/boe/dias/2007/03/23/pdfs/A12611-12645.pdf>
- Lyons J. A., & Wheeler-Cox T.** (1999). MMPI, MMPI-2, and PTSD: Overview of scores, scales, and profiles. *Journal of Traumatic Stress*, 12, 175–183. <http://dx.doi.org/10.1023/A:1024710803042>
- Observatorio contra la Violencia Doméstica y de Género.** (2011). *Datos de denuncias, procedimientos penales y civiles registrados, órdenes de protección solicitadas en los Juzgados de Violencia sobre la Mujer (JVM) y sentencias dictadas por los órganos jurisdiccionales en esta materia en el año 2011* [Data of complaints registered in criminal and civil proceedings, protective orders ordered by the of Violence against Women Courts (JVM) and judgments of the courts in this area in 2011]. Madrid, Spain: Author. Retrieved from http://www.poderjudicial.es/cgpj/es/Temas/Estadistica_Judicial/Informes_estadisticos/Informes_periodicos/La_violencia_sobre_la_mujer_en_la_estadistica_judicial_Datos_anuales_de_2011
- O'Donnell M. L., Creamer M., Bryant R. A., Schnyder U., & Shalev A.** (2006). Posttraumatic disorders following injury: Assessment and other methodological considerations. In G. Young, A. W. Kane, & K. Nicholson (Eds.), *Psychological knowledge in courts. PTSD, pain and TBI* (pp. 70–84). New York, NY: Springer.
- Pérez-Pareja J., Sesé A., González-Ordi H., & Palmer A.** (2010). Fibromyalgia and chronic pain: Are there discriminating patterns by using the Minnesota Multiphasic Personality Inventory-2 (MMPI-2)? *International Journal of Clinical and Health Psychology*, 10, 41–56.
- Resnick P. J., West S., & Payne J. W.** (2008). Malingering of posttraumatic disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd Ed., pp. 109–127). New York, NY: The Guilford Press.
- Rogers R.** (2008a). Structured interview and dissimulation. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd Ed., pp. 301–322). New York, NY: The Guilford Press.
- Rogers R.** (2008b). An introduction to response styles. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd Ed., pp. 3–13). New York, NY: The Guilford Press.
- Rogers R.** (2008c). Researching response styles. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd Ed., pp. 411–434). New York, NY: The Guilford Press.
- Rogers R., Bagby R. M., & Dickens S. E.** (1992). *Structured Interview of Reported Symptoms (SIRS) and professional manual*. Odessa, FL: Psychological Assessment Resources.
- Rogers R., Sewell K. W., Martin M. A., & Vitacco M. J.** (2003). Detection of feigned mental disorders: A meta-analysis of the MMPI-2 and malingering. *Assessment*, 10, 160–177. <http://dx.doi.org/10.1177/1073191103010002007>
- Rogers R., Sewell K. W., & Salekin R. T.** (1994). A meta-analysis of malingering on the MMPI-2. *Assessment*, 1, 227–237.
- Rogers R., & Shuman D. W.** (2005). *Fundamentals of forensic practice: Mental health and criminal law*. New York, NY: Springer.
- Spitzer R. L., Williams J. B., Gibbon M., & First M. B.** (1995). *Structured clinical interview for DSM-IV*. Washington, DC: American Psychiatric Press.
- United Nations.** (1988). *Committee on crime prevention and control. Report on the tenth session*. Vienna, Switzerland: Author.
- Vilariño M., Arce R., & Fariña F.** (2013). Forensic-clinical interview: Reliability and validity for the evaluation of psychological injury. *The European Journal of Psychology Applied to Legal Context*, 5, 1–21.
- Vilariño M., Fariña F., & Arce R.** (2009). Discriminating real victims from feigners of psychological injury in gender violence: Validating a protocol for forensic settings. *The European Journal of Psychology Applied to Legal Context*, 1, 221–243.