FORENSIC-CLINICAL INTERVIEW: RELIABILITY AND VALIDITY FOR THE EVALUATION OF PSYCHOLOGICAL INJURY

Manuel Vilariño*, Ramón Arce*, and Francisca Fariña**

*Department of Social Psychology, University of Santiago de Compostela (Spain)
**AIPSE Department, University of Vigo (Spain)

Abstract	Resumen
Forensic evaluation of psychological injury involves the use of a multimethod approximation i.e., a psychometric instrument, normally the MMPI-2, and a clinical interview. In terms of the clinical interview, the traditional clinical interview (e.g., SCID) is not valid for forensic settings as it does not fulfil the triple objective of forensic evaluation: diagnosis of psychological injury in terms of Post Traumatic Stress Disorder (PTSD), a differential diagnosis of feigning, and establishing a causal relationship between allegations of intimate partner violence (IPV) and psychological injury. To meet this requirement, Arce and Fariña (2001) created the forensic-clinical interview based on two techniques that do not contaminate the contents i.e., reinstating the contexts and free recall, and a methodic categorical system of contents analysis for the diagnosis of psychological injury and a differential diagnosis of feigning. The reliability and validity of the forensic-clinical interview designed for the forensic evaluation of psychological injury was assessed in 51 genuine cases of (IPV) and 54 mock victims of IPV who were evaluated using a forensic-clinical interview and the MMPI-2. The result revealed that the forensic-clinical interview was a reliable instrument (α = .85 for diagnostic criteria of psychological injury, and α = .744 for feigning strategies). Moreover, the results corroborated the predictive validity (the diagnosis of PTSD was similar to the expected rate); the convergence validity (the diagnosis of PTSD in the interview strongly correlated with the Pk Scale of the MMPI-2), and discriminant validity (the diagnosis of PTSD in the interview did not correlate with the Pk Scale in feigners). The feigning strategies (differential diagnosis) also showed convergent validity (high correlation with the Scales and indices of the MMPI2 for the measure of feigning) and discriminant validity (no genuine victim was classified as a feigner). Notwithstanding, feigning strategies failed to correctly classify all of the feigners indicating they must be complemented with other measures (multimethod approximation) to meet the requirements of forensic settings.	La evaluación forense del daño psicológico implica la utilización de una aproximación multimétodo: instrumentación psicométrica, generalmente el MMPI-2, y una entrevista clínica. Como entrevista clínica, la entrevista clínica tradicional (p.e., la SCID) no es válida para el campo forense ya que no cumple con el triple objetivo de la evaluación forense: diagnosticar el daño psicológico (Trastorno de Estrés Postraumático, TEP), un diagnóstico diferencial de simulación, y establecer una relación causa-efecto entre los hechos denunciados y el daño. Para este propósito Arce y Fariña (2001) crearon la entrevista clínico forense basada en dos técnicas que no contaminan los contenidos: la reinstauración de contextos y el recuerdo libre, y un sistema categorial metódico de análisis de contenido para el diagnóstico del daño psicológico y diferencial de simulación. Se diseñó un estudio con el objeto de contrastar la fiabilidad y validez de la entrevista clínico forense en la evaluación forense del daño psicológico en casos de violencia contra la mujer. 51 víctimas reales de violencia de género y 54 simuladoras fueron sometidas a la entrevista clínico forense y al MMPI-2. Los resultados mostraron que la entrevista clínico forense es un instrumento fiable (α = .85 para los criterios diagnósticos del daño psicológico, α = .744 para la estrategias de simulación). Asimismo, los resultados avalaron la validez predictiva (el diagnóstico del TEP fue igual al esperado); convergente (el diagnóstico de TEP en la entrevista correlacionaba altamente con la Escala Pk del MMPI-2) y discriminante (el diagnóstico de TEP en la entrevista no correlacionaba con la Escala Pk entre las simuladoras). Por su parte, las estrategias de simulación (diagnóstico diferencial) también se mostraron validez convergente (correlación elevada con las escalas e índices del MMPI2 de medida de la simulación) y discriminante (ninguna víctima real fue informada como simuladora). No obstante, falló al clasificar correctamente a todos los simuladores por lo que no es prueba forense suficiente habiendo de complementarse con otras medidas (aproximación mutimétodo).

Keywords: forensic evaluation; psychological injury; simulation; clinical interview; MMPI-2; multimethod approach.

Palabras clave: evaluación forense; daño psicológico; simulación; entrevista clínica; MMPI-2; aproximación multimétodo.

Correspondence: Manuel Vilariño. Facultad de Psicología. Campus Vida, s/n. 15782 Santiago de Compostela (Spain). E-mail: manuel.vilarino@usc.es

Introduction

According to a review of 600 judicial judgements of cases of intimate partner violence (IPV) in Spain, Arce, Alonso, & Vilariño (2010), invalid or insufficient evidence of psychological injury accounted for approximately 42% of acquittals. This finding underscores the need for raising the standard of proof of psychological injury in cases of IPV, and in other violent offences where evidence of psychological injury may be a statutory requisite (e.g., psychological violence, assault, economic loss or substantial impairment, sexual harassment or assault, threatening behaviour, duress, forcible restriction of individual freedom or privacy, kidnapping).

Evidence of psychological injury is mandatory if the prosecution is to secure a conviction for an offence involving psychological violence. It is axiomatic that no injury implies no victim, and by definition an offence of IPV must have a victim. Similarly, physical or sexual IPV may inflict psychological injury. Thus, proof of psychological injury is vital prosecution evidence underpinning a conviction in case of IPV. As for gender violence, a victim is defined as a woman who has suffered harm, including physical or mental injury, emotional suffering, economic loss or substantial impairment of their fundamental rights, through acts or omissions that are in violation of criminal laws (United Nations, 1988). Psychological injury inflicted by a criminal offence is determined by the forensic evaluation of the victim´s mental and emotional health. In legal contexts, it is mandatory for the prosecution to establish an unequivocal causal relationship between the offence and the injury. The forensic psychologist´s attempts to fulfilling this statutory requirement are often hindered in cases of IPV since victims are emotionally fraught with circumstances that have an adverse impact on their mental health or emotional wellbeing (e.g., breaking up with a partner, family breakdown, serious financial hardship, life disruption, fear of an uncertain future, loss of control, social alienation). Hence, the task of the forensic psychologist is twofold: to evaluate psychological injury, and to establish a causal relationship between observable injury and allegations of IPV. The Posttraumatic Stress Disorder (PTSD) criteria stipulated in the ICD-10 (World Health Organization, 1992), and the DSM-IV (American Psychiatric Association, 2000) are extensively used for diagnosing psychological injury, particularly in cases of IPV (e.g., Bargai, Ben-Shakhar, & Shalev, 2007; Kessler, Sonnega, Hughes, & Nelson, 1995; Sarasua, Zubizarreta, Echeburúa, & Corral, 2007; Vilariño, Fariña, & Arce, 2009). Nevertheless, forensic and clinical evaluation based on psychometric instruments and the standard clinical interview

do not meet the requirements of forensic evaluation as they fail to establish a causal relationship i.e., each specific PTSD criterion must be linked to IPV in order to rule out other concurring causes. The prevalence of PTSD arising from IPV fluctuates considerably from 31% to 84%, with modal rates ranging from 45% to 60% (Cascardi, O'Leary, & Schlee, 1999). The prevalence PTSD among victims of IPV varies from culture to culture (American Psychiatric Association, 2000) with Spanish forensic samples (Vilariño et al., 2009), and those receiving psychological treatment in a non forensic setting (Echeburúa & Corral, 1998) estimated to be around 55%. Psychological injury is comorbid with several disorders, secondary or indirect injury in the context of forensic evaluation such as depression, social alienation, anxiety, and sexual dysfunctions (Bargai et al., 2007; Kessler et al., 1995). Nevertheless, the diagnosis of these disorders in the absence of PTSD does not constitute sufficient forensic evidence of psychological injury (O'Donnell, Creamer, Bryant, Schnyder, & Shalev, 2006). Moreover, forensic evaluation inextricably requires a differential diagnosis of feigning i.e., to suspect and therefore rule out alternative hypothesis to psychological injury by establishing differential diagnosis (American Psychiatric Association, 2000). The dual task of diagnosing psychological injury and a differential diagnosis of feigning underlines the need for a multimethod approach (Arbisi, 2005; Polusny & Arbisi, 2006) based on a clinical interview and a psychometric instrument, generally the MMPI-2 (Graham, 2006; Greene, 2008; Pope, Butcher, & Seelen, 2006; Resnick, West, & Payne, 2008; Rogers, Sewell, Martin, & Vitacco, 2003). Though the MMPI-2 includes measures to control protocol validity which may be of great value for establishing a differential diagnosis, it does not in itself constitute sufficient evidence as the diagnosis of feigning is compatible with other alternative hypotheses; it does not provide diagnosis, but rather diagnostic impressions (Graham, 2006), and it does not correctly classify all of the feigners (Rogers et al., 2003). Moreover, the MMPI-2 does not fulfil the legal requirement of establishing a forensically valid causal relationship between the allegations of IPV and psychological injury as it does not rest on objective criteria, but rather on clinical intuition or the good faith of a victim´s unsubstantiated allegations as opposed to scientifically valid evidence (Steller, Raskin, Yuille, & Esplin, 1990). As for the interview, the structured clinical interviews are the standard of reference for forensic assessment of psychological injury, specifically the Structured Clinical Interview of the DSM-IV (SCID-IV) (Spitzer, Williams, Gibbon, & First, 1995), the Clinician Administered PTSD Scale of the DSM-IV (CAPS) (Blake et al., 1998), the Structured Interview for PTSD (SIP) (Davidson,

Malik, & Travers, 1997), the PTSD Symptom Scale-Interview (PSS-I) (Foa, Riggs, Daneu, & Rothbaum, 1993), and in Spain the "Escala de Gravedad de Síntomas del Trastorno de Estrés Postraumático" (Echeburúa, Corral, Amor, Zubizarreta, & Sarasua, 1997). These structured interviews based on checklists rely on the interviewee´s self-report of symptoms of PTSD. This interview format is not valid for forensic contexts as it lacks any specific and efficacious means of establishing a differential diagnosis of feigning, and is vulnerable to feigning even by subjects naive to PTSD criteria with feigning rates ranging from 86 to 94%, or 100% in trained populations (Resnick et al., 2008; Vilariño et al., 2009). As for the differential diagnosis of feigning, the DSM-IV-TR does not in effect *diagnose* but rather *suspects* feigning if any combination of the following criteria are present: 1) medicolegal context of presentation; 2) marked discrepancy between the person's claimed stress or disability and the objective findings; 3) lack of cooperation during the diagnostic evaluation and in complying with the diagnostic evaluation; and 4) the presence of antisocial personality disorder. These criteria do not enable the diagnosis of PTSD or a differential diagnosis of feigning; hence, they fail to respond to the requirements of forensic settings, are inefficacious for detecting feigning, and serve only to suspect feigning (Rosenfeld, Green, Pivovarova, Dole, & Zapf, 2010). In fact, standard structural clinical interviews are unsatisfactory for meeting the dual objective of clinical diagnosis and the control of feigning since inherently they do not presuppose feigning from the outset (Rogers, 2008a). Furthermore, this checklist type structured interview format cannot establish a causal relationship between victims allegations and psychological injury as this nexus is never contested, and it is assumed it can be accurately ascertained from the allegations themselves, which runs counter to the forensic precept of scientifically attesting criteria (Steller et al., 1990). This task is further exasperated in cases of IPV where PTSD symptoms may be due to a multiplicity of concurring causal factors such as separation or divorce, anxiety when faced with life disruption, fear of an uncertain future, and loss of control following separation or divorce (e.g., financial hardship, sole responsibility for the care of children, low self-esteem linked to self-perceptions of little ability to deal with the situation), or anxiety regarding the legal predicaments which the victim encounters. Thus, victims affirmatively pinpointing specific symptoms on these interviews can be misleading given that symptoms have common pathways with many other conditions, and one cannot safely ascertain they are the consequence IPV. The Structured Interview of Reported Symptoms, SIRS (Rogers, Bagby, & Dickens, 1992),

has demonstrated to be a reliable and valid psychometric instrument for the differential diagnosis of feigning. Notwithstanding, the SIRS has not been validated for the assessment of psychological injury in terms of PTSD, and specifically for IPV. Moreover, the SIRS is not sensitive to the forensic psychologist's statutory requirement of establishing a causal relationship. In order to overcome the limitations of structured interviews, a forensic-clinical interview was devised for forensic settings (Arce and Fariña, 2001). This interview format consists of narrative models whereby individuals create narrative accounts to describe an event or situation, and in particular *Anchored Narratives* (Wagenaar, 1995 Wagenaar, Van Koppen, & Crombag, 1993), that assert that narrative accounts of episodes and events under anomalous conditions (e.g., feigning, lying) in contrast to honest accounts are intrinsically different. The validity of these models for the forensic evaluation of psychological injury have systematically shown that, upon demand, individuals are capable of building a narrative account of their mental health (Rogers, 2008a; Arce, Pampillón, & Fariña, 2002), to the extent that the diagnostic criteria of the DSM is derived from the narratives (and the statistical data) of patients. This interview format is an extension of the cognitive interview (CI) of mental health (Fisher & Geiselman, 1992). Designed for obtaining testimony in judicial contexts, the CI is more productive than the standard structured interviews for obtaining information (Köhnken, Milne, Memon, & Bull, 1999; Memon, Meissner, & Fraser, 2010), and for discriminating between honest and false narrative accounts of reality (Vrij, 2005). The CI consists of 4 techniques: mentally reinstating the context, in-depth reporting of events, reporting the events from different perspectives, and recalling in reverse order. The CI aims to elicit the testimonies of witnesses and victims without contaminating the evidence i.e., retrieved memories, and to obtain testimonies that comply with legal and procedural safeguards (Colwell, Hiscock, & Memon, 2002). The first two techniques are free from any external contamination, whereas the two latter techniques involve interrogation techniques, and are concerned with past events that are not a characteristic of mental health (i.e., the objective of the evaluation of mental health is not to ascertain the facts but the symptoms). The use of the first two techniques i.e., mentally reinstating the context, and in-depth reporting of events, offers results similar to the full CI (Davis, McMahon, & Greenwood, 2005), and controls any potential contamination derived from interrogations. Based on the tenets of CI described in the literature, the forensic-clinical interview focuses on subjects using a free recall format to inform of all the personal changes in their life (i.e., symptoms, behaviours, thoughts, personal feelings, and emotions) prior to and following the

traumatic event, and the improvements victim have experienced since traumatic victimization has ceased (this contingency is frequent in cases of continued long-term IPV). Having completed this interview technique, interviewers encourage the reinstatement of the contexts of reference for reporting symptomatology taken from the V axis of the DSM-IV using an in-depth free narrative technique to report the impact on interpersonal and social relationships, family relationships, academic/occupational or other important areas of functioning. This interview also entails a procedure for the differential diagnosis of feigning. This procedure has demonstrated to be reliable and valid for the forensic diagnosis of psychological injury, and the differential diagnosis of feigning in cases of psychological injury sustained in motor vehicle accidents (Arce, Fariña, Carballal, & Novo, 2006).

Bearing in mind the goal of forensic evaluation is to diagnose psychological injury with a differential diagnosis of feigning, and to establish a causal relationship between the alleged IPV and PTSD, the aim of this study was to assess the efficacy of the forensic-clinical interview in discriminating between genuine victims and feigners of IPV; the differential diagnosis of feigning; the prevalence of psychological injury in terms of PTSD; and to compare reported rates with predicted rates of PTSD.

Method

Participants

A total of 105 women aged 19 to 73 years ($M = 33.56$, $SEM = 1.09$), participated in the study. Of the participants, 51 women, age range 19 to 64 years ($M = 37.61$; $SEM = 1.48$), were real victims of IPV who had reported the offence and secured a firm conviction against their aggressor in a court of law. All of the cases of real IPV involved both physical and psychological violence, given that the type of violence mediates the prevalence of psychological injury i.e., PTSD (Pico-Alfonso, 2005). The remaining 54 women, age range 21 to 73 years ($M = 29.74$; $SEM = 1.41$), were living with a partner, and had no history of IPV.

Measurement instruments

All participants underwent a forensic-clinical interview (Arce & Fariña, 2001). In order to contrast the efficacy of the forensic-clinical interview, the Spanish version of

the MMPI-2 (Hathaway & McKinley, 1999) was administered. Since the primary objective was to measure psychological injury inflicted by IPV i.e., PTSD, and the differential diagnosis of feigning, the validity scales and indices of reference for the assessment of feigning, and the Pk Scale (Keane, Malloy, & Fairbank, 1984) to measure psychological injury were estimated. The list of scales and indices of the MMPI-2 for the assessment of feigning, and the decision criteria were taken from Graham (2006) and Rogers et al. (2003).

Design and procedure

A quasi-experimental research methodology was used with archive data and data from the normal population. The experimental design aimed to measure psychological injury with a differential diagnosis of feigning in genuine and mock victims of IPV using a psychometric instrument involving a symptom recognition task, and a forensic-clinical interview involving a knowledge task. As for design sensitivity analysis for a sample of 105 participants, the results showed 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 chi-squared test $(df = 1)$ is $> 80\%$.

The evaluations of genuine victims were drawn from the archives of the Forensic Psychology Institute of the University of Santiago de Compostela (Spain). The inclusion criteria for the real victims 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 in a court of law. There was no evidence in judicial files for real victims of previous psychological distress. 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 screening questionnaire on instances of IPV. Mock victims were contacted and assessed individually to establish a sociodemographic correspondence (e.g., age, social status, number of children) with the real victim group. Feigners were informed about the purpose of the study and freely volunteered to participate in the study. Mock victims received feigning instructions in line with the recommendations of Rogers (2008b) 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; Arce, Fariña, Carballal, & Novo,

M. Vilariño et al.

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 for 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 (though participants were requested to withdraw from the study if they unwilling to comply with the instructions, they all freely volunteered 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. Furthermore, all mock victims showed the ability to discriminate between expected and unexpected symptoms the consequence of IPV. Moreover, 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 Scale T score $\geq$ 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.

The interviews were recorded on video for subsequent content analysis. The clinical protocols were obtained by 11 interviewers. The order of data gathering (MMPI-2 and the forensic-clinical interview) was rotated.

Analysis of protocols

The audiovisual recording of the free narrative interviews underwent systematic content analysis to detect diagnostic criteria of psychological injury using the categories of analysis in the DSM-IV, a categorical system referred to by Weick (1985) as a methodic system of categories. Thus, the categories of analysis were composed of diagnostic criteria for PTSD described in the DSM-IV (American Psychiatric Association, 2000) with the exception of the Criterion A1 ('the person experienced,

witnessed, or was confronted with an event or events that that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others') that in our study was assumed to be the victimization of IPV; and Criterion E ('the symptoms on Criteria B, C and D last for more than one month') which is not applicable to forensic settings given that in judicial terms the criterion is injury regardless of time duration i.e., a discrepancy between clinical and legal criteria. This is in line with the subsection 'Use of DSM-V in forensic settings' in the introduction of the DSM-IV (American Psychiatric Association, 2000) refers to the potential misuse or misunderstanding of clinical criteria in judicial setting.

The detection of content categories in the interview is grounded on two complementary methods: the subject reports in the interview, and the encoder's observations. Thus, deteriorated memory may be the subject's direct observation or the encoder's inference from protocols analysis. The analysis of internal consistency of the measurement scale revealed a Cronbach alpha coefficient of .850 ($N = 105$).

Moreover, the two encoders too valuated the protocols to detect the 9 feigning strategies. An exhaustive list of strategies was selected from the literature on feigning strategies (see for a review Rogers, 2008b), which was combined with a procedure of successive approximations. The procedure for creating a methodic categorical system i.e., mutual exclusion, homogeneity, objectivity, adequacy or pertinence, exhaustion and productivity, was applied to obtain a reliable and valid system (Bardin, 1977). This procedure ensured that the feigning strategies of rare or quasi-rare symptoms described in the literature as two distinct categories were fused into one joint category, as were symptom combination and spurious patterns of psychopathological categories. The resulting list of strategies and their corresponding definitions are listed below:

a) Subtle symptoms. Subtle symptoms are not real symptoms, but everyday problems which are regarded as symptoms associated to mental illness (i.e., to be unorganized, lack of motivation, and difficulty in ordinary decision-making).

b) Improbable/absurd symptoms. Improbable symptoms are fantastic or ridiculous in nature (opinions, attitudes or bizarre beliefs) and do not respond to real referents, with the exclusion of rare symptoms.

c) Obvious symptoms. These are psychotic symptoms related to what is vulgarly known as madness or mental illness.

d) (Quasi) rare symptoms. This category refers to symptoms described by participants though they are rarely observed even in real psychiatric populations or infrequently found in normative samples.

e) Symptom combination and spurious patterns of psychopathology. This indicator of feigning includes real symptoms reported by participants but rarely occur simultaneously; or when the participant describes an indiscriminate array of symptoms that have no internal consistency among them; or configurations that are uncommon in clinical populations.

f) Severity of symptoms. As the term indicates, the category analyzes the degree of symptom severity. Feigners frequently over exaggerated their symptom severity.

g) Inconsistency of symptoms (observed or manifest). The category analyses the association between the symptoms described by the participant and the encoder's observation regarding the concordance between the symptoms and the participant's attitude, composure and/or behaviour.

h) Erroneous stereotypes. This category refers to common misconceptions about which clinical characteristics are commonly associated with mental disorders.

i) Indiscriminate symptom endorsement. This category implies the tendency to endorse a large proportion of symptoms.

Analysis of the reliability of productive strategies for detecting feigning, revealed an α of .744 ($n = 54$, strategies were only productive for the subsample of feigners).

The order for encoding the clinical diagnostic criteria and feigning strategies was rotated. The whole unit of analysis in all of the categories was the protocol, and both the criteria and strategies were registered as present or absent. As most of the categories were not related to the overall narrative account, the correspondence between the encodings was confirmed. In fact, the direct encoding of the categories was not sensitive to the corresponding measures. In other words, the encoding of a given category as present did not imply that the measure was referring to exactly the same incident or cognition. Thus, encoders did not only register the symptom, but the incidents and cognitions where the symptom was registered. This enables us to check if a symptom registered for both encoders is referring precisely to the same contingencies (it was registered as consistent if they were referring to the same incident or cognition, and inconsistent if not so). Moreover, the encoders registered each symptom to determine if there was a causal relationship with the allegations of IPV. Following content analysis

of each interview, the encoders evaluated if the detected criteria constituted diagnosis for PTSD, and whether each of these criteria were derived from IPV.

Table 1. Within- and Between-Encoder Consistency of PTSD Criteria and Feigning Strategies: Kappa Coefficient.

Criterion/Strategy	Within1	Within2	Between
Responses involved intense fear	1	1	.92
Recurrent or intrusive recollections of the event	1	1	1
Recurrent distressing dreams of the event	1	1	.8
Acting as the traumatic event was recurring	.73	1	1
Psychological distress at exposure to reminders	1	1	1
Physiological reactivity on exposure to reminders	1	.8	.61
Efforts to avoid thoughts about the trauma	1	1	.64
Efforts to avoid places that remind the event	1	.64	1
Inability to recall part of the event	1	1	1
Diminished interest in significant activities	.86	.65	.64
Feelings of detachment	1	.65	1
Restricted affect	1	1	1
Foreshortened future	1	.8	.65
Falling or staying asleep	.81	1	.64
Irritability or anger	1	1	1
Difficulty concentrating	1	1	.81
Hypervigilance	1	1	.1
Exaggerated startle responses	1	1	.81
Clinically significant distress	1	1	.65
Subtle symptoms	1	1	1
Severity of symptoms	1	1	.8
Indiscriminant symptom endorsement	1	1	1
Causal nexus IPV and PTSD	1	1	1

Note. The feigning categories inconsistency of symptoms (observed or manifest), erroneous stereotypes, improbable/absurd symptoms; (quasi)rare symptoms, symptom combination and spurious patterns, and obvious symptoms, were unproductive.

Training of encoders

The protocols were encoded by two encoders with extensive experience in encoding this type of material, and with a sound knowledge of psychopathological evaluation (Arce et al., 2006, 2009). The encoders were exhaustively trained in the encoding procedure i.e., each category of analysis was presented and exemplified with material from other claimants not pertaining to this study, using the concordance as the instrument to verify between-encoder correspondence in order to detect inconsistencies, and correct them by homogenizing the criteria.

The categories of analysis of the clinical disorders were defined according to the diagnostic criteria of the DSM-IV whereas the definitions and examples of the categories of feigning strategies were based on the work of Rogers (2008b). Each encoder was supplied with a copy of the manuals with the definitions for each category as well as our manual with examples for each category.

Analysis of the reliability of the interviews and encodings

The interviews were conducted by 11 forensic psychologists with extensive experience in using this type of interview format based on free recall, and techniques designed to enhance the process of memory retrieval. In order to assess the effects of interviewer on the interview (reliability of the measure), the protocols of real victims and feigners were randomly assigned to one of two groups. Thus, an equal quantity of symptomatology observed in the protocols would indicate the interviewer had no effect on the interviews. The results showed a similar number of symptoms were registered in the protocols of real victims, $F(1, 49) = 0.32$; *ns*, $\eta^2 = .006$, 1-ß $= .085$, and feigners, $F(1, 52) = 0.38$; *ns*, $\eta^2 = .009$; 1-ß $= .093$. This indicated the interviews were not contaminated by the interviewer factor. Likewise, the interviewers have proven to be consistent and productive in other studies (e.g., Arce et al., 2006).

The material was distributed equally between the encoders by randomly combining the interviews of genuine victims with mock victims.

To estimate intra-encoder reliability, the encoders repeated the encoding of 20 of their own interviews a week after having completed all of the initial encodings. Similarly, to estimate inter-encoder reliability, the encoders repeated the encoding of 20 of each other's interviews a week after having completed all of the initial encodings. The results (see Table 1) revealed encoding reliability (kappa > .61).

Results

No instances of feigning strategies were registered among genuine victims of IPV. However, the forensic-clinical interviews of mock victims showed that 31 of 54 feigners had used at least one feigning strategy (two strategies were jointly used in 3 cases: indiscriminate grouping of symptoms and symptoms severity) i.e., around 50% of feigners, $\chi^2(1) = 0.49$, *ns*, were not detected by feigning strategies. Feigners employed three strategies to feign clinical injury: reporting subtle symptomatology (do not report real symptoms, but rather everyday problems which are confused with symptomatology associated to mental disorders); indiscriminate grouping of symptoms (indiscriminate reporting of clinical problems) and symptom severity (maximum symptom severity). Of the three productive strategies, feigners exceeded significantly, i.e., exceeded the statistically admissible margin of error, .05, in reporting subtle symptoms, and maximum symptoms severity (see Table 2) i.e., these strategies were present with a greater than the admissible margin of error, whereas symptom combination was marginal (within the limits of the statistically admissible margin of error). Moreover, feigners did not resort to strategies characteristic of psychotic conditions [i.e., improbable/absurd symptoms; (quasi)rare symptoms improbable symptoms, obvious symptoms], nor report symptom combination and spurious patterns of psychopathology or erroneous stereotypes, as well as exhibiting consistency between the reported and observed symptoms.

Table 2. Z Scores for the Feigning Strategies Registered in Forensic-Clinical Interview (Test Value = .05).

Strategy	*f(%)*	**Z**	*p*
Subtle symptoms	23(42.6)	12.68	.001
Indiscriminant symptom endorsement	3(5.6)	0.20	*ns*
Severity of symptoms	8(14.8)	3.30	.001

Note. *n* = 54. *f(%)* = frequency(percentage). The non-listed categories were unproductive. Among genuine victims, no feigning strategy was detected.

The corrected correlation coefficient by predictor and criterion reliability of the scales and indices of the MMPI-2 indicating feigning (F, Fb, Fp, F-K, S-O, FBS, Ds) was .478 for subtle symptom strategy, and .569 for symptom severity strategy, with a

large effect size for both contingencies (*r* around .50). Hence, these results illustrate the convergent and discriminant validity (no genuine victims were informed as feigners) of the feigning strategies detection technique in the forensic-clinical interview.

The comparison of PTSD criteria observed in genuine IPV victims with feigners, revealed real victims (*M* = 9.3) reported more clinical symptomatology related to PTSD, $F(1, 103) = 42.89$, $p < .001$, $\eta^2 = .294$, 1-ß = 1, than feigners (*M* = 4.52). The contrast of genuine and mock victims showed a greater prevalence of PTSD criteria (see Table 3) among the former in: responses involved intense fear; recurrent or intrusive recollections of the event that produce anxiety including images, thoughts and perceptions; recurrent distressing dreams of the event; acting as the traumatic event was recurring; intense psychological distress at exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event; physiological reactivity on exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event; efforts to avoid thoughts, feelings, or conversations associated with the trauma; efforts to avoid activities, places, or people that arouse recollections of the trauma; markedly diminished interest or participation in significant activities; sense of a foreshortened future; difficulty falling or staying asleep; irritability or outbursts of anger; hypervigilance; exaggerated startle responses; the disturbance causes clinically significant distress or impairment in social, occupational, or other important areas of functioning. In relation to expected psychological injury, the diagnosis of PTSD was greater in real victims of IPV, (58.8%), $\chi^2(1) = 32.11$, $p < .001$, $\varphi = .537$, than feigners (5.6%), with a large effect size ($r > .50$). This prevalence among genuine victims was similar to the rate reported in the Spanish literature on IPV, (55%), $Z(n = 51) = 0.54$, *ns*. Nevertheless, the measure of the psychological injury on the Pk scale of the MMPI-2 showed a higher rate of PTSD among feigners (*n* = 41, 75.9%) than among real victims (*n* = 28, 54.9%) of IPV, $\chi^2(1) = 4.83$, $p < .05$, $\varphi = .221$, with a small effect size ($r > .10$). Nonetheless, the efficacy of feigners in the forensic-clinical interview was 5.6% in comparison to 75.9%, for the psychometric instrument i.e., the Pk Scale of the MMPI-2 (T ≥ 70), indicating the psychometric evaluation was vulnerable to feigning, $\chi^2(1) = 955.38$, $p < .001$, $\varphi = -.716$, with a large effect size ($r > .50$). As for the corrected correlation coefficient by predictor and criterion reliability between the diagnosis of PTSD in the forensic-clinical interview and the Pk Scale of the MMPI-2 in real victims of IPV was .923 ($p < .001$) with a large effect size ($r > .50$), and .159 (*ns*) for feigners. In short, the results substantiate the validity of the forensic-clinical interview both in

predictive validity (diagnosed PTSD was similar to the expected rate); convergent validity (PTSD diagnosed in the interview correlated with the diagnostic impression on the Pk Scale), and discriminant validity (PTSD diagnosed in the forensic-clinical interview was not correlated to the Pk Scale in feigners, in which the rate of PTSD was higher than for real victims of IPV).

Table 3. χ² Test of PTSD Criteria (Registered vs. Non-Registered) by Sample (Genuine vs. Mock Victims).

Criterion	%mv	%rv	χ^2	p	φ
Responses involved intense fear	72.2	94.1	7.38	.007	-.290
Recurrent or intrusive recollections of the event	27.8	51	5.00	.025	-.238
Recurrent distressing dreams of the event	13	43.1	10.48	.001	-.337
Acting as the traumatic event was recurring	3.7	17.6	4.05	.044	-.228
Psychological distress at exposure to reminders	40.7	80.4	15.57	.000	-.405
Physiological reactivity on exposure to reminders	16.7	45.1	8.71	.003	-.309
Efforts to avoid thoughts about the trauma	25.9	60.8	11.63	.001	-.352
Efforts to avoid places that remind the event	16.7	51	12.40	.000	-.364
Inability to recall part of the event	0	7.8	2.52	.112	-.205
Diminished interest in significant activities	44.4	70.6	6.29	.012	-.264
Feelings of detachment	27.8	41.2	1.54	.215	-.141
Restricted affect	20.4	39.2	3.62	.057	-.206
Foreshortened future	18.5	39.2	4.54	.033	-.229
Falling or staying asleep	27.8	64.7	12.96	.000	-.370
Irritability or anger	13	49	14.44	.000	-.391
Difficulty concentrating	14.8	31.4	3.19	.074	-.197
Hypervigilance	13	35.3	6.03	.014	-.262
Exaggerated startle responses	5.6	29.4	8.90	.003	-.316
Clinically significant distress	50	74.5	5.68	.017	-.252

Note. df(1). %mv = % in mock victims; %rv = % in genuine victims.

Notwithstanding, the forensic-clinical interview in itself failed to provide sufficient evidence for the dual task of diagnosing PTSD and establishing a differential diagnosis of feigning given that feigners who managed to successfully feign PTSD in the interview were not detected by the feigning strategies.

Discussion

The limitations of this study should be borne in mind and caution should be exercised in generalizing the findings to other populations and case types. First, allegations of IPV are considered to be genuine if a conviction is secured in a court of law; however, this conjecture may be tentative rather than absolute. Second, it is assumed real victims of IPV suffering from psychological injury are conscious and able to appraise and report their psychological injuries, and more specifically in a recognition task, but this may not be the case. Thus, there is no guarantee that psychological injury in cases of IPV can be effectively ascertained. In judicial terms, however, what cannot be proven by attestable evidence simply does not exist. Third, a judge's rulings may dismiss genuine allegations on the grounds of insufficient evidence or the victims may invalidate their own allegations, indicating the sample of genuine victims may not be truly representative of the general population of victims. Forth, it is assumed that there is parity between feigning under real conditions and laboratory high fidelity recreations yet both circumstances are not exactly identical (Fariña, Arce, & Real, 1994; Konecni & Ebbesen, 1992). Fifth, this study exclusively assessed IPV thus caution should be taken in extrapolating these results to other populations and case types. Sixth, the results obtained for feigners were neither based on archived data nor replicated the legal process, which poses certain limitations in terms of the ecological validity of the design. Bearing in mind the aforementioned limitations, the following conclusions for forensic practice may be drawn:

a) The forensic-clinical interview is a reliable and valid instrument for measuring psychological injury in cases of IPV. Moreover, it is also valid in fulfilling judicial requirement of establishing a causal relationship between allegations of IPV and PTSD symptoms.

b) The forensic-clinical interview is a reliable and valid measure instrument for the differential diagnosis of feigning.

c) The forensic-clinical interview does not on its own enable forensic diagnosis i.e., evaluation of psychological injury in terms of PTSD, and a differential diagnosis of feigning by controlling false negatives (identifying feigners as honest) that are entirely inadmissible in forensic settings.

Acknowledgements

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arbisi, P. A. (2005). Use of the MMPI-2 in personal injury and disability evaluations. In J. N. Butcher (Ed.), *Practitioners handbook for the MMPI-2* (pp. 407-42). Washington, DC: American Psychological Association.

Arce, R., & Fariña, F. (2001). *La entrevista clínico forense* [The clinic-foresenc interview]. Santiago de Compostela, Spain: Universidad de Santiago de Compostela.

Arce, R., Alonso, M. A., & Vilariño, 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 judgements 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* (pp. 241-252). Santiago de Compostela, Spain: Nino.

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., Pampillón, M. C., & Fariña, F. (2002). Desarrollo y evaluación de un procedimiento empírico para la detección de la simulación de enajenación mental en el contexto legal [Development and validation of an empirical procedure to

detect malingering of criminal insanity in legal context]. *Anuario de Psicología*, *33*, 385-408.

Bardin, L. (1977). *L'analyse de contenu* [Content analysis]. Paris, France: PUF.

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. doi: 10.1007/s10896-007-9078-y

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Charney, D. S., & Keane, T. M. (1998). *Clinican-administered PTSD scale for DSM-IV*. Boston, MA: National Center for Posttraumatic Stress Disorder.

Cascardi. M., O'Leary, K. D., & Schlee, K. A. (1999). Co-occurrence and correlates of posttraumatic stress disorder and major depression in physically abused women. *Journal of Family Violence, 14*, 227-247. doi: 10.1023/A:1022827915757

Colwell, K., Hiscock, C. K., & Memon, A. (2002). Interviewing techniques and the assessment of statement credibility. *Applied Cognitive Psychology, 16*, 287-300. doi: 10.1002/acp.788

Davidson, J., Malik, M., & Travers, J. (1997). Structured Interview for PTSD (SIP): Psychometric validation for DSM-IV criteria. *Depression and Anxiety*, *5*, 127-129. doi: 10.1002/(SICI)1520-6394(1997)5:3<127::AID-DA3>3.0.CO;2-B

Davis, M. R., McMahon, M., & Greenwood, K. M. (2005). The efficacy of mnemonic components of the cognitive interview: Towards a shortened variant for time-critical investigations. *Applied Cognitive Psychology*, *19*, 75-93. doi: 10.1002/acp.1048

Echeburúa, E., & Corral, P. (1998). *Manual de violencia familiar* [Handbook of domestic violence] Madrid, Spain: Siglo XXI.

Echeburúa, E., Corral, P., Amor, P. J., Zubizarreta, I., & Sarasua, B. (1997). Escala de Gravedad de Síntomas del Trastorno de Estrés Postraumático: Propiedades psicométricas [The Severity of Symptom Scale of Posttraumatic Stress Disorder: Psychometric properties]. *Análisis y Modificación de Conducta, 23*, 503-526.

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.

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interview*. Sprinfield, IL: Charles C. Thomas.

Foa, E. B., Riggs, D. S., Daneu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing posttraumatic stress disorder. *Journal of Traumatic Stress*, *6*, 459-473. doi: 10.1002/jts.2490060405

Graham, J. R. (2006). *MMPI-2: Assessing personality and psychopathology* (4th ed.). New York, NY: Oxford University 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.

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. doi: 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.

Köhnken, G., Milne, R., Memon, A., & Bull, R. (1999). The cognitive interview: A meta-analysis. *Psychology, Crime & Law, 5*, 3-27. doi: 10.1080/10683169908414991

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.

Memon, A., Meissner, C. A., & Fraser. J. (2010). Cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law, 16,* 340-372. doi: 10.1037/a0020518

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.

Pico-Alfonso, M. A. (2005). Psychological intimate partner violence: The major predictor of posttraumatic stress disorder in abused women. *Neuroscience and Biobehavioral Reviews, 29*, 181-193. doi: 10.1016/j.neubiorev.2004.08.010

Polusny, M. A., & Arbisi, P. A. (2006). Assessment of psychological distress and disability after sexual assault in adults. In G. Young, A. W. Kane, & K. Nicholson (Eds.), *Psychological knowledge in courts. PTSD, pain and TBI* (pp. 97-125). New York, NY: Springer.

Pope, K. S., Butcher, J. N., & Seelen, J. (2006). *The MMPI, MMPI-2, and MMPI-A in court. A practical guide for expert witnesses and attorneys*. Washington, DC: American Psychological Association.

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). Current status of clinical methods. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 391-410). New York, NY: The Guilford Press.

Rogers, R. (2008b). 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. doi: 10.1177/1073191103252349

Rosenfeld, B., Green, D., Pivovarova, E., Dole, T., & Zapf, P. (2010). What to do with contradictory data? Approaches to the integration of multiple malingering measures. *International Journal of Forensic Mental Health, 9,* 63-73. doi: 10.1080/14999013.2010.499559

Sarasua, B., Zubizarreta, I., Echeburúa, E., & Corral, P. (2007). Psychopathological profile of battered women according to age. *Psicothema*, *19*, 459-466.

Spitzer, R. L., Williams, J. B., Gibbon, M., & First, M. B. (1995). *Structured clinical interview for DSM-IV*. Washington, D.C.: American Psychiatric Press.

Steller, M., Raskin, D. C., Yuille, J. C., & Esplin, P. (1990). *Child sexual abuse: Forensic interviews and assessment*. New York, NY: Springer.

United Nations. (1988). *Committee on crime prevention and control. Report on the tenth session*. Vienna, Switzerland: Author.

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.

Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy and Law, 11*, 3-41. doi: 10.1037/1076-8971.11.1.3

Wagenaar, W. A. (1995). Anchored narratives: A theory of judicial reasoning. In G. Davies, S. Lloyd-Bostock, M. McMurram, & C. Wilson (Eds.), *Psychology, law and criminal justice* (pp. 267-285). Berlin, Germany: Walter de Gruyter.

Wagenaar, W. A., Van Koppen, P. J., & Crombag, H. F. M. (1993). *Anchored narratives. The psychology of criminal evidence*. New York, NY: St Martin's Press and Hertfordshire: Harvester Wheatsheaf.

Weick, K. E. (1985). Systematic observational methods. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology bulletin* (Vol. 1, pp. 567-634). Hilldsale, N.J.: LEA.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: Author.