Psicothema 2009. Vol. 21, nº 2, pp. 241-247
www.psicothema.com

ISSN 0214 - 9915 CODEN PSOTEG
Copyright © 2009 Psicothema

Creación y validación de un protocolo de evaluación forense de las secuelas psicológicas de la violencia de género

Ramón Arce, Francisca Fariña*, Alicia Carballal y Mercedes Novo
Universidad de Santiago y * Universidad de Vigo

El daño psíquico conforma una de las secuelas de la victimación de un delito. La literatura ha identificado el Trastorno de Estrés Postraumático (TEP) como la huella psicológica. Ahora bien, no es suficiente con diagnosticar un TEP dado que, en el ámbito legal, es preciso descartar la simulación. Para conocer las aptitudes y modos de detección de la simulación en casos de violencia de género, planificamos una investigación en la que solicitamos a 101 mujeres que simularan la huella psíquica del maltrato en el MMPI-2 y en una entrevista clínico-forense. Los resultados mostraron que las mujeres gozaban de una capacidad general (84.2%) de simulación en el MMPI-2. Por su parte, las medidas de validez fueron efectivas en la detección de la simulación, pero permitían un margen de error muy elevado. En la entrevista clínico-forense, la tarea de simulación resultó poco accesible: sólo 3 participantes lograron simular un TEP. Además, en dichos protocolos no se observó estrategia alguna de simulación. Por ello, contrastamos una aproximación multimétodo que clasificó correctamente todas las simulaciones. Finalmente, formulamos un protocolo de evaluación forense de la huella psíquica en casos de violencia de género.

Creation and validation of a forensic protocol to assess psychological harm in battered women. Psychological harm is one of the sequela of the victimization of a crime. Nevertheless, a diagnosis of psychological harm, that is, posttraumatic stress disorder (PTSD), does not constitute sufficient evidence alone given that, in legal terms, faking or false testimony must be detected and eliminated before an expert testimony can be admitted. To assess the aptitudes and strategies for faking psychological harm associated with domestic violence, 101 women were asked to fake psychological harm using the MMPI-2 and a forensic clinical interview. The results showed that most women (84.2%) were capable of faking in MMPI-2. The original validity scales of the MMPI-2 and their configurations were effective for the detection of faking though with a wide margin of error was observed. In the forensic clinical interview, it was much more difficult to simulate since only 3 malingerers, in which protocols were not observed any feigning strategy, were capable of simulating the psychological injury linked to domestic violence. Consequently, a multimethod approach (MMPI and interview) was contrasted and observed to classify correctly all subjects. Thus, a protocol for forensic assessment of the psychological harm consequence of domestic violence against women was formulated.

La ONU define a la víctima como *aquella persona que ha sufrido un perjuicio, entendiendo por ello una lesión física o mental, sufrimiento emocional, pérdida o daño material, o un menoscabo importante en sus derechos, como consecuencia de una acción u omisión que constituya un delito con arreglo a la legislación nacional o el derecho internacional, o bien sea un acto de abuso de los poderes públicos. Esto se puede aplicar no sólo a la persona victimada, sino también a colectivos, tales como entidades o asociaciones, así como a los familiares de la víctima y aquellas que intentaron auxiliarla durante la perpetración del delito* (United Nations, 1988). En consecuencia, la victimación engloba cinco tipos de lesiones: física, mental, sufrimiento emocional, pérdida o daño material y pérdida o menoscabo en los derechos, pero como carga de prueba los daños se reducen fundamentalmente a físicos y psíquicos (v.gr., Milner y Crouch, 2004). El daño o huella psíquica se obtiene a través de la medida de los efectos de un acto delictivo en la salud mental, siendo el trastorno primario en violencia de género y en la práctica forense el Trastorno de Estrés Postraumático (TEP), y los secundarios la depresión, inadaptación social, ansiedad y disfunciones sexuales (i.e., Sarasua, Zubizarreta Echeburúa y Corral, 2007; Kessler et al., 1995). Pero la observación de los trastornos secundarios sin el primario, el TEP, no se puede considerar una secuela del hecho traumático (O'Donnell et al., 2006). Identificado el daño psíquico o huella psíquica de la violencia de género, el siguiente paso es su medida. Pero esta medida no es asumible en los términos clínicos ordinarios. De hecho,

Fecha recepción: 4-6-08 • Fecha aceptación: 6-11-08
Correspondencia: Ramón Arce
Facultad de Psicología
Universidad de Santiago de Compostela
15872 Santiago de Compostela (Spain)
E-mail: ramon.arce@usc.es

la evaluación clínica tradicional nunca ha diagnosticado simulación porque no la sospecha (i.e., Rogers, 1997), algo que debe desconfiarse y, por tanto, someterse a prueba en el contexto médico-legal (American Psychiatric Association, 2000). En suma, en la práctica de la prueba judicial no sólo hay que observar lesión psíquica o sufrimiento emocional, sino que también es imprescindible demostrarlo, controlando, al mismo tiempo, los efectos de la simulación. Así, si bien las cifras de simulación son desconocidas, se estima que la probabilidad de simulación de lesiones está en torno al 50% (Miller y Cartlidge, 1972), siendo especialmente elevada en casos que implican procesos de separación y divorcio, como ocurre mayoritariamente en los de violencia de género (Mikkelsen, Guthel y Emens, 1992).

Como consecuencia de lo señalado anteriormente y del contexto legal vigente, la Ley 1/2004, de 28 de diciembre, de Medidas de Protección Integral contra la Violencia de Género, se hace necesaria la creación y validación de un protocolo forense de medida del daño psíquico consecuencia de la violencia de género. Para ello, se precisa la adopción de una aproximación multimétodo, mostrándose como las más efectivas aquellas que combinan entrevista y el MMPI (Gothard, Rogers y Sewell, 1995; Rogers, 1997). En este contexto nos planteamos un estudio dirigido a verificar la capacidad de simulación de las consecuencias psicológicas asociadas a la vivencia de malos tratos en el MMPI-2 y una entrevista clínico-forense, al tiempo que someter a prueba la validez predictiva de los procedimientos de control de la simulación y el engaño disponibles en ambas medidas.

Método

Participantes

La muestra la compusieron un total de 101 mujeres, todas ellas mayores de edad, legas en Psicología, mentalmente sanas y que manifestaron no haber sufrido violencia de género. En relación con la edad, el rango oscilaba entre los 19 años y los 61 años (*M*= 32; *DT*= 12.5). Con respecto al estado civil, el 2.97% eran viudas, el 41.58% estaban casadas y el 55.45% solteras, teniendo o habiendo tenido todas ellas una pareja estable con convivencia.

Procedimiento y diseño

La metodología de investigación empleada fue del tipo experimental y en un ambiente de simulación de alta fidelidad. De hecho, para dotar de mayor realismo a este experimento y para lograr una mayor implicación de las participantes, recurrimos al ofrecimiento de una recompensa de tipo económico (150€) a las cuatro mejores simulaciones. A su vez, se planificó un contraste de la ejecución de los sujetos en dos formatos de tarea distintos: tarea de reconocimiento en el MMPI-2 (Hathaway y McKinley, 1999) y tarea de conocimiento en la *entrevista clínico-forense* (Arce y Fariña, 2001). En un primer momento, la tarea de las participantes consistió en responder al cuestionario MMPI-2 bajo las *instrucciones estándar* del propio MMPI-2 para un cotejo de su estado clínico. Se les comunicó que se iba a efectuar una evaluación clínica de cada una de ellas y que, si lo deseaban, se les haría un informe de su estado clínico. Todas participaron de forma voluntaria. Los pases se produjeron individualmente. Al terminar de cumplimentar el MMPI-2 se les explicó que volveríamos a reevaluarlas transcurrida una semana. Para esta nueva medida se las instruía en lo que denominamos *instrucciones de simulación*. Éstas consistían en solicitarles que se pusiesen en el lugar de una mujer que simulaba haber sufrido malos tratos en el ámbito doméstico y a la que se le van valorar las secuelas psicológicas de los mismos. La denuncia falsa de malos tratos se justificaba en base a la obtención de algún beneficio asociado (v.gr., guarda y custodia de los hijos, venganza, compensaciones económicas). Al cabo de una semana, las participantes fueron evaluadas por medio de la *entrevista clínico-forense*. Dos fueron los entrevistadores que recabaron todos los protocolos clínicos y que se repartieron el total de las entrevistas a partes alícuotas. Finalmente, en torno a una semana después, las participantes fueron evaluadas de nuevo a través del MMPI-2 bajo las *instrucciones de simulación*.

Instrumentos

Como instrumento de medida en la tarea de reconocimiento se utilizó el MMPI-2 baremado para España por TEA (Hathaway y McKinley, 1999). Dados los propósitos de medida clínica del daño psíquico directo e indirecto asociado a la victimación de violencia de género y de simulación, se tomaron las diez escalas clínicas básicas, las dos escalas adicionales de medida del TEP, y las escalas de validez. Por último, a partir de los resultados obtenidos de las escalas originales de validez del MMPI-2 se calcularon los índices que se habían mostrado efectivos en la detección de la simulación de trastornos (Duckworth y Anderson, 1995): el *índice F-K* y el *perfil en V invertida*. El *índice F-K*, también conocido como *índice de Gough*, se ha observado que es un indicador robusto de los intentos de proporcionar una mala imagen, en tanto el *perfil en V invertida* advierte de la exageración de síntomas (Nicholson et al., 1997). Finalmente, el estudio de los resultados nos llevó a prestar atención al estudio del *Perfil en V* porque, si bien se dirige a la medida de la disimulación, también puede desempeñar un papel significativo a través de la validez discriminante en el análisis de la simulación.

En la tarea de conocimiento, las participantes fueron sometidas a una *entrevista clínico-forense*, esto es, a una entrevista de orden clínico en formato de discurso libre que se ha mostrado fiable y productiva en el contexto forense (por ejemplo, Arce, Fariña, Carballal y Novo, 2006). El procedimiento consiste en pedir a los sujetos que relaten todo aquello que ha cambiado en su vida (i.e., síntomas, conductas, pensamientos) en relación con antes de haber sufrido malos tratos, o lo que ha mejorado tras dejar de sufrirlos. Si no informan de motu propio sobre las relaciones interpersonales, el contexto de trabajo o académico, relaciones familiares y, en su caso, de pareja, el entrevistador procedía a reinstaurar esos contextos específicos de evaluación (eje V del DSM-IV-TR).

Análisis de los protocolos

Las entrevistas, tras ser grabadas en vídeo, fueron sometidas a un análisis de contenido mediante un sistema categorial metódico. El objetivo de este análisis de contenido era detectar criterios, bien autoinformados bien observados, diagnósticos del daño psíquico primario, así como de cualquier otro trastorno clínico. Para ello tomamos como categorías de análisis los criterios diagnósticos del TEP y los restantes criterios diagnósticos recogidos en el DSM-IV-TR (American Psychiatric Association, 2000). Así, creamos un sistema categorial mutuamente excluyente, fiable y válido (Weick, 1985). Una vez analizado cada protocolo, los codificadores eva-

luaban si se cumplían o no los criterios diagnósticos de los trastornos mentales relacionados con el daño psíquico.

A su vez, se codificaron, por los mismos evaluadores pero en distintas tareas, 9 estrategias de simulación a las que los potenciales simuladores recurren a la hora de simular, siendo la unidad de análisis el protocolo. El sistema categorial fue tomado de Arce et al. (2006). Las categorías de análisis fueron: no cooperación con la evaluación; síntomas sutiles; síntomas improbables; síntomas obvios; síntomas raros; combinación de síntomas; severidad de síntomas; inconsistencia de síntomas (observados y manifestados); y agrupación indiscriminada de síntomas.

Entrenamiento de codificadores

En este estudio participaron dos codificadores con conocimientos de evaluación psicopatológica, uno de ellos con experiencia previa en codificación de este tipo de material (Arce et al., 2006). Los codificadores fueron exhaustivamente entrenados. El entrenamiento consistió, tras presentarles y ejemplificarles cada categoría de análisis, en la ejecución con material desechado del estudio, utilizando el índice de concordancia como instrumento de cotejo de la ejecución, lo que permitía, constatada inconsistencia, subsanar errores de codificación a través de la contrastación y homogeneización de criterios. Además, se les proporcionó un manual de creación propia con la definición y ejemplos de cada categoría de análisis.

Análisis de la fiabilidad de las entrevistas y codificaciones

Los dos entrevistadores del estudio, entrenados y con experiencia en la *entrevista clínico-forense*, obtuvieron unos protocolos similares, tanto en términos de la sintomatología observada, $F(1;99)= 0.97$; ns, como en las estrategias de simulación detectadas, $F(1;99)= 1.67$; ns. En suma, las entrevistas no están contaminadas por el factor entrevistador. A su vez, estos entrevistadores también se habían mostrado consistentes y productivos en otros estudios (por ejemplo, Arce et al., 2006).

Los protocolos clínicos procedentes de las entrevistas se dividieron en dos mitades equivalentes, de modo que cada codificador examinó en torno al 50%. Para calcular la consistencia intra-jueces, los codificadores repitieron, una semana después del fin de la codificación original, 10 (en torno al 20%) de las entrevistas que ellos mismos habían evaluado. Asimismo, para obtener la medida de la consistencia inter-jueces, cada codificador examinó otros 10 protocolos que inicialmente había registrado el otro codificador.

Se consideran evaluaciones concordantes aquellas que superan el punto de corte de IC[acuerdos/(acuerdos+desacuerdos]$\geq$.80 (Tversky, 1977), que es más restrictivo que los valores kappa. Los resultados mostraron en todas las variables unos índices superiores a .80, esto es, los datos son fiables.

Análisis de datos

Para establecer si los diversos síntomas clínicos del TEP, que conforma la huella psíquica de la violencia de género, eran accesibles o no a la simulación, tomamos como punto de contraste el concepto de significatividad estadística, el .05. De esta forma, aquellos síntomas informados en las entrevistas clínicas por las participantes bajo instrucciones de simulación cuya prevalencia superara significativamente el punto de corte de .05 se entiende que son accesibles a la simulación. Complementariamente, aquellos síntomas que no sobrepasaran este criterio no son accesibles a la simulación.

Resultados

Estudio del cambio en la evaluación clínica en el MMPI-2

Encontramos diferencias multivariadas en las escalas clínicas básicas del MMPI-2 mediadas por el factor instrucciones (estándar vs. simulación), $F_{multivariada}$ (10;91)= 52.466; p<.001; eta^2= .852. Además, la magnitud del efecto de las instrucciones fue muy elevada, explicando el 85% de la varianza.

Los efectos univariados (tabla 1) muestran diferencias significativas en todas las variables clínicas, a excepción de la Masculinidad-Feminidad. A su vez, los efectos se producen en la dirección esperada; esto es, un aumento en las puntuaciones en la condición de simulación, con lo que las participantes, en principio, presentan una aptitud general de simulación. Dicha capacidad es, además, efectiva en los trastornos secundarios asociados al daño psíquico producido por el maltrato: la Depresión y la Inadaptación Social (Introversión Social en el MMPI-2) (i.e., Sarasua et al., 2007; Freyd, 1996; Taylor y Koch, 1995). Asimismo se produce una elevación significativa y de tal dimensión que se estaría ante un indicio de patología severa (T>70), en línea con la simulación, en la tríada psicótica (Paranoia, Esquizofrenia y Psicastenia) al igual que en las medidas de Desviación Psicopática, Histeria e Hipocondriasis. Estas patologías no están directamente ligadas a las secuelas psíquicas de los malos tratos, por lo que, en este caso, no han ejecutado correctamente. Finalmente, se observa un cambio hacia la patología en la variable Hipomanía, pero sin llegar a entrar en la región de impresión de patología (T<70). En otras palabras, no discriminan efectivamente entre los esperados síntomas depresivos e introversión social y los no esperados de orden psicótico, Desviación Psicopática, Histeria e Hipocondriasis.

Por su parte, en las escalas específicas de medida del TEP, que conforma el cuadro clínico del diagnóstico primario de las secuelas de malos tratos en la práctica forense (Sarasua et al., 2007; Freyd, 1996; Taylor y Koch, 1995; Kessler et al., 1995), las escalas Pk, $F(1;100)= 614.45$; p<.001; eta^2=.860, y Ps, F(1;100)=

Tabla 1
Efectos univariados en las escalas clínicas básicas del MMPI-2 por el factor instrucciones. Efectos intra-sujetos

Escalas	F	p	Eta^2	M_{ie}	M_{is}	Potencia
Hipocondriasis	242.47	.000	.708	54.46	81.08	1.000
Depresión	248.41	.000	.713	50.03	77.21	1.000
Histeria	207.13	.000	.674	52.61	73.99	1.000
Desviación psicopática	348.96	.000	.777	49.3	73.00	1.000
Masculinidad-feminidad	1.98	.162	.019	51.92	50.12	.286
Psicastenia	214.25	.000	.682	51.38	74.75	1.000
Paranoia	260.9	.000	.723	50.09	85.09	1.000
Esquizofrenia	443.45	.000	.816	49.15	88.8	1.000
Hipomanía	132.38	.000	.570	51.26	63.61	1.000
Introversión social	202.16	.000	.669	50.48	71.56	1.000

Nota: G.L.(1;100). M_{ie}= Media de la condición *instrucciones estándar*. M_{is}= Media de la condición *instrucciones de simulación*

530.21; p<.001; eta²= .841, constatamos un efecto significativo modulado por el factor instrucciones. En concreto, las simuladoras pasan de la normalidad (T±50), en la condición respuestas bajo instrucciones estándar (*M*s= 49.91 y 50.09, para Pk y Ps, respectivamente), a una impresión diagnóstica de padecer un TEP (T>70) en la condición de respuestas de simulación (*M*s= 80.39 y 78.63, en Pk y Ps). En resumen, las mujeres de la muestra no padecían este trastorno y fueron capaces de simularlo a la vez que entendieron que está ligado a la huella psíquica producida por malos tratos.

El estudio de casos nos advierte que 85 (84.2%) mujeres pudieron simular en las escalas Pk y Ps que padecían un TEP sin realmente sufrirlo. Si el objeto de contraste fuera un cuadro comórbido compuesto por TEP y Depresión, la capacidad de simulación sería la misma, en tanto que si se coteja un cuadro compuesto por TEP, Depresión e Introversión Social, el potencial de simulación estaría presente en 59 mujeres (58.4%).

Estudio de la simulación en las escalas originales de control de validez del MMPI-2

El modelo criminológico (Bagby, Rogers, Buis, Nicholson, Cameron, Rector, Schuller y Seeman, 1997) y la American Psychiatric Association (2000) en el DSM-IV-TR defienden como estrategia central de simulación la falta de cooperación del sujeto en la evaluación. Ésta se mide a través de la Escala ? (No Respuestas), siendo el punto de corte óptimo 10 no respuestas (Graham, 2006). Los resultados al respecto no invalidan ningún protocolo ni en la condición de respuestas bajo *instrucciones estándar* ni bajo *instrucciones de simulación*. Por tanto, esta medida de la simulación resulta totalmente inefectiva para el caso que nos ocupa.

En la Escala F (Incoherencia), en la que puntuaciones elevadas (T>70) son indicativas de una potencial simulación (Bagby, Rogers, Buis y Kalemba, 1994), observamos la correcta clasificación del 81.2% de las simuladoras, lo que lo convierte en un indicador fiable, $\chi^2(1)= 39.29$; p<.001. Ahora bien, el 18.8% de las participantes sobrepasaría este control de validez, superando el margen de error admisible estadísticamente (<.05), Z(101)= 6.36; p<.001. Por su parte, en la Escala K, en la que puntuaciones T<40 sugieren que el evaluado finge tener serias dificultades emocionales (Graham, 2006), el estudio de casos pone de manifiesto que el 89.1% de las simuladoras son correctamente identificadas por esta escala como tales, de lo que se desprende que es un estimador fiable, $\chi^2(1)= 61.79$; p<.001. Aún así, el 10.9% de las participantes superaría este control de validez, dato que sobrepasa el margen de error estadísticamente admisible (<.05), Z(101)= 2.72; p<.01. Complementariamente, la Escala K informa de disimulación (T≥65, Graham, 2006). En nuestro caso, ninguna participante presentó una puntuación que advirtiera de disimulación. Finalmente, en la Escala L (Mentira), que estima la tendencia a la deseabilidad social en las respuestas, el estudio de casos determinó que ninguna participante alcanzó, en la condición de simulación, una puntuación invalidante (T>70; Graham, 2006). De este modo, una puntuación acorde o en dirección a la deseabilidad social pasa a ser un indicador potente de no simulación.

Configuraciones de las escalas de validez

El *índice F-K* informa, si la diferencia es positiva y mayor o igual a 30 (Rogers, Sewell y Salekin, 1994), de intentos de dar una mala imagen. En la clasificación de casos, el *índice F-K* identifica correctamente al 79.2% de las simuladoras que lo sitúa como un índice fiable de simulación, $\chi^2(1)= 34.46$; p<.001, pero permite un margen de error, el 20.8%, que supera lo estadísticamente admisible, Z(101)= 7.28; p<.001. En suma, el *índice de Gough* es un estimador fiable de simulación de la huella psíquica en casos de violencia de género pero, al mismo tiempo, deja un margen de error significativo al no informar de simulación estando ésta presente. Además, el *índice de Gough* puede dar entrada a otra hipótesis interpretativa: intentos de presentar una buena imagen de sí mismos, cuando es fuertemente negativo (PD<-9; Jiménez y Sánchez, 2003). El estudio de casos no ha registrado ninguno bajo esta contingencia. Por tanto, una puntuación fuertemente negativa se convierte en un indicador robusto de no simulación.

El *perfil en V invertida*, con L y K por debajo de 50 y F por encima de 80, sugiere la posibilidad de exageración de síntomas (Jiménez y Sánchez, 2003). En nuestro caso, esta configuración clasifica correctamente a 43 participantes como simuladoras (42.5%), lo que le confiere un estatus de estimador para la detección de simulación de huella de malos tratos igual al azar, $\chi^2(1)= 2.22$; ns. La medida complementaria del *perfil en V invertida*, el *perfil en V* (L y K >70 y F<50; Graham, 2006), que advierte de disimulación, no clasifica ningún protocolo dentro de la disimulación, esto es, la observación de un *perfil en V* conforma un pronóstico muy fiable de no simulación.

Estudio de las escalas adicionales de control de la validez del MMPI-2

La Escala Fb, en la que puntuaciones T>70 son indicativas de simulación, clasifica correctamente al 70.5% de las participantes, lo que lo convierte en un indicador fiable, $\chi^2(1)= 19.29$; p<.001. No obstante, cabe destacar que, tras un estudio de casos, se observa que las sujetos detectadas por la Escala Fb eran las mismas que las no validadas por la Escala F original. En suma, la Escala Fb, que es la continuación de la F original en las escalas Suplementarias y de Contenido, se solapa con ésta, no presentando valor añadido alguno, por lo que tomar las dos escalas supondría una duplicidad de medidas. Las escalas TRIN y VRIN (PD>13) validan, casi en su totalidad, los protocolos: TRIN invalida 1 protocolo y VRIN 2.

Estudio global de la detección de la simulación a través de los indicadores de validez del MMPI-2

Para que un perfil sea invalidado se requiere que en, al menos, dos de las escalas originales de validez relevantes para la simulación (F y K) o de sus combinaciones (el *índice F-K* y *perfil en V invertida*) los sujetos presenten puntuaciones invalidantes (Graham, 2006). Por ello procedimos a estudiar el efecto acumulativo de estos índices y sus configuraciones. Los resultados muestran que dos o más indicadores coinciden en la calificación como simuladoras en 87 participantes (86.14%), confiriéndole un alto valor diagnóstico de simulación, $\chi^2(1)= 52.76$; p<.001. No obstante, 4 no son detectadas por ningún indicador y 10 sólo por uno. En otras palabras, la falta de concordancia inter-índices abre una región de duda en torno a la calificación como simuladoras de 14 sujetos (13.86%), que es mayor que el margen de error asumible estadísticamente, Z(101)= 4.08; p<.001. Además, estas 14 simularon en la medida directa del daño esperado, el TEP, y 9 también los trastornos secundarios, depresión e introversión social. Ahora

bien, también simularon otros daños no esperados (por ejemplo, desviación psicopática) pero, al mismo tiempo, discriminaron otros correctamente (v.gr., esquizofrenia). En resumen, el análisis de la consistencia interna del perfil clínico observado puede resultar de interés, pero no sería una prueba definitiva.

Estudio de la simulación en una tarea de conocimiento: la entrevista clínico-forense

En las entrevistas simuladas de daño psíquico asociado a un maltrato (tabla 2) se observa que las participantes presentan los siguientes criterios propios de un TEP más allá de lo esperado por azar (p>.05), o sea, constituyen sintomatología accesible: a) respuestas intensas de miedo; b) recuerdos del acontecimiento recurrentes e intrusos que provocan malestar y en los que se incluyen imágenes, pensamientos o percepciones; c) malestar psicológico al exponerse a estímulos internos o externos que simbolizan o recuerdan los malos tratos; d) reducción acusada del interés o la participación en actividades significativas; e) sensación de desapego; f) restricción de la vida efectiva; g) futuro desolador; h) dificultades para conciliar o mantener el sueño; i) dificultades de concentración; y j) malestar clínico significativo o deterioro social, laboral o de otras áreas importantes en la actividad de la persona consecuencia de los malos tratos. En resumen, de los criterios diagnósticos del TEP recogidos en el DSM-IV-TR resultan accesibles a la simulación los Criterios A; dos de los Criterios B (recuerdos recurrentes; malestar psicológico al exponerse a estímulos que recuerdan los malos tratos); cuatro de los Criterios C (reducción del interés en actividades significativas; sensación de desapego; restricción de la vida afectiva; futuro limitado); dos de los Criterios D (dificultades para conciliar y mantener el sueño; dificultades para concentrarse); y el Criterio F. En términos operativos, son accesibles a la simulación los criterios A (los dos síntomas), B (más de un síntoma), C (tres o más síntomas), D (uno o más) y F con lo que, en general, los criterios diagnósticos para la simulación de un TEP son asequibles para las simuladoras de malos tratos.

Tabla 2
Análisis de la sintomatología accesible y no-accesible en el TEP.
Valor de prueba .05

VARIABLE	Proporción observada	Z	p
CRITERIO A			
Temor intenso	.495	20.5	.001
CRITERIO B			
Recuerdos recurrentes	.128	3.59	.001
Sueños recurrentes	.019	-1.42	ns
Flashback	.000	–	–
Malestar psicológico intenso	.093	1.98	.05
Respuestas fisiológicas	.029	-0.96	ns
CRITERIO C			
Evitación de pensamientos	.000	–	–
Evitación de actividades	.079	1.33	ns
Incapacidad recuerdo	.000	–	–
Reducción interés en actividades	.594	20.06	.001
Sensación de desapego	.336	13.17	.001
Restricción vida afectiva	.237	8.61	.001
Futuro limitado	.099	2.25	.05
CRITERIO D			
Dificultades para mantener sueño	.128	3.59	.001
Irritabilidad o ira	.079	1.33	ns
Dificultades de concentración	.148	4.51	.001
Hipervigilancia	.019	-1.42	ns
Respuestas exageradas sobresalto	.029	-0.96	ns
CRITERIO F			
Malestar clínico significativo u otros	.815	35.25	.001

Nota: N= 101. El criterio A1, *haber experimentado una herida grave o un riesgo a la integridad física*, se asume que es el maltrato denunciado

De estos mismos datos (tabla 2) también se desprende que hay una sintomatología no-accesible a la simulación ($p \leq .05$): a) sueños recurrentes sobre los acontecimientos que provocan malestar; b) sensación de que el hecho traumático está ocurriendo; c) respuestas fisiológicas al exponerse a estímulos asociados al trauma; d) esfuerzos para evitar pensamientos, sentimientos o conversaciones sobre el trauma; e) esfuerzos para evitar actividades, lugares o personas que recuerdan el trauma; f) incapacidad para recordar un aspecto importante del trauma; g) irritabilidad o ataques de ira; h) hipervigilancia; e i) respuestas exageradas de sobresalto. En consecuencia, la verificación de estos síntomas en casos forenses ha de entenderse como un signo de realidad.

El estudio de casos clínicos nos advierte que 3 (2.9%) de las participantes fueron capaces de simular un TEP, esto es, las secuelas psíquicas de la victimación de violencia de género. En suma, la accesibilidad de la simulación en la tarea de conocimiento es igual al margen de error estadísticamente aceptable (.05), $Z(101)= 0.95$; ns., y el grado de dificultad de simulación es mayor en la tarea de conocimiento (.029%) que en la de reconocimiento (.842), $\chi^2(1,n= 101)= 76.4$; $p<.001$.

Estudio de las estrategias de simulación en la entrevista clínico-forense

Sólo 3 de las estrategias de simulación resultaron productivas en este estudio. De ellas, las simuladoras sólo acudieron a 2 con una frecuencia significativa (>.05): *síntomas sutiles* (.475) (no informa de síntomas reales sino de problemas cotidianos, $Z(101)= 21.88$; $p<.001$, y *severidad de síntomas* (.217) (atribución sistemática a los síntomas padecidos de una severidad extrema), $Z(101)= 7.69$ $p<.001$. Adicionalmente, hallamos inconsistencia entre la sintomatología manifestada y la observada (.039), pero su presencia es marginal (=.05), $Z(101)= -0.5$; ns. Acumulativamente, las participantes utilizaron alguna de las estrategias habituales de simulación en el 57.4% de los protocolos, con lo que estas herramientas para la detección de la simulación no son más fiables que el azar ($p= .5$), $\chi^2(1,n= 101)= 2.22$; ns. Más en concreto, en las tres simuladoras efectivas en la tarea de conocimiento no se encontró que recurrieran a alguna de las estrategias de simulación.

Estudio global de la detección de la simulación

Para que judicialmente una prueba se tome como válida precisa de su demostración y de consistencia en la medida. Partiendo de la base de que sólo tres participantes lograron simular TEP en la *entrevista clínico-forense*, encontramos que éstas también simularon las medidas directas e indirectas en la medida psicométrica. Por su lado, las medidas de control de la validez de la entrevista no las detectaron como simuladoras pero sí las medidas de control de la validez del MMPI. En concreto, no menos de 3 medidas de con-

trol informaron de posible simulación, lo que es un indicio fiable de la misma. A su vez, un análisis de la consistencia inter-medidas, más allá de la medida de la huella psíquica de los malos tratos, advierte de otras patologías en la medida psicométrica no presentes en la entrevista. Sucintamente, se observa que en la medida psicométrica, estas participantes asumieron todo tipo de sintomatología a la vez que atribuyeron una gran severidad a la misma, en tanto en la tarea de conocimiento no dieron muestras de estos otros cuadros. En consecuencia, la falta de consistencia inter-medidas puede tomarse como un criterio en el estudio de la simulación.

Discusión y conclusiones

Los anteriores resultados merecen una serie de matizaciones sobre el alcance de los mismos previamente a la discusión. Primero, las participantes cumplimentaron una tarea de simulación, con lo que la implicación en la tarea es presumiblemente menor que en casos auténticos y los resultados han de generalizarse a contextos reales con ciertas salvedades (Konecni y Ebbesen, 1992). Segundo, los resultados de este estudio no pueden generalizarse a otros instrumentos de medida. Tercero, los resultados no pueden generalizarse a otras tipologías de delitos que produzcan otras secuelas psíquicas. Cuarto, nuestros sujetos son sólo simuladoras, con lo que no disponemos de verdaderas víctimas para contrastarlas con las simuladoras y acercarnos a los *falsos positivos*.

Con estas salvaguardas en mente, de los anteriores resultados se pueden derivar las siguientes conclusiones:

a) Ninguno de los procedimientos de control de la simulación sometidos a prueba es por sí mismo totalmente efectivo y válido para la detección de la simulación.
b) Las participantes fueron capaces de simular en el MMPI-2 las secuelas psíquicas asociadas a la violencia de género, tanto el daño primario (84.2%), el TEP, como un estado comórbido (58.4%) compuesto por TEP, depresión e introversión social. Esta facilidad para la simulación la atribuimos a que en los instrumentos psicométricos, por su propia naturaleza, los sujetos son sometidos a una tarea de reconocimiento de síntomas. A su vez, las medidas de control de la validez darían paso a un 13.86% de falsos negativos.
c) En la *entrevista clínico-forense,* que presupone una tarea de conocimiento de síntomas por parte del sujeto, la capacidad de simulación de las participantes se redujo al 2.9%. A su vez, en este tipo de tarea, las medidas de control de la validez de la misma, el análisis de las estrategias de simulación, tampoco detectó éstas entre las que lograron simular efectivamente. Aun así, para la práctica forense, con diseños de n= 1, es preciso controlar totalmente la capacidad de simulación, por lo que este procedimiento no es suficiente por sí solo para evaluar la simulación.
d) Dado que ni la aproximación psicométrica, ni la basada en la entrevista son prueba válida y suficiente por sí sola, es necesario adoptar una perspectiva multimétodo. Los anteriores resultados permiten dibujar la existencia de dos agrupaciones de criterios en torno a la simulación: positivos y negativos. Son criterios positivos todos aquellos indicadores de simulación (v.gr., falta de consistencia inter-medidas, puntuaciones invalidantes en las escalas F [T>70] y K [T<50], en el índice F-K (T>30), un *perfil en V invertida*; detección de estrategias de simulación en la *entrevista clínico-forense*), mientras que son negativos los que se vinculan directamente con la no simulación (una puntuación ?>10, L>70, K>70, un diferencial en puntuaciones directas <-9 en el *índice de Gough*, la observación de un *perfil en V*, y la obtención de sintomatología no-accesible).
e) Acorde a los resultados de este estudio, para invalidar un protocolo por simulación es imprescindible que, al menos, converjan 3 criterios positivos de simulación.
f) Finalmente, para el control de los falsos positivos (auténticos enfermos clasificados como simuladores), el perito psicólogo deberá proceder con el Modelo de Decisión Clínica (Cunnien, 1997) para la detección de falsos positivos; el contraste de criterios positivos; y con el registro de sintomatología no-accesible en la entrevista clínico-forense.

Agradecimientos

Esta investigación fue financiada por Ministerio de Educación y Ciencia, Dirección General de Investigación en el Proyecto de Referencia: SEJ2007-63436/PSIC).

Referencias

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

Arce, R., y Fariña, F. (2001). *Construcción y validación de un procedimiento basado en una tarea de conocimiento para la medida de la huella psíquica en víctimas de delitos: la entrevista clínico-forense*. Manuscrito inédito, Universidad de Santiago de Compostela.

Arce, R., Fariña, F., Carballal, A., y Novo, M. (2006). Evaluación del daño moral en accidentes de tráfico: desarrollo y validación de un protocolo para la detección de la simulación. *Psicothema, 18*, 278-283.

Bagby, R.M., Rogers, R., Buis, T., Nicholson, R.A., Cameron, S.L., Rector, N.A., Schuller, D.R., y Seeman, M.V. (1997). Detecting feigned depression and schizophrenia on the MMPI-2. *Journal of Personality Assessment, 68*, 650-664.

Cunnien, A.J. (1997). Psychiatric and medical syndromes associated with deception. En R. Rogers (Ed.): *Clinical assessment of malingering and deception* (pp. 23-46). New York: Guilford Press.

Duckworth, J.C., y Anderson, W.P. (1995). *MMPI and MMPI-2: Interpretation manual for counsellors and clinicians*. Bristol: Accelerated Development.

Echeburúa, E., Corral, P., y Amor, P.J. (2002). Evaluación del daño psicológico en víctimas de crímenes violentos. *Psicothema, 14*, 139-146.

Edleson, J.L. (1999). Children's witnessing of adult domestic violence. *Journal of Interpersonal Violence, 14*, 839-870.

Freyd, J. (1996). *Betrayal trauma: The logic of forgetting childhood abuse*. Cambridge, Mass.: Harvard University Press.

Gothard, S., Rogers, R., y Sewell, K.W. (1995). Feigning incompetency to stand trial. An investigation of the Georgia court competency test. *Law and Human Behavior, 19*(4), 363-373.

Graham, J.R. (2006). *MMPI-2: Assessing personality and psychopathology* (4ª ed.). NewYork: Oxford University Press.

Hathaway, S.R., y McKinley, J.C. (1999). *Inventario multifásico de personalidad de Minnesota-2. Manual*. Madrid: TEA.

Jiménez, F., y Sánchez, G. (2003). *Evaluación psicológica forense. Contribución de las técnicas de Minnesota y Millon*. Salamanca: Amarú Ediciones.

Kessler, R.C., Sonnega, A., Hughes, M., y Nelson, C.B. (1995). Posttraumatic stress disorder in the national comorbidity survey. *Archives of General Psychiatry*, *52*, 1048-1060.

Konecni, V.J., y Ebbesen, E.B. (1992). Methodological issues in research on legal decision-making, with special reference to experimental simulations. En F. Lösel, D. Bender y T. Bliesener (Eds.): *Psychology and law. International perspectives* (pp. 413-423). Berlín: Walter de Gruyter.

Mikkelsen, E., Gutheil, T., y Emens, M. (1992). False sexual abuse allegations by children and adolescents: Contextual factors and clinical subtypes. *American Journal of Psychotherapy, 55*, 556-570.

Miller, H., y Cartlidge, N. (1972). Simulation and malingering after injuries to the brain and spinal cord. *Lancet, 1*, 580-585.

Milner, J.S., y Crouch, J.L. (2004). El perfil del niño víctima de violencia. En J. Sanmartín (Coord.): *El laberinto de la violencia. Causas, tipos y efectos* (pp. 195-203). Barcelona: Ariel.

Nicholson, R.A., Mouton, G.J., Bagby, R.M., Buis, T., Peterson, S.A., y Buigas, R.A. (1997). Utility of MMPI-2 indicators of response distortion: Receiver operating characteristic analysis. *Psychological Assessment, 9*, 471-479.

O'Donnell, M.L., Creamer, M., Bryant, R.A., Schnyder, U., y Shalev, A. (2006). Posttraumatic disorders following injury: Assessment and other methodological considerations. En G. Young, A.W. Kane y K. Nicholson (Eds.): *Psychological knowledge in courts. PTSD, pain and TBI* (pp. 70-84). New York: Springer.

Rogers, R. (1997). Current status of clinical methods. En R. Rogers (Ed.): *Clinical assessment of malingering and deception* (pp. 373-397). New York: Guilford Press.

Rogers, R., Sewell, K.W., y Salekin, R.T. (1994). A meta-analysis of malingering on the MMPI-2. *Assessment, 1*, 227-237.

Sarasua, B., Zubizarreta, I., Echeburúa, E., y Corral, P. (2007). Perfil psicopatológico diferencial de las víctimas de violencia de pareja en función de la edad. *Psicothema, 19,* 459-466.

Taylor, S., y Koch, W.J. (1995). Anxiety disorders due to motor vehicle accidents: Nature and treatment. *Clinical Psychology Review, 15*, 721-738.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.

United Nations (1988). *Committee on crime prevention and control. Report on the tenth session*. Viena: United Nations Publications.

Wicker, A.W. (1975). An application of a multipletrait-multimethod logic to the reliability of observational records. *Personality and Social Psychology Bulletin, 4*, 575-579.

Weick, K.E. (1985). Systematic observational methods. En G. Lindzey y E. Aronson (Eds.): *The handbook of social psychology bulletin* (vol. 1, pp. 567-634). Hillsdale, N.J.: LEA.