Arce, R., y Fariña, F. (2015). Evaluación psicológico-forense de la credibilidad y daño psíquico mediante el Sistema de Evaluación Global. En P. Rivas y G.L. Barrios (Dirs.), *Violencia de género: Perspectiva multidisciplinar y práctica forense* (pp. 411-441). Navarra: Thomson Aranzadi.

Evaluación psicológica de la realidad del testimonio y de la secuela psicológica: Estado del arte y prueba forense

Autores: Ramón Arce(1) y Francisca Fariña(2)

Afiliación: 1) Departamento de Psicología Organizacional, Jurídica-Forense y Metodología de las CC. de Comportamiento, Universidad de Santiago de Compostela (España).
2) Facultad de Ciencias de la Educación y el Deporte, Universidad de Vigo (España).

Introducción

Desde la entrada en vigor de la Ley 1/2004, de 28 de diciembre, de Medidas de Protección Integral contra la Violencia de Género, cuyo objeto era erradicar la violencia contra la mujer, las cifras de denuncias, con datos de 2007 a 2013[1], se han mantenido más o menos estables, M = 132214.6, IC95%[127747.4–136681.8]. Estos parámetros nos pueden permitir, en un futuro, conocer la evolución real del impacto de la Ley ya que los datos anteriores a la Ley, así como iniciales de la misma (i.e., 2005 y 2006), no están disponibles o no son comparables porque las metodologías y rigor de cómputo no son equiparables. No obstante, de los mismos sí se puede obtener un indicador de la relación entre denuncias y condenas, que no es viable año a año ya que los procesos pueden alargarse más del año natural. Dicha relación, medida en *Odds Ratio*, M_{OR} = 4.62, IC95%[3.93–5.31], es realmente muy preocupante (i.e., tamaño del efecto grande, que se corresponde con 1.29 desviaciones estándar), esto es, sólo aproximadamente 1 de cerca de cada 5 mujeres (4.62) que denuncian ser víctimas de violencia de género obtiene respaldo en sentencia judicial a dicha denuncia (±22%), lo que nos obliga a buscar las causas de ésta. Aunque inicialmente se hubiera puesto el interés en que podría ser la absolución la causa de esta disfunción, sólo alrededor de 1 de cada 7

[1] Todas estas estadísticas se computaron a partir de las estadísticas del estadísticas del Observatorio contra la Violencia Doméstica y de Género del Consejo General del Poder Judicial, el Ministerio de Igualdad y el Ministerio de sanidad, Servicios Sociales e Igualdad.

denuncias, M_{OR} = 6.89, IC95%[6.22–7.56], termina con la absolución del denunciado (un tamaño del efecto grande, que se corresponde con 1.70 desviaciones estándar). Es decir, la absolución, con ser una causa importante (±14%), no es la principal de la baja tasa de condenas. En las mismas estadísticas se da cuenta de otras dos causas importantes: la renuncia a la denuncia y el sobreseimiento. Las renuncias, como una causa de extrema transcendencia desde el punto de vista judicial y de las políticas de igualdad, y a pesar de que no es admisible una tasa que no sea cero, numéricamente puede llegar a explicar en torno al 12% de casos denuncias no enjuiciadas (la relación no es directa al poder continuar la Fiscalía con el caso por lo que podría concluir el caso en sobreseimiento o absolución fundada mayormente en falta de persistencia en la incriminación, o incluso, aunque en menor medida, en condena). El sobreseimiento de la causa sí explica, en gran medida, la baja relación entre denuncia, por una parte, y enjuiciamiento y condena, por otra. En torno al 40% de las denuncias terminan en el sobreseimiento de la causa. Si bien los análisis de las decisiones judiciales apuntan numerosas motivaciones judiciales para el desistimiento del caso o la absolución (e. g., falta de persistencia en la incriminación, error en la apreciación de la prueba, infracción por aplicación indebida del artículo 153 o del 172.3), las dificultades dimanantes de la carga de la prueba parecen ser las más importantes (Novo y Seijo, 2010; Arce, Alonso, y Novo, 2010). Como quiera que la práctica totalidad de los sobreseimientos se producen en los Juzgados de Violencia contra la Mujer (>95%), y las estadísticas también informan de los medios de prueba practicados en el momento de la denuncia, procede un análisis y estudio de la potencia de los mismos. Cinco son los medios de prueba que se relacionan al respeto: denuncia de la víctima, denuncia de familiares, denuncia por servicios asistenciales o terceros, atestados policiales y parte de lesiones. El testimonio de la víctima no acostumbra a ser prueba suficiente porque choca fácilmente con el criterio jurisprudencial de incredibilidad subjetiva en cuanto que la víctima puede mostrarse parte en la causa. Para dotarlo de valor de prueba, éste ha de estar rodeado de ciertas corroboraciones periféricas de carácter objetivo que le doten de aptitud probatoria (criterio jurisprudencial de verisimilitud). A este respecto Novo y Seijo (2010) y Arce et al. (2010) hallaron, tras análisis de contenido de sentencias judiciales, que las corroboraciones periféricas que se presentaban a jueces y tribunales para dotar de valor de prueba al testimonio de la víctima eran fundamentalmente el atestado policial, los partes de lesiones, otros testimonios y la prueba pericial psicológica sobre la realidad (i.e., credibilidad) del testimonio o de secuelas

psicológicas. El atestado policial, con ser la más frecuente registrada, no resulta ser prueba periférica suficiente para dotar de valor de prueba al testimonio de la denunciante porque se asociaba por un igual a sentencias condenatorias y absolutorias. Los testimonios de terceros, menos frecuentes, tampoco resultaron clarificadores a excepción del de los hijos menores pero eran poco usuales al tiempo que implican un coste por sobrecarga para los menores difícilmente aconsejable. Así pues las dos grandes pruebas de carga son las evaluaciones médico- y psicológico-forenses. Las evaluaciones médicos forenses sobre las lesiones resultaban en una corroboración periférica suficiente para avalar el testimonio de la víctima en el 72.7% de los casos; el informe psicológico forense sobre la credibilidad del testimonio en el 93.3%; y aquel sobre las secuelas psicológicas en el 80%. En los casos de violencia de género, hallamos que los atestados policiales son prueba que acompaña a la denuncia en casi la totalidad de los casos denunciados, M_{OR} = 1.28, IC95%[1.26–1.30]; los partes médico-forenses de lesiones, M_{OR} = 8.83, IC95%[8.56–9.10], en aproximadamente el 11%; en tanto que los informes psicológico-forenses no son prueba habitual, de modo que ni se aportan estadísticas al respecto. En suma, los sobreseimientos recaen fundamentalmente en la carencia de pruebas que doten de valor probatorio al testimonio de la víctima, esto es, informe psicológico forense de credibilidad del testimonio, y de prueba de daño (los informes de lesiones sólo de practican en el 11% de los casos y, de ningún modo, tienen por objeto demostrar la victimización de violencia psicológica), específicamente la secuela psicológica que, como se puede inferir, no forma parte de las pruebas practicadas con asiduidad en acompañamiento de la denuncia. Aunque pudiera creerse que en otros momentos procesales se corregirían estas limitaciones en la práctica de la prueba, estudios de archivo con sentencias judiciales (Arce et al., 2010) hallaron que sólo en el 20% de los casos enjuiciados se habían practicado evaluaciones psicológico-forenses sobre la credibilidad del testimonio o la secuela psicológica. Abundando aún más, Novo y Seijo (2010) informaron, tras un análisis de sentencias judiciales, que en el 32% de los casos la única prueba de cargo era la declaración de la víctima. Como consecuencia de todo ello se perdería, por una parte, la posibilidad de dotar de valor de prueba al testimonio de la víctima a través de la corroboración periférica del mismo mediante la prueba de credibilidad del testimonio, y, por otra, la demostración de la victimización en términos de la secuela psicológica.

Arce, R., y Fariña, F. (2015). Evaluación psicológico-forense de la credibilidad y daño psíquico mediante el Sistema de Evaluación Global. En P. Rivas y G.L. Barrios (Dirs.), *Violencia de género: Perspectiva multidisciplinar y práctica forense* (pp. 411-441). Navarra: Thomson Aranzadi.

Tabla 1. Tamaños del efecto en estudios que comparan la discriminación entre declaraciones verdaderas e inventadas con los criterios del CBCA.

Estudio	**Declaraciones**	***d***
Arce, Fariña y Vilariño (2010)	50	1.39
Bensi, Gambetti, Nori, y Giusberti (2009)	80	0.98
Blandón-Gitlin, Pezdek, Lindsay y Hagen (2009)[a]	56	0.75
Blandón-Gitlin, Pezdek, Lindsay y Hagen (2009)[b]	102	0.53
Blandón-Gitlin, Pezdek, Lindsay y Hagen (2009)[c]	102	0.28
Bogaard, Meijer y Vrij (2014)	64	0.48
Gödert, Gamer, Rill y Vossel (2005)[d]	64	0.59
Gödert, Gamer, Rill y Vossel (2005)[e]	64	0.83
Lee, Klaver y Hart (2008)	90	0.31
Merckelbach (2004)[f]	40	-0.23
Merckelbach (2004)[g]	36	-0.28
Schelleman-Offermans y Merckelbach (2010)	120	0.77
Vrij y Mann (2006)[h]	80	0.47
Vrij, Akehurst, Soukara y Bull (2004)	52	0.77
Vrij, Edward, Roberts y Bull (2000)	73	1.15
Vrij, Kneller y Mann (2000)[i]	30	0.22
Vrij, Kneller y Mann (2000)[j]	30	1.15
Vrij, Mann, Kristen y Fisher (2007)[k]	40	0.54
Willén y Strömwall (2011)	60	0.45

Nota. k = 19; se eliminaron los outliers positivos, datos repetidos, análisis de contenido no fiables y declaraciones no obtenidas con entrevistas basadas en el recuerdo libre y reinstauración de contextos. [a]experimento 1; [b]memoria real vs. sugerida; [c]memoria real vs. fabricada; [d]memoria falsa (agresor) vs. verdadera (testigo); [e]memoria falsa (sujeto no-implicado) vs. verdadera (testigo); [f]sujetos con alta propensión a la fantasía; [g]sujetos con baja propensión a la fantasía; [h]fase 1; [i]mentirosos informados sobre el CBCA; [j]mentirosos no informados; [k]info-gathering interview.

Para dotar de valor probatorio al testimonio del denunciante, los sistemas de análisis de contenido basado en criterios de las declaraciones son los que se han mostrado realmente eficaces. Éstos se basan en la *hipótesis Undeutsch*, que establece

que la memoria de una experiencia vivida difiere en contenido y calidad de una memoria de lo no experimentado, bien sea inventado o imaginado (Undeutsch, 1967, 1989). Dicha hipótesis ha sido confirmada de forma sistemática por la literatura científica. Así, Vrij (2005), en una revisión de la evidencia empírica, halló que el 92% de los estudios confirmaban esta hipótesis con las categorías de realidad del CBCA (Criteria Based Content Analysis), el sistema de análisis de contenido de las declaraciones de referencia basado en esta hipótesis (Steller y Köhnken, 1989). En una revisión posterior, el mismo Vrij (2008) no encontró ningún estudio de campo en el que criterio alguno del CBCA estuviera más presente en declaraciones inventadas que en reales. Sin embargo, esta hipótesis y sistemas categoriales de contenido se crearon para casos de abusos sexuales y poblaciones de menores. Por ello, lo primero procede llevar a cabo una revisión meta-analítica de los estudios realizados con los criterios del CBCA con poblaciones de adultos (ver en la Tabla 1 el listado de estudios y tamaños del efecto) a fin de conocer si dicha hipótesis es igualmente válida o no para esta población. Los resultados, que se resumen en la Tabla 2, prestan un apoyo a la validez de la hipótesis y criterios del CBCA para poblaciones de adultos: tamaño del efecto significativo, 95% CI_d[0.48/0.70], positivo y de magnitud moderada (δ = 0.67). Como quiera que el tipo de casuísticas es muy variable, los resultados avalan la pertinencia de la hipótesis y criterios de realidad, además de para poblaciones de adultos, para todo tipo de casuísticas, 90%CI_δ[0.32/1.02].

Tabla 2. Estadísticos del meta-análisis de la hipótesis *Undeutsch* con los criterios de realidad del CBCA en población adulta.

TOTAL CBCA	***k***	***N***	***d***	SD_d	**δ**	SD_δ	**%VE**	**95% CI_d**	**90% CI_δ**
	19	1305	0.59	0.351	0.67	0.276	52	0.48/0.70	0.32/1.02

Sin embargo, esto no significa que los criterios de realidad del CBCA sean muy potentes en casos de violencia de género y, lo que no es objeto de la investigación científica, que se pueda fijar un criterio de decisión forense. Así, Arce et al. (2010) hallaron que los criterios del CBCA discriminaban significativamente pero con poca potencia entre casos reales e inventados de violencia de género de modo, y que se precisarían más 14 o más criterios del CBCA para garantizar que una declaración estaba basada en hechos vividos (las declaraciones inventadas contenían hasta 13 criterios por

lo que el criterio de decisión forense, en el que no es admisible clasificar un caso falso como verdadero, requeriría de, al menos, 14 criterios por lo que el 44% de las declaraciones verdaderas, al contener menos de 14 criterios de realidad, no podrían ser clasificadas como tales). Sin embargo, la generalización de los criterios de realidad del CBCA para casos de violencia de género es la práctica común en España (Novo y Seijo, 2010). En concreto, todos aquellos casos en los que la decisión forense esté sustentada en la presencia de menos de 14 criterios de realidad no cumplirían con el estándar judicial requerido a la prueba forense. Por el contario si se tomara como *gold standard*[2] la presencia de más de 7 criterios de realidad en la declaración, el 80% de las declaraciones inventadas serían clasificadas como reales. En suma, la hipótesis Undeutsch es válida para casos de violencia de género, pero no así la extensión de la robustez de los criterios y los estándares de decisión de otros contextos.

Para dar una respuesta a este estado de la cuestión, creamos (Arce y Fariña, 2009) creamos un sistema categorial metódico (i.e., exclusión mutua, homogeneidad, objetividad, adecuación o pertinencia, exhaustividad y productividad) proveniente de la combinación de categorías del CBCA con otras de los sistemas SRA (Undeutsch, 1967, 1988) y Reality Monitoring (criterios ampliados por Sporer, 1997) (objetividad y objetividad), de una redefinición de las categorías cuando fue necesario (exclusión mutua y objetividad), y de la concreción de nuevas categorías por medio de un procedimiento por aproximaciones sucesivas (exhaustividad y pertinencia) (Bardin, 1996). El sistema resultante puede verse en la Tabla 3. Hasta la categoría 24 son categorías universales; esto es, aplicables a todo tipo de delitos, en tanto de la 25 a la 31 sólo son específicas de casos de violencia de género. Sometido el sistema general (SEG-G) a prueba se observó que todas las categorías fueron productivas; se identificaron aquellas que discriminaban significativamente entre declaraciones reales e inventadas en casos de violencia de género (objetividad) (ver Tabla 3) y se halló que el sistema era fiable (homogeneidad), $\alpha = .789$. El sistema categorial compuesto por las categorías productivas que no se registraron entre las declaraciones inventadas y las que discriminaban significativamente entre declaraciones reales e inventadas, denominado sistema categorial ajustado (SEG-A), también se mostró consistente internamente

[2] Aunque no se ha formulado un criterio de decisión estricto, dejando recaer la decisión en un modelo de decisión clínico (Steller, 1989; Köhnken, 2004), se ha apuntado que las declaraciones verdaderas contienen más de 7 criterios (Steller, 1989; Vrij y Akehurst, 1998).

(homogeneidad), α = ,856). Dado que en las declaraciones inventadas de violencia de género se han llegado a registrar hasta 13 criterios del SEG-G y 5 del SEG-A (Arce, Fariña, y Vivero, 2007; Arce, Fariña, Novo, y Vilariño, 2013; Vilariño, Novo, y Seijo, 2011), el estándar de decisión forense (no se admiten los falsos positivos, esto es, la identificación de un caso inventado como real) ha de estar sustentado en la presencia en la declaración de más de 13 y 5 criterios para el SEG- y el SEG-A, respectivamente, para sostener que la memoria del evento está sustentada en eventos vividos. Los falsos negativos esperados para el SEG-G y el SEG-A son del 20% y el 8%, por lo que el SEG-A es preferible por ser más eficaz.

Tabla 3. Criterios de Realidad del SEG.

CARACTERÍSTICAS GENERALES. Los criterios aquí englobados se refieren a la declaración tomada en su totalidad.
1.- Estructura lógica (incluye la reconstrucción de la historia y el realismo del RM, esto es, plausibilidad de reconstrucción del evento, realismo y sentido de la historia).*
2.- Elaboración inestructurada.*
3.- Cantidad de detalles.*
4.- Concreción (SRA, concreción frente a estereotipos, incluye claridad del RM, esto es, viveza en vez de vaguedad)*.
CONTENIDOS ESPECÍFICOS. Se evalúan partes específicas del testimonio referidas a la presencia o fuerza de ciertos tipos de descripciones.
5.- Engranaje contextual (incluye los criterios información espacial e información temporal del RM).
6.- Descripción de interacciones.
7.- Reproducción de conversaciones.
8.- Complicaciones inesperadas durante el incidente*.
9.- Información perceptual (RM, esto es, información sensorial tal como sonidos, gustos o detalles visuales).
10.- Operaciones cognitivas (RM, codificación inversa)
PECULIARIDADES DEL CONTENIDO. Se incluyen aquí aquellas características de una declaración que aumentan su concreción o viveza.
11.- Detalles inusuales.
12.- Detalles superfluos.

13.- Incomprensión de detalles relatados con precisión.
14.- Originalidad de las expresiones (SRA)*.
15.- Asociaciones externas relacionadas (se añade la información no relacionada directamente con los hechos, pero sí con el nodo de memoria recuperado).*
16.- Relatos del estado mental subjetivo (incluye el criterio afecto del RM).*
17.- Atribución del estado mental del autor del delito.
CONTENIDOS REFERENTES A LA MOTIVACIÓN. Estos criterios desvelan la motivación del testigo para hacer la declaración.
18.- Correcciones espontáneas (se añade especificaciones y complementaciones de la información del SRA)*.
19.- Admisión de falta de memoria.*
20.- Plantear dudas sobre el propio testimonio.*
21.- Auto-desaprobación.*
22.- Perdón al autor del delito.
ELEMENTOS ESPECÍFICOS DE LA AGRESIÓN. Elementos del testimonio que no se relacionan con la viveza general de la declaración, sino con el delito.
23.- Detalles característicos de la agresión (un ejemplo en casos de violencia de género sería que no es consciente de determinados tipos de agresiones, tal como las sexuales, pero describe relaciones sin consentimiento).+
DAÑO PSICOLÓGICO DERIVADO DE LA ACCIÓN OBJETO DE JUICIO.
24.- Síntomas clínicos sutiles que forman parte de la huella psíquica (ver en Arce et al., 2009 la relación de síntomas sutiles para casos de violencia de género. Para otras casuísticas el relatorio difiere).*
CRITERIOS ESPECÍFICOS DE VIOLENCIA DE GÉNERO
25.- Justificación de la agresión (provocadora).
26.- Intento de ocultar hechos que agravan la agresión.+
27.- Comportamiento protector del agresor (que no vaya a la cárcel, que no le pase nada).+
28.- Exculpación del agresor (no me quería hacer daño, me pegaba para que me relajara, autoincriminación para exculpar al agresor).+
29.- Medida educativa o represora del agresor (que no beba, métanle miedo, hagan que se porte bien).+
30.- Echar de menos la agresión.+

31.- "Vuelta atrás" en la acusación (falta de cooperación, retirada de la denuncia, obstrucción a la acción judicial).+

Nota. *Criterios que discriminaron significativamente entre declaraciones reales e inventadas en casos de violencia de género; +Estos criterios no se registraron en casos inventados; RM = Criterio del Reality Monitoring; SRA = Criterio del SRA.

Se define la víctima de un delito, en nuestro caso objeto de violencia de género, como aquella mujer que ha sufrido un perjuicio por parte de su cónyuge o persona con la que tenga o haya tenido una relación análoga, entendiendo por ello una lesión física o mental, sufrimiento emocional, pérdida o daño material, o un menoscabo importante en sus derechos (United Nations, 1988). En consecuencia, todo caso de violencia de género implica indefectiblemente la existencia de una víctima y, para tal, la verificación de algún tipo de perjuicio (i.e., victimización): físico, mental, emocional, material o menoscabo en los derechos. La lesión mental y el sufrimiento emocional conforman la secuela o huella psicológica. El siguiente dilema a resolver es la identificación del daño psicológico consecuencia de la violencia de género. La literatura forense ha establecido el Trastorno de Estrés Postraumático (TEP) como la huella psicológica al comprobarse sistemáticamente que este trastorno estaba ligado sistemáticamente con situaciones traumáticas de muy diversa índole (Chivers-Wilson, 2006; Orcutt, Ericsson, y Wolfe, 2004). La concreción de un evento traumático, real o amenaza, estresante (e.g., agresiones físicas, sexuales, secuestros) como la causa del trastorno va a ser la que permita cumplir con el requerimiento judicial de establecer una relación causa-efecto entre la acción a enjuiciar y los daños. Sin embargo, en el TEP las causas están limitadas a estresores altamente estresantes tales como agresiones sexuales, desastres naturales, guerras, secuestros, tortura, o accidentes graves, pero no todos los casos judiciales, específicamente en violencia de género (ténganse en mente, por ejemplo, la violencia psicológica), son consecuencia de un acontecimiento extremadamente estresante. Para estos casos la huella psicológica es el Trastorno Adaptativo que comparte los criterios diagnósticos del TEP con la salvedad de que el estresor es de menor intensidad, esto es, no cumple con el Criterio A del TEP. La huella psicológica suele darse de modo comórbido con otros trastornos que varían según el tipo de acontecimiento y que, en el caso de la violencia de género, se relaciona con depresión, inadaptación social, ansiedad y disfunciones sexuales (Bargai, Ben-Shakhar, y Shalev, 2007). Ahora bien, estos trastornos secundarios no se pueden considerar, en ausencia de un TEP, como una secuela del hecho traumático (O'Donnell, Creamer, Bryant,

Schnyder, y Shalev, 2006). Asimismo, la presencia de un subsíndrome del TEP no es prueba suficiente de secuela psicológica. La prevalencia de TEP o Trastorno Adaptativo entre las víctimas de agresiones de género oscila entre el 45 y el 60% (Vilariño, Fariña, y Arce, 2009), pero se espera que sea mayor entre las denunciantes porque el daño psicológico facilita la denuncia y su ausencia la inhibe. Esta patología cursa de un modo que acarrea implicaciones forenses muy importantes: los síntomas no aparecen de forma contigua a la vivencia del evento sino que se van desarrollando poco a poco por lo que el cuadro raramente está antes de los 15 días; la mayoría de los diagnósticos se producen a los 3 meses del evento; 1/3 de los diagnosticados a los 12 meses, no lo padecían a los 3 (O'Donnell, Creamer, Pattison, y Atkin, 2004). Sucintamente, esta prueba no es aplicable para juicios rápidos, y de la no verificación de secuela no se puede concluir que no lo produjera.

Adicionalmente, en la evaluación forense la evaluación de la secuela, TEP o Trastorno Adaptativo, no es suficiente, siendo obligado *siempre* un diagnóstico diferencial de simulación (DSM-IV; American Psychiatric Association, 2002). Por su parte, en la evaluación sanitaria, el clínico, psicólogo o psiquiatra, ha de diagnosticar, si se cumplen los criterios, otros trastornos, en vez de (o en adición a) TEP (DSM-V; American Psychiatric Association, 2013), siendo la tasa de comorbilidad del 92% (Brown, Campbell, Lehman, Grisham, y Mancill, 2001). Esto implica que no tiene por qué haber correspondencia entre el diagnóstico sanitario y la evaluación forense; y que si se cumplen los criterios de otro trastorno, el sanitario no debería diagnosticar únicamente un TEP, es decir, su diagnóstico no es primario en la evaluación sanitaria. Además, el diagnóstico sanitario de TEP no se puede trasvasar al campo forense, pues en ese contexto no se sospecha simulación. De hecho, la evaluación clínica tradicional no informa sistemáticamente de simulación porque no la sospecha; por los efectos negativos de un error en este diagnóstico; y porque la efectividad de los criterios que maneja para el diagnóstico diferencial de simulación es prácticamente nula (Rogers, 2008; Rosenfeld, Green, Pivovarova, Dole, y Zapf, 2010).

Para este doble objetivo, el diagnóstico clínico y control de la simulación, no es efectiva la evaluación clínica ordinaria. La literatura ha informado que los métodos tradicionales de evaluación de las simulación (p.ej., instrumentación psicométrica con escalas de control de validez; entrevista estructurada, checklist) no son totalmente efectivos en la detección de la simulación, por lo se ha concluido que la detección de la simulación precisa de una aproximación multimétodo (Arbisi, 2005). En esta línea,

nosotros (Arce, Carballal, Fariña, y Seijo, 2004; Arce, Fariña, Carballal, y Novo, 2009) creamos y validamos científica (i.e., sustentada en evidencia empírica) y legalmente (cumple los criterios *Daubert*; Daubert v. Merrell Dow Pharmaceuticals, 1993) una técnica forense basada en un procedimiento multimétodo que se ha mostrado efectivo el control de la simulación en muy diversos contextos (p.ej., violencia de género, enajenación mental, daño moral, acoso moral, agresiones sexuales). Dicha técnica se basa en la combinación de dos métodos distintos de obtención de la información: Una tarea de reconocimiento y una tarea de conocimiento. En la tarea de conocimiento, los sujetos son sometidos a una *entrevista clínico forense* (Arce y Fariña, 2001), o sea, a una entrevista de orden clínico en formato de discurso libre, cuyo procedimiento desgranaremos posteriormente. La entrevista grabada en vídeo es sometida a análisis de contenido, siendo las categorías de análisis los criterios diagnósticos del DSM-V del TEP y cualquier otro síntoma clínico que se detecte en la misma, especialmente de aquellos trastornos y síntomas asociados al TEP (American Psychiatric Association, 2013). El control de la fiabilidad de la entrevista, esto es, la consistencia interna, se aborda a través de un análisis de contenido dirigido a controlar si el sujeto sigue una o más de las estrategias de simulación en tareas de este tipo. Como instrumento de medida en tarea de reconocimiento, recurrimos al MMPI-2 (Hathaway y McKinley, 1999) dado que es el instrumento más utilizado para la evaluación forense del daño psíquico (Greene, 2011) y cuenta con escalas de control de la validez muy contrastadas en eficacia (Rogers, Sewell, Martin, y Vitacco, 2003). Con estas dos medidas que se han de aplicar por este orden, validamos científicamente (Arce et al., 2009) una técnica de actuación que se ha mostrado efectiva en la discriminación entre daño psicológico real y simulado asociado a la victimización de violencia de género (la técnica puede verse con mayor detalle en Arce y Fariña, 2007). Dicha técnica parte de la existencia de criterios positivos, esto es, contrarios a la simulación, y negativos o, lo que es lo mismo, que informan de potencial simulación. Son *criterios positivos* (contrarios a la simulación) la no evitación de respuestas y la deseabilidad social (un patrón de repuestas inconsistentes, TRIN, VRIN y F-Fb tampoco es característico de la simulación, pero se ha de ligar a una patología). Así, aquellos sujetos que sean evaluados por las escalas de control del MMPI en el sentido de que se abstienen significativamente de dar respuesta (Escala ?) y que tienden a dar respuestas de deseabilidad social (Escalas L y K), están emitiendo respuestas honestas. Si bien, la *no respuesta* se ha interpretado en términos médico legales como un indicador fiable de simulación (v.gr., American Psychiatric

Association, 2013), los datos provenientes de simuladores nunca han mostrado que éstos acudan a tal estrategia. Por su parte, los *criterios negativos* son: a) Los instrumentos de medida de la salud mental (esto es, MMPI-2 y entrevista), no detectan, en protocolos válidos, un daño psicológico compatible con la victimización de violencia de género (en otras palabras, si los instrumentos de medida no detectan daño psíquico, no se puede sostener el mismo); b) las escalas de control de validez del MMPI-2 y sus combinaciones relacionadas con la simulación informan de ésta; c) la detección de alguna estrategia de simulación en la entrevista; y d) ausencia de concordancia inter-medidas. El primer criterio es eliminatorio, o sea, si la enfermedad mental no es mensurable, no se puede sostener la existencia de un daño psíquico consecuencia de violencia de género. Asimismo, la falta de concordancia inter-medidas también lleva a desestimar la evaluación tanto por razones puramente científicas como legales (i.e., falta de persistencia). Los restantes, en sí mismos, no son determinantes, con lo que se requerirá la concurrencia de varios criterios y el estudio de las hipótesis alternativas para concluir con relación a la (in)validez del protocolo obtenido. Para estos últimos indicios de no validez es para los que formulamos el concepto de *invalidez convergente*, requiriéndose tres o más criterios de invalidez para anular los protocolos por no poder descartarse indicios sistemáticos de simulación. Con toda esta evidencia, confeccionamos (Arce y Fariña, 2007) el siguiente protocolo de actuación:

a) Consistencia inter-medidas: Recurso a sistemas de medida complementarios y concordantes, no fácilmente simulables y con control de la validez del protocolo (MMPI-2, SCL-90-R, observación y registro conductual, entrevista clínico-forense) que impliquen tareas distintas, reconocimiento, MMPI-2 y SCL-90-R, y de conocimiento (p.ej., entrevista clínico-forense). Ha de tenerse en mente que la tasa de consistencia no va a ser total (e.g., en las evaluaciones test-retest se admiten como consistentes correlaciones > .70).
b) Análisis de la consistencia interna de las medidas: Escalas de control del MMPI-2 y SCL-90-R, y consistencia interna de la entrevista clínico forense mediante el análisis de contenido y registro de las estrategias de simulación [no cooperación con la evaluación, síntomas sutiles, síntomas improbables/absurdos, síntomas obvios, síntomas (cuasi)raros, combinación de síntomas y patrones espurios de psicopatología, severidad de síntomas, inconsistencia de síntomas, estereotipos erróneos, agrupación indiscriminada de síntomas].

c) Análisis de la consistencia inter-evaluadores. Para el control de los sesgos y errores de la evaluación y del análisis de contenido, éstos son llevados a cabo por dos psicólogos forenses por separado de modo que se puede someter a prueba la consistencia intra- e inter-evaluador, e inter-contextos, si, al menos uno de ellos, tiene experiencia previa compartida y consistente con otros psicólogos forenses.

d) Control de falsos positivos, esto es, enfermos reales identificados como potenciales simuladores, a través de un estudio de los antecedentes e historia general del sujeto; de las hipótesis alternativas en cada indicador de no validez (véase Graham, 2011; Greene, 2011; Hathaway y Mckinley, 1999); del cumplimento de los criterios del *Modelo de Decisión Clínica* para el establecimiento de la simulación (Cunnien, 1997); del hallazgo de sintomatología no accesible[3] a la simulación en la entrevista clínico-forense (Arce et al., 2009), y del estudio de los criterios positivos (contarios a la simulación).

e) Estudio de los antecedentes, pruebas documentales y otros. Afortunadamente, en la realidad contamos con más información acerca de la persona lo que facilita la evaluación. Así, se puede reforzar la evaluación con el estudio de los antecedentes, datos del entorno, un estudio de su comportamiento, de pruebas documentales, otros testimonios, etc.

f) Contraste de la validez discriminante. En otras palabras, tomamos una medida no relacionada con el caso, como la personalidad (en concreto, aplicamos 16PF-5), en la que nos detenemos a observar si se registra una personalidad anómala y, sobre todo, en la manipulación de la imagen, aquiescencia e infrecuencia en las respuestas. De estos indicadores de validez nos servimos para contrastar intentos de distorsión sistemáticos de las respuestas. Adicionalmente, la validez

[3] Los *síntomas poco accesibles* a la simulación de un TEP originado por victimización de violencia de género son (Arce et al., 2009):

a) Sueños de carácter recurrente sobre la violencia sufrida que provocan malestar.
b) Malestar psicológico intenso cuando se expone a estímulos internos o externos que simbolizan o recuerdan algún aspecto de la violencia.
c) Esfuerzos para evitar pensamientos, sentimientos o conversaciones sobre la violencia de género.
d) Evitación de actividades, lugares o personas que provocan el recuerdo de la violencia de género.
e) Incapacidad para recordar un aspecto importante de la violencia de género sufrida.
f) Irritabilidad o ataques de ira.
g) Hipervigilancia.
h) Respuestas exageradas de sobresalto..

discriminante se aborda a través de la presencia (o no-presencia) de cuadros clínicos no relacionados con la victimización de violencia de género.

g) El sistema de evaluación se ciñe a las siguientes categorías:

I. Sobre la secuela psicológica: "Observamos un daño en la salud mental compatible con la huella psíquica consecuencia de la victimización de violencia de género" vs. "No observamos daño en la salud mental compatible con la huella psíquica consecuencia de la victimización de violencia de género".

II. Sobre la validez de las medidas de la secuela psicológica: "Observamos indicios de simulación en el daño psíquico informado de modo que no se puede garantizar la realidad del mismo[4]"; "No hallamos indicios de simulación en el daño psíquico informado" o, en su caso, "indefinido".

El Sistema de Evaluación Global en casos de Violencia de Género

Sobre la base de las técnicas forenses que hemos creado y descritas previamente, hemos concretado un protocolo de ejecución global de las mismas adaptado a la práctica forense y a las demandas judiciales, el Sistema de Evaluación Global (Arce y Fariña, 2006). Los pasos de éste ajustados a casos de violencia de género son los siguientes[5]:

1. Obtención del testimonio

Para la evaluación de la realidad del testimonio y la evaluación de las secuelas psicológicas precisamos es necesario obtener el testimonio de la denunciante por medio de una entrevista narrativa, la *entrevista cognitiva mejorada* (Fisher y Geiselman, 1992) que es la versión forense de la *entrevista cognitiva* (Geiselman et al., 1984). La pertinencia de ésta para la evaluación forense e investigativa no se sustenta únicamente en la superioridad en términos de producción sobre otras modalidades de entrevista, sino también en que la aplicación de la misma no está limitada a la reproducción de un suceso de episodio único (las mujeres a evaluar pueden haber sufrido o presenciado previamente episodios de

[4] Es importante tener presente que el sistema no garantiza que el daño sea realmente falso, sino que no se puede garantizar con total certeza que no sea falso. La conclusión de daño falso trae aparejadas implicaciones legales que no se derivan de la técnica como que el evaluado falseó intencionadamente las respuestas o que no haya otras hipótesis alternativas a la simulación, cuando realmente sí concurren otras.

[5] Los peritos psicólogos pueden solicitar a los autores un modelo de informe para casos de violencia de género basado en el SEG.

violencia familiar o de género), ya que permite extender su utilización para recordar hechos que ocurren o se han observado frecuentemente de una manera similar (Mantwill, Köhnken, y Ascherman, 1995), se puede aplicar en series repetidas (Memon, Wark, Bull, y Köhnken, 1997), esto es, posibilita someter a prueba la persistencia en el testimonio/incriminación; y favorece la manifestación de los criterios de realidad (Vrij, 2005). La entrevista cognitiva comprende cuatro técnicas generales de recuperación de memoria (se limita a facilitar la recuperación de la información sin contaminar la prueba): reinstauración mental de contextos, narrar todo lo recordado y recuerdo de los sucesos en una variedad de órdenes temporales distintos. La entrevista cognitiva incluye técnicas suplementarias, *memory jogs*, para facilitar el recuerdo de información específica, que puede ser crítica para la reconstrucción de los hechos denunciados.

La entrevista cognitiva presenta tres grandes limitaciones que el psicólogo o investigador ha de tener siempre presentes: La complejidad en su aplicación, por lo que requiere de un programa exhaustivo de entrenamiento; precisa de la colaboración del entrevistado; y las destrezas cognitivas que exige al entrevistado. Como quiera que algunas víctimas pueden presentar déficits en las funciones cognoscitivas que afectan a la capacidad para testimoniar, se han concretado dos protocolos de actuación adaptados a estas dos poblaciones, en lo que se denominan *Entrevistas Cognitivas Modificadas*. El desarrollo de la entrevista cognitiva y otras técnicas de entrevista cognitiva modificadas, que supera el alcance de este ensayo, puede verse en Arce y Fariña (2012).

Una última fuente de obtención de información con implicaciones para la evaluación, en este caso, forense, de las víctimas de violencia de género proviene de la valoración dela huella psicológica de la victimización de tal violencia. A este respecto, hemos desarrollado (Arce y Fariña, 2001) y validado científicamente una técnica de entrevista, incluido el contexto de la evaluación forense de casos de víctimas de violencia de género (Arce et al., 2009; Vilariño, Arce, y Fariña, 2013), la *entrevista clínico forense*, que es una entrevista de orden clínico narrativo en formato de discurso libre, esto es, una *entrevista cognitiva modificada* para la evaluación clínico-forense. Dicha modificación consiste en solicitar a las víctimas de violencia de género que relaten, en contraste con el momento anterior al sufrimiento de la victimización o posterior a la presentación de la denuncia, los síntomas que advirtieron en sí mismas, los cambios comportamentales y los pensamientos que las invadieron (esto es, EEAG en el eje V del DSM-IV-TR; American Psychiatric Association, 2002). Posteriormente y de no informar de motu propio, se les requiere por medio de preguntas abiertas y de

acuerdo con el eje V del DSM-IV-TR (American Psychiatric Association, 2002), que revelen igualmente los efectos de la violencia de género a la que fueron sometidas en las relaciones familiares (EEGAR), sociales (EEASL) y laborales (EEASL). Además de esta modificación, la entrevista clínico-forense incluye un ajuste para el diagnóstico diferencial de simulación. Para ello se concretó un sistema categorial metódico con las siguientes 10 categorías para la detección de estrategias de simulación:

a) No cooperación con la evaluación. Se codifica esta posibilidad cuando el sujeto no muestra interés en la evaluación o no responde a la misma (American Psychiatric Association, 2013).
b) Síntomas sutiles. Los síntomas sutiles no son síntomas reales, sino problemas cotidianos que se informan como síntomas asociados a una enfermedad mental (p.ej., ser desorganizado, falta de motivación, dificultades en la toma de decisiones ordinaria).
c) Síntomas improbables/absurdos. Son aquellos que tienen un carácter fantástico o ridículo (opiniones, actitudes o creencias extrañas) y que no gozan de referentes reales (se excluyen aquí los síntomas raros).
d) Síntomas obvios. Éstos son síntomas de carácter psicótico que se relacionan con lo que vulgarmente se conoce como locura o enfermedad mental.
e) Síntomas (cuasi)raros. Esta contingencia se da cuando el sujeto dice padecer síntomas que raramente se observan incluso en poblaciones psiquiátricas reales y muy infrecuentemente en las normativas.
f) Combinación de síntomas y patrones espurios de psicopatología. Este indicador de simulación se codifica cuando el sujeto informa de síntomas reales que difícilmente se dan juntos (v.gr., esquizofrenia y extroversión), o cuando el sujeto describe un conjunto indiscriminado de síntomas sin consistencia interna entre ellos o configuraciones inusuales en poblaciones clínicas.
g) Severidad de síntomas. Esta categoría analiza el grado de severidad de los síntomas manifestados. Es frecuente que los simuladores atribuyan a la sintomatología que dicen padecer una severidad extrema.
h) Inconsistencia de síntomas (observados y manifestados). Esta categoría tiene por objeto analizar la correspondencia entre los síntomas elicitados por el sujeto y las observaciones del forense sobre si esos síntomas se corresponden con la actitud, presencia y comportamiento del sujeto.

i) Estereotipos erróneos. Esta categoría se refiere a las concepciones equivocadas sobre las características clínicas comúnmente asociadas con los trastornos mentales.

j) Agrupación indiscriminada de síntomas. El sujeto informa de problemas psiquiátricos de forma indiscriminada.

Las entrevistas de hechos y clínica han de ser grabadas para los subsecuentes análisis de contenido y guardada para que pueda ser sometida a contraste externo.

2. Repetición de la obtención del testimonio

La repetición del testimonio posibilita cotejar la continuidad del relato en el tiempo, criterio para la estimación de la fiabilidad del testimonio tanto científico, consistencia intra-medida (i.e., Weick, 1985), como legal, persistencia en el testimonio/incriminación (Sentencia del Tribunal Supremo de España, Sala de lo Penal, de 28 septiembre 1988, RJ 7070). Al respecto, se ha encontrado que la repetición de la entrevista en formato de discurso libre, como las incluidas en el SEG, no contamina el contenido (v.gr., Memon et al., 1997). Ahora bien, ha de tenerse presente que la consistencia, de acuerdo con la hipótesis Undeutsch (1967, pág. 125) y con la jurisprudencia (Novo y Seijo, 2010), debe entenderse en función de la centralidad o periferia del material que entra en contradicción. Sucintamente, sólo es relevante la contradicción si afecta a detalles centrales para la acción de juicio. Por su parte, la inconsistencia u omisión de información periférica relacionada con los hechos es asumible en una memoria de realidad.

3. Estudio de la motivación

El estudio de la motivación se deriva de tres pasos: Contraste de las declaraciones hechas a lo largo del proceso judicial; el contexto de la denuncia original, y los motivos/intereses para presentar una denuncia falsa. Para este estudio es preciso recabar todas las declaraciones del procedimiento judicial. Ahora bien, el valor de éstas es relativo. Es preciso tener en mente que muchas de ellas son transcripciones de lo que la denunciante ha dicho con lo cual no reflejan fehacientemente lo testificado. Además, el tipo de interrogatorio puede haber contaminado la respuesta. No en vano, los interrogatorios llevados a cabo por jueces provoca una presión hacia la conformidad en un conjunto de materias percibidas por los testigos como judicialmente deseables (Jones, 1987). A su vez,

nuestra experiencia nos advierte sistemáticamente que las denunciantes no refieren todos los hechos objeto de la denuncia. Así, las agresiones sexuales, la deprivación económica e incluso muchos episodios delictivos no son recogidos en las denuncias y diligencias judiciales. Asimismo, las declaraciones se suelen referir a expresiones (v.gr., me violó, pegó), no a narraciones de hechos (esto es, un evento narrativo de los hechos) con lo que no se puede contrastar su fiabilidad y validez, ni, en ocasiones, conocer del alcance de los daños. Por tanto, la falta de persistencia entre las declaraciones obtenidas ante los peritos y otras recogidas en el procedimiento tiene un valor relativo. En su caso, el forense debe dar cuenta que una hipotética falta de persistencia entre las declaraciones forenses y las obrantes en el procedimiento no tienen por qué ser causa de falta de persistencia de las declaraciones, sino que puede deberse a otras causas como el tipo de interrogatorio. Asimismo, el forense o perito ha de tener presente las presiones o intereses para presentar una declaración/denuncia falsa. Ahora bien, el estudio de la motivación no es consustancial para la estimación del valor de la prueba, que es función de la validez y fiabilidad del testimonio, sino que el forense o perito ha de conocer de ella pues puede ser trascendental para la defensa de la pericial o para la argumentación de la validez de un testimonio con inconsistencias entre la declaraciones ante los forenses y jueces/policías. En otras palabras, la motivación no es criterio estricto para la refutación de la prueba, sino para la justificación de posibles inconsistencias y para que el perito o forense tenga conocimientos para la defensa de la pericial.

4. Evaluación de la validez del testimonio

Previamente a proceder al estudio de la realidad del testimonio es necesario verificar previamente si el testimonio ha sido obtenido de modo que permita dicho estudio, esto es, en línea con el Sistema de Evaluación Global (Arce y Fariña, 2006), si es prueba suficiente y válida, esto es, si cumple con el criterio legal de persistencia en el testimonio/incriminación (Novo y Seijo, 2010). Para verificar si la *Prueba es (in)suficiente*, el psicólogo forense ha de plantearse las siguientes cuestiones ¿Supera la capacidad de memoria del testigo? ¿Contiene toda la información necesaria del evento? Además, es preciso que los testimonios sean prueba válida para proceder con el estudio de la realidad. De hecho, si una prueba no es válida ya no tiene sentido la estimación de la realidad de ésta. Para confrontar si la *Prueba es (in)válida* el perito o forense ha de responder a las siguientes preguntas:

a) (In)consistencia interna (¿tiene contradicciones internas en el relato?).
b) (In)consistencia externa (¿es consistente con otras pruebas robustas o incontrovertibles?).
c) (In)consistente con la anterior (¿hay consistencia en la información central inter-declaraciones?).
d) Persistencia en las declaraciones (¿son estables las declaraciones en el tiempo en los contextos?).
e) (In)consistencia con las leyes científicas y de la naturaleza (¿contiene el relato hechos incompatibles con las leyes científicas o de la naturaleza?).

Si la prueba es válida y suficiente para un análisis de la realidad se pasa a éste, pero de no ser válida o suficiente se detiene el análisis. Bajo esta última contingencia, es importante que el forense sea consciente de que de la desestimación de la prueba por ser insuficiente o inválida para la práctica de la prueba psicológico-forense no se puede inferir que sea falsa, y que el hecho de que la prueba no sea válida o suficiente para un estudio de la realidad, no implica que carezca de validez legal.

5. Análisis de la realidad de las declaraciones

Se aplica el sistema categorial ajustado SEG-A descrito anteriormente referidos. Siempre ha de tenerse en mente que en la práctica forense no son admisibles los falsos positivos; y que el registro de criterios de realidad se interpreta como soporte de la realidad del testimonio, en tanto que de la ausencia de criterios no se deriva que el testimonio sea falso.

6. Análisis de la fiabilidad de las medidas

Si bien los protocolos de evaluación referenciados de la realidad del testimonio y de la huella psíquica son científicamente válidos, al tiempo que han sido aceptados como prueba válida en las Salas de Justicia, esto no garantiza la fiabilidad del uso de ambos en los casos forenses. Por ello, el SEG incluye un procedimiento para estimar la fiabilidad de la evaluación en diseños de $n = 1$. En concreto, la fiabilidad de la medida se deriva de la estimación de la consistencia inter- e intra-medidas, inter-evaluadores e inter-contextos (Weick, 1985). Mediante el recurso a dos evaluadores con, al menos, uno de ellos entrenado y fiable en evaluaciones anteriores, que ejecuten la tarea por separado, se puede obtener una aproximación a la consistencia inter-evaluadores e inter-contextos. Como

herramienta estadística de análisis de la consistencia inter-evaluadores se toma el índice Kappa para las evaluaciones categóricas y la correlación intraclase para las continuas[6]. La fiabilidad inter-contexto se obtiene recurriendo a un evaluador entrenado que haya sido efectivo y consistente en otros contextos previos, o sea, en pericias anteriores. La consistencia inter-e intra-medidas viene de la mano de la consistencia interna de las medidas (v.gr., de las declaraciones en el tiempo, con los criterios diagnósticos; escalas e índices de validez; estrategias de simulación), de la consistencia entre distintas medidas (v.gr., con pruebas documentales, con otras pruebas, evaluación psicométrica y contenidos de la entrevista). Constada la consistencia inter- e intra-medidas, inter-evaluadores e inter-contextos se puede sostener que la réplica del procedimiento daría los mismos resultados.

7. Medida de la secuela psicológica del evento traumático

Para la evaluación forense de las secuelas psicológicas de la violencia de género, en el SEG se procede con la técnica forense previamente descrita. En todo caso, el psicólogo forense ha de tener en cuenta las siguientes limitaciones: No toda agresión de género produce secuela psicológica y la no detección de secuela psicológica no implica que no hubiera violencia de género sino que, en su caso, no le causó daño psicológico.

8. Evaluación del testimonio de testigos

Si bien en un principio la técnica fue creada para la evaluación del testimonio de la supuesta víctima o denunciante, el análisis de contenido en casos de violencia de género también se puede aplicar a testigos.

9. Análisis de las características psicológicas de los actores implicados

El estudio de la personalidad de los actores implicados puede ser de suma trascendencia ya que por medio de él podemos encontrar cualquier daño psíquico con implicaciones jurídicas relevantes tal como la ausencia o presencia de problemas psicológicos que afecten o no a la credibilidad de los testimonios recabados. Adicionalmente, se evalúan las capacidades cognitivas de los testigos (y, de ser necesario, se procederá con una exploración neuropsicológica) a fin de conocer la capacidad para testimoniar y ser objeto de una evaluación psicológico-forense. Por último, se evaluará

[6] El forense no familiarizado con estas herramientas estadísticas podría tomar el índice de concordancia [IC = Acuerdos/(acuerdos+desacuerdos)] contrastando, para las medidas continuas, la exacta correspondencia entre los registros, y tomando como punto de corte .80 (Tversky, 1977). No obstante, se advierte que, con esta estimación, la potencia de la fiabilidad pueda verse sensiblemente afectada.

cualquier otra área psicológica que pudiera resultar de interés para el caso (p.ej., valoración de riesgo de renuncia de la denuncia).

10. Implicaciones para la redacción del informe pericial

En las conclusiones de los informes periciales, los tribunales exigen la seguridad plena, no la alta probabilidad. No obstante, toda medida, y muy especialmente la psicológica, está sujeta a error. Evidencia que ha de reconocerse, pero sin que por ello se establezca una graduación de certeza que lleva a una mayor confusión para el decisor. De este modo, las categorías de evaluación de las realidad del testimonio serían: "declaración (muy) probablemente cierta/real/vivida/creíble", "declaración carente de criterios de realidad", "declaración o prueba inválida" y, en su caso, "declaración indeterminada" (también puede referirse como prueba insuficiente). Las categorías de evaluación de la secuela psicológica y del diagnóstico diferencial de simulación pueden verse en la descripción de la técnica forense.

Conclusiones

La actuación como psicólogo forense acorde a las prácticas descritas requiere de un alto grado de especialización, conocimiento del contexto y normativa legal aplicable, así como un exhaustivo entrenamiento. El último debe incluir: a) entrenamiento en los modos de obtención de la información en todas sus modalidades; b) entrenamiento en el análisis de la validez y realidad de los testimonios; c) entrenamiento en evaluación forense de la secuela psicológica; d) entrenamiento en la detección de la simulación de la secuela psicológica; y e) ejecución de las primeras evaluaciones forenses en compañía de un perito con experiencia.

Referencias[7]

American Psychiatric Association. (2002). *Manual diagnóstico y estadístico de los trastornos mentales* (IV Edición-Texto Revisado). Barcelona, España: Masson.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5a. ed.). Washington, DC: American Psychiatric Association.

[7] Las referencias marcadas con un asterisco (*) indican que forman parte de la revisión meta-analítica.

Arbisi, P. A. (2005). Use of the MMPI-2 in personal injury and disability evaluations. En J. N. Butcher (Ed.), *Practitioners handbook for the MMPI-2* (pp. 407-442). Washington, DC: American Psychological Association.

Arce, R., Alonso, M. A., y Novo, M. (2010). Estudio de sentencias de violencia de género con menores implicados: Carga de la prueba y decisiones sobre los menores. En F. Fariña, R. Arce, M. Novo, y D. Seijo (Eds.), *Separación y divorcio: Interferencias parentales* (pp. 241-252). Santiago de Compostela: Nino.

Arce, R., Carballal, A., Fariña, F., y Seijo D. (2004). Can mock battered women malinger psychological evidence in a recognition task? En A. Czerederecka, T. Jaskiewic-Obydzinska, R. Roesch, y J. Wójcikiewicz (Eds.), *Forensic psychology and law. Facing the challenges of a changing world* (pp. 327-336) Kraków: Institute of Forensic Research Publishers.

Arce, R., y Fariña, F. (2009). Evaluación psicológica forense de la credibilidad y daño psíquico en casos de violencia de género mediante el Sistema de Evaluación Global. En F. Fariña, R. Arce, y G. Buela-Casal (Eds.), *Violencia de género. Tratado psicológico y legal* (pp. 147-168). Madrid: Biblioteca Nueva

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 forense*. Manuscrito Inédito, Universidad de Santiago de Compostela, España.

Arce, R., y Fariña, F. (2006). Psicología del testimonio: Evaluación de la credibilidad y de la huella psíquica en el contexto penal. En Consejo General del Poder Judicial (Ed.), *Psicología del testimonio y prueba pericial* (pp. 39-103). Madrid, España: Consejo General de Poder Judicial.

Arce, R., y Fariña, F. (2007). Evaluación forense de la huella psíquica consecuencia de la violencia de género. En R. Arce, F. Fariña, E. Alfaro, C. Civera, y F. Tortosa (Eds.), *Psicología Jurídica. Evaluación e intervención* (pp. 47-57). Valencia: Servicio de Publicaciones de la Diputació de València.

Arce, R., y Fariña, F. (2012). La entrevista psicológico forense a niños, adultos y discapacitados. En S. Delgado, F. Bandrés, y A. Tejerina (Coords.), *Tratado de medicina legal y ciencias forenses: Vol. V. Pediatría legal y forense. Violencia. Víctimas* (pp. 795-817). Barcelona, España: Bosch.

Arce, R., Fariña, F., Carballal, A., y Novo, M. (2009). Creación y validación de un protocolo de evaluación forense de las secuelas psicológicas de la violencia de género. *Psicothema, 21*, 241-247.

Arce, R., Fariña, F., Novo, M., y Vilariño, M. (2013). Contraste de la efectividad en la detección de declaraciones fabricadas de los criterios de validez del Sistema de Evaluación Global en casos de violencia contra la mujer. En F. Expósito, I. Valor-Segura, M. Vilariño, y A. Palmer (Eds.), *Psicología jurídica aplicada a los problemas sociales* (pp. 153-161). Santiago de Compostela: Sociedad Española de Psicología Jurídica y Forense.

Arce, R., Fariña, F., y Vivero, A. (2007). Estudio exploratorio de la efectividad de las técnicas de análisis de contenido de las declaraciones en casos de violencia de género. En C. Guillén y R. Guil (Coords.), *Psicología social: Un encuentro de perspectivas* (Vol. 1, pp. 590-604). Cádiz, España: Asociación de Profesionales de la Psicología Social.

*Arce, R., Fariña, F., y Vilariño, M. (2010). Contraste de la efectividad del CBCA en la evaluación de la credibilidad en casos de violencia de género. *Intervención Psicosocial, 19*, 109-119.

Bardin, L. (1996). *El análisis de contenido* (2a. ed.). Madrid: Akal.

Bargai, N., Ben-Shakhar, G., y 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.

*Bensi, L., Gambetti, E., Nori, R., y Giusberti, F. (2009). Discerning truth from deception: The sincere witness profile. *The European Journal of Psychology Applied to Legal Context, 1*, 101-121.

*Blandón-Gitlin, I., Pezdek, K., Lindsay, D. S., y Hagen, L. (2009). Criteria-Based Content Analysis of true and suggested accounts of events. *Applied Cognitive Psychology, 23*, 90-917.

*Bogaard, G., Meijer, E. H., y Vrij, A. (2014). Using an example statement increases information but does not increase accuracy of CBCA, RM, and SCAN. *Journal of Investigative Psychology and Offender Profiling, 11*, 151-163.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., y Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology, 110*, 585-599.

Chivers-Wilson, K. A. (2006). Sexual assault and posttraumatic stress disorder: A review of the biological, psychological and sociological factors and treatments. *Mcgill Journal of Medicine, 9*, 111-118.

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

Daubert vs. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786 (1993).

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

Geiselman, R. E., Fisher, R. P., Firstenberg, I., Hutton, L. A. Sullivan, S., Avetissian, I., y Prosk, A. (1984). Enhancement of eyewitness memory: An empirical evaluation of the cognitive interview. *Journal of Police Science and Administration, 12*, 74-80.

*Gödert, H. W., Gamer, M., Rill, H. G., y Vossel, G. (2005). Statement Validity Assessment: Inter-rater reliability of Criteria-Based Content Analysis in the mock-crime paradigm. *Legal and Criminological Psychology*, *10*, 225-245.

Graham, J. R. (2011). *MMPI-2: Assessing personality and psychopathology* (5a. ed.). Nueva York, NY: Oxford University Press.

Greene, R. L. (2011). *The MMPI-2/MMPI-2-RF: An interpretive manual* (3a. ed.). Boston, MA: Allyn & Bacon.

Hathaway, S. R., y McKinley, J. C. (1999). *MMPI-2. Inventario Multifásico de Personalidad de Minnesota-2. Manual*. Madrid, España: TEA.

Jones, S. E. (1987). Judge- versus attorney-conducted voir dire: An empirical investigation of juror candor. *Law and Human Behavior, 11*, 131-146.

Köhnken, G. (2004). Statement validity analysis and the 'detection of the truth'. En A. Granhag y L. A. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 41-63). Cambridge, Inglaterra: Cambridge University Press.

*Lee, Z., Klaver, J. R., y Hart, S. D. (2008). Psychopathy and verbal indicators of deception in offenders. *Psychology, Crime & Law, 14*, 73-84.

Ley 1/2004, de 28 de diciembre, de Medidas de Protección Integral contra la Violencia de Género *Boletíon Oficial del Estado*, 303, 42166-42197. Recuperado de http://www.boe.es/boe/dias/2004/12/29/pdfs/A42166-42197.pdf

Mantwill, M., Köhnken, G., y Ascherman, E. (1995). Effects of the cognitive interview on the recall of familiar and unfamiliar events. *Journal of Applied Psychology, 80*, 68-78.

Memon, A., Wark, L., Bull, R., y Köhnken, G. (1997). Isolating the effects of the cognitive interview techniques. *British Journal of Psychology, 88*, 179-198.

*Merckelbach, H. (2004). Telling a good story: Fantasy proneness and the quality of fabricated memories. *Personality and Individual Differences, 37*, 1371-1382.

Novo, M., y Seijo, D. (2010). Judicial judgement-making and legal criteria of testimonial credibility. *European Journal of Psychology Applied to Legal Context, 2*, 91-115.

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). Nueva York, NY: Springer.

O'Donnell, M. L., Creamer, M., Pattison, P., y Atkin, C. (2004). Psychiatric morbidity following injury. *American Journal Psychiatry, 161*, 507-514.

Orcutt, H. K., Erickson, D. J., y Wolfe, J. (2004). The course of PTSD symptoms among Gulf War veterans: A growth mixture modelling approach. *Journal of Traumatic Stress, 17*, 195-202.

Rogers, R. (2008). Current status of clinical methods. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3a. ed., pp. 391-410). Nueva York, NY: The Guilford Press.

Rogers, R. Sewell, K. W., Martin, M. A., y Vitacco, M. J. (2003). Detection of feigned mental disorders: A meta-analysis of the MMPI-2 and malingering. *Assessment, 10*, 160-177.

Rosenfeld, B., Green, D., Pivovarova, E., Dole, T., y 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.

*Schelleman-Offermans, K., y Merckelbach, H. (2010). Fantasy proneness as a confounder of verbal lie detection tools. *Journal of Investigative Psychology and Offender Profiling, 7*, 24-260.

Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, 11, 373-397.

Steller, M. (1989). Recent developments in statement analysis. En J. C. Yuille (Ed.), *Credibility assessment* (pp. 135-154). Dordrecht, Holanda: Kluwer Academic Publishers.

Steller, M., y Köhnken, G. (1989). Criteria based statement analysis. En D. C. Raskin (Ed.), *Psychological methods for criminal investigation and evidence* (pp. 217-245). Nueva York, NY: Springer. (trad. castellano: *Métodos psicológicos en la investigación y pruebas criminales*. Bilbao, España: Desclée de Brouwer, 1994).

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

Undeutsch, U. (1967). Beurteilung der glaubhaftigkeit von zeugenaussagenn[La evaluación de la credibilidad de los testigos]. En U. Undeutsch (Ed.), *Handbuch der psychologie, Vol. II: Forensische psychologie* (pp. 26-181). Göttingen: Verlag für Psychologie.

Undeutsch, U. (1989). The development of statement reality analysis. En J. Yuille (Ed.), *Credibility assessment* (pp. 101-119). Dordrecht, Holanda: Kluwer.

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

Vilariño, M., Arce, F., y Fariña, F. (2013). Forensic-clinical interview: Reliability and validity for the evaluation of psychological injury. *The European Journal of Psychology Applied to Legal Context, 5*, 1-21.

Vilariño, M., Fariña, F., y 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.

Vilariño, M., Novo, M., y Seijo, D. (2011). Estudio de la eficacia de las categorías de realidad del testimonio del Sistema de Evaluación Global (SEG) en casos de violencia de género. *Revista Iberoamericana de Psicología y Salud, 2*, 1-26.

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

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2a. ed.). Chichester, Inglaterra: John Wiley and Sons.

Vrij, A., y Akehurst, L. (1998). Verbal communication and credibility: Statement validity assessment. En A. Memon, A. Vrij, y R. Bull (Eds.), *Psychology and law. Thuthfulness, accuracy and credibility* (pp. 3-31). Londres: McGraw-Hill.

*Vrij, A., Akehurst, L., Soukara, R., y Bull, R. (2004). Detecting deceit via analyses of verbal and nonverbal behavior in children and adults. *Human Communication Research, 30,* 8-41.

*Vrij, A., Edward, K., Roberts, K. P., y Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior, 24,* 239-263.

*Vrij, A., Kneller, W., y Mann, S. (2000). The effect of informing liars about Criteria-Based Content Analysis on their ability to deceive CBCA-raters. *Legal and Criminological Psychology, 5*, 57-70.

*Vrij, A., y Mann, S. A. (2006). Criteria-Based Content Analysis: An empirical test of its underlying processes. *Psychology, Crime & Law, 12,* 337-349.

*Vrij, A., Mann, S. A., Kristen, S., y Fisher, R. P. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior, 31*, 449-518.

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

*Willén, R. M., y Strömwall, L. A. (2011). Offender's uncoerced false confessions: A new application of statement analysis? *Legal and Criminological Psychology, 17,* 346-359.