

Contraste de la Efectividad del CBCA en la Evaluación de la Credibilidad en Casos de Violencia de Género

Contrasting the Efficiency of the CBCA in the Assessment of Credibility in Violence Against Women Cases

Ramón Arce
Universidad de Santiago de Compostela

Francisca Fariña
Universidad de Vigo

Manuel Vilariño
Universidad de Santiago de Compostela

Resumen. Los datos de archivo apuntan a que la probabilidad de condena en casos de violencia de género es aproximadamente del 70% (Ministerio de Igualdad, 2008), mientras en otros delitos está alrededor del 90% (Novo, 2000). A esta disfunción contribuyen múltiples causas, destacando entre ellas la falta de pruebas. Ésta viene derivada de que el delito se comete en un ámbito privado por lo que, en muchas ocasiones, el testimonio de la víctima es central. Para dotar de valor de prueba a éste, la jurisprudencia requiere de corroboraciones periféricas que lo avalen (criterio legal de verosimilitud), siendo crítica, entre ellas, la pericial psicológica de la credibilidad del testimonio. En un estudio de archivo, Arce, Vilariño y Alonso (2008) hallaron que esta prueba se practicaba en cerca del 20% de los casos. Por analogía se está procediendo con el sistema estándar de evaluación de la credibilidad del testimonio, el *Criteria Based Content Analysis* (CBCA; Steller y Köhnken, 1994), que no está validado para este contexto. Por ello, nos planteamos un estudio con 50 mujeres (25 víctimas reales y 25 irreales de violencia de género) para contrastar la eficacia de este procedimiento en la discriminación entre víctimas reales e irreales. Los resultados mostraron que las declaraciones reales contenían más criterios de realidad que las inventadas. No obstante, el procedimiento, acorde a las demandas de la práctica forense (no es admisible ningún falso positivo, esto es, caso irreal identificado como real), clasificaría erróneamente como falsos negativos al 44% de las declaraciones reales. Se discuten las implicaciones de los resultados para la práctica psicológico-forense. *Palabras clave:* violencia de género, maltrato, credibilidad, CBCA, testimonio, práctica forense.

Abstract. Archival data reveal that in Spain the conviction rate for cases of gender violence is approximately 70%, whereas the conviction rate for other crimes is in the region of 90%. This dysfunction arises from a multiplicity of factors, chiefly the lack of evidence. As most gender violence occurs within the privacy of the home, the burden of proof rests exclusively or primarily on the victim's testimony. Testimonies admitted in a court of law are often corroborated by circumstantial evidence (e.g., legal criteria of plausibility), and in particular the psychological report on testimonial credibility that plays a critical role in verdict outcome. An archive study of cases of gender violence revealed that psychological reports were admitted in 20% of the cases reviewed. Although Criteria Based Content Analysis (CBCA; Steller and Köhnken, 1994) is not valid for legal contexts, it has been observed to be the standard procedure for the evaluation of testimonial credibility in the archive cases of gender violence under review. Thus, a study involving 50 women (25 real victims and 25 feigners of gender violence) was undertaken to assess the efficacy of this procedure for the discrimination of real victims from feigners. The results reveal that real testimonies contained more reality criteria than faked testimonies. Nevertheless, the procedure, in line with the demands of forensic evidence in terms of the unacceptability of false positives (i.e., feigned case classed as real), erroneously detected 44% of real testimonies as false negatives (i.e., detected real case as feigned). The results are discussed in terms of the practical implications for forensic psychology. *Key words:* violence against women, battered woman, credibility, CBCA, testimony, expert testimony.

Introducción

El maltrato a la mujer, por parte de su pareja, representa una de las formas de violencia contra las mujeres más frecuente en las sociedades occidentales, encarnando un grave problema de derechos humanos y de salud pública (Matud, Bermúdez, y Padilla,

2009). En el contexto mundial, los malos tratos constituyen, tras la diabetes y los problemas de parto, la tercera causa que provoca más muertes prematuras y más secuelas físicas y psíquicas en las mujeres (Lorente, 2001). Dentro del Estado Español, entre los años 2000 y 2009, han sido asesinadas 629 mujeres, una media de 63 mujeres al año (68 desde la entrada en vigor de la Ley 1/2004, de 28 de diciembre, de Medidas de Protección Integral contra la Violencia) (Centro Reina Sofía para el Estudio de la Violencia,

La correspondencia sobre este artículo debe enviarse al primer autor
al e-mail: ramon.arce@usc.es

2010). Sin embargo, el número de denuncias ha aumentando gradualmente (Centro Reina Sofía para el Estudio de la Violencia, 2007; Ministerio de Igualdad, 2008). En consecuencia, las actuaciones legales no han logrado reducir los femicidios, pero sí visibilizar la violencia oculta. En lo tocante a las decisiones judiciales, las estadísticas del Observatorio contra la Violencia Doméstica y de Género del Consejo General del Poder Judicial (Europa Press, 2010) y del Ministerio de Igualdad (2008) apuntan a que no más del 70% de los casos llevados a juicio resultan en una sentencia condenatoria, cuando la tasa media en otro tipo de delitos es del 90% (Novo, 2000). Si bien no se puede apuntar a una única causa de esta reducida probabilidad de condena (e. g., falta de persistencia en la incriminación, error en la apreciación de la prueba, infracción por aplicación indebida del artículo 153 ó del 172,3), las dificultades dimanantes de la carga de la prueba parecen ser las más importantes. El problema probatorio radica en que, acorde a la doctrina del Tribunal Supremo (v. gr., Sentencia del TS de 29 de abril de 1997), cuando el testimonio de la víctima sea la única o la prueba central de cargo, éste debe reunir las tres siguientes características: ausencia de incredibilidad subjetiva, alguna corroboración periférica de carácter objetivo y persistencia en la incriminación, esto es, en el tiempo sin ambigüedades ni contradicciones. Entre las corroboraciones periféricas, en un informe de campo, Novo y Seijo (2010) hallaron que jueces y magistrados solicitan en el 32% de los casos un informe psicológico sobre la credibilidad del testimonio. Además, observaron que cuando el informe psicológico respalda la credibilidad del testimonio acusador, el caso se resuelve con la condena del encausado en el 93,3% de los casos. Por el contrario, si el informe concluye que el testimonio no es creíble se absuelve al acusado en el 100% de los casos. También, es habitual que el estudio de la credibilidad se complementa con una evaluación del daño psíquico ocasionado por los hechos denunciados. En el citado estudio, Novo y Seijo (2010) encontraron que se había efectuado una pericial psicológica sobre las secuelas psíquicas de la victimación en el 24% de los casos. Para la evaluación forense de la victimación psicológica por violencia de género se han creado y validado protocolos de actuación específicos (Arce, Fariña, Carballal, y Novo, 2009; Vilariño, Fariña, y Arce, 2009), aunque los psicólogos forenses siguen usando las técnicas tradicionales de evaluación psicológica que son totalmente ineficaces en la evaluación de la simulación (Rogers, 1997a), de la que se requiere indefectiblemente un diagnóstico diferencial en la evaluación forense (American Psychiatric Association, 2002). Asimismo, se ha desarrollado y validado un procedimiento de evaluación forense de la credibilidad del testimonio específico para casos de violencia de género (Arce y Fariña, 2009), pero los psicólogos forenses utilizan, mayormente, los procedimientos

estándar de evaluación de la credibilidad del testimonio que fueron ideados para menores víctimas de agresiones sexuales.

Las evaluaciones psicológicas de la credibilidad del testimonio se realizan a través de sistemas de análisis de contenido. Las dos técnicas más conocidas y empleadas son el *Reality Monitoring* (RM; Johnson y Raye, 1981) en la versión ampliada por Sporer (1997); y el *Statement Validity Assessment* (SVA; Steller, 1989; Steller y Boychuck, 1992). El RM se fundamenta en la asunción de que los recuerdos varían en una serie de rasgos. Así, aquellos que poseen un origen externo, esto es, que se basan en la percepción de un suceso, presentan más información sensorial, mayor cantidad de detalles contextuales y menos referencias a procesos cognitivos, que los que tienen un origen interno, es decir, inventados. Por su parte, el SVA, en la actualidad, representa probablemente el instrumento más usado en el estudio de la credibilidad (Bensi, Gambetti, Nori, y Giusberti, 2009). De hecho, las evaluaciones a través del SVA son aceptadas como prueba en algunas cortes norteamericanas y en varios países del norte de Europa como Suecia, Alemania y Holanda (Vrij, 2005). Su aplicación se compone de tres etapas: a) se obtiene el testimonio mediante una entrevista semi-estructurada, b) se analiza la credibilidad de la declaración mediante el *Criteria Based Content Analysis* (CBCA; Steller y Köhnken, 1994), y c) se aplica un listado de validez.

El estudio de la credibilidad de la declaración a través del CBCA constituye el elemento central del SVA. Este procedimiento parte de la *hipótesis de Undeutsch*, quien sostiene que las declaraciones derivadas de la vivencia de hechos reales se diferencian cualitativamente de aquellas que no se basan en la experiencia directa y que son fruto de la fantasía o inventiva del sujeto (Undeutsch, 1967, 1988). Inicialmente, se desarrolló con el objetivo de analizar el testimonio de niños víctimas de abusos sexuales. Aun así, diversas investigaciones han evidenciado que puede detectar el engaño en declaraciones de adultos (Köhnken, Schimossek, Aschermann, y Hofer, 1995; Sporer, 1997) y en casos diferentes al abuso sexual (Steller, 1989; Tye, Amato, Honts, Devitt, y Peters, 1999; Yuille, 1988).

El CBCA se compone de 19 criterios de realidad divididos en 5 categorías (ver Tabla 1). Estos criterios pueden evaluarse como presentes o ausentes, o bien, en función de la fuerza o grado en que aparecen en el testimonio (v.gr., ausente; presente; fuertemente presente). De este modo, cuanto mayor sea la presencia de estos criterios mayor será la probabilidad de que la declaración sea real. Sin embargo, de su ausencia no se desprende que la declaración sea falsa, sino carente de criterios de realidad.

La relevancia del CBCA para el estudio de la credibilidad ha motivado que haya sido objeto de multitud de investigaciones. Algunas ponen de manifiesto cier-

Tabla 1. Categorías y criterios del CBCA

Características generales. Los criterios aquí englobados se refieren a la declaración tomada en su totalidad.

1. Estructura Lógica. La declaración es coherente y consistente lógica y psicológicamente.
2. Elaboración inestructurada. La información se presenta en un orden no-cronológico.
3. Cantidad de detalles. La declaración es rica en detalles (lugares, sensaciones, información perceptual, etc.).

Contenidos Específicos. Se evalúan partes específicas del testimonio referidas a la presencia o fuerza de ciertos tipos de descripciones.

4. Engranaje Contextual (los hechos se sitúan en un tiempo y espacio).
5. Descripción de interacciones (la declaración contiene información en la que el agresor y la víctima interactúan).
6. Reproducción de conversaciones (diálogos específicos entre los actores de los hechos).
7. Complicaciones inesperadas durante el incidente (p.e., interrupción imprevista, incidente inesperado).

Peculiaridades del contenido. Se incluyen aquí aquellas características de una declaración que aumentan su concreción o viveza.

8. Detalles inusuales (detalles inesperados en el contexto de la declaración).
9. Detalles superfluos (detalles que no son relevantes para los hechos denunciados).
10. Incomprensión de detalles relatados con precisión (mención de detalles fuera del alcance de comprensión de la persona).
11. Asociaciones externas relacionadas.
12. Relatos del estado mental subjetivo (descripción de un cambio en los sentimientos, pensamientos, emociones o actitudes de la víctima durante el incidente).
13. Atribución del estado mental del autor del delito (descripciones por parte de la víctima del estado mental, motivos, estados fisiológicos o reacciones afectivas del agresor).

Contenidos Referentes a la Motivación. Estos criterios desvelan la motivación del testigo para hacer la declaración.

14. Correcciones espontáneas (corrección del testimonio ya prestado).
15. Admisión de falta de memoria (verbaliza que no recuerda información de los hechos).
16. Plantear dudas sobre el propio testimonio (la víctima manifiesta dudas sobre la exactitud del testimonio que está prestando).
17. Auto-desaprobación (información autoincriminatoria o desfavorable).
18. Perdón al autor del delito (la víctima emite afirmaciones favorecedoras o que excusan al acusado).

Elementos Específicos de la Agresión. Elementos del testimonio que no se relacionan con la viveza general de la declaración, sino con el delito.

19. Detalles característicos del delito (características de los hechos que contradicen las creencias habituales sobre cómo se producen este tipo de agresiones).

tas limitaciones del instrumento. Así, Porter y Yuille (1996) encontraron que sólo tres de las categorías del SVA/CBCA distinguían entre verdad y mentira; en tanto Blandon-Gitlin, Pezdek, Rogers, y Brodie (2005) y Pezdek et al. (2004) hallaron que la familiaridad del evento relatado parece desempeñar un papel más significativo en las puntuaciones obtenidas en el CBCA que la propia veracidad. Por su parte, Vrij, Akehurst, Soukara, y Bull (2002) demostraron que el conocimiento y entrenamiento en las categorías de realidad del CBCA merman su efectividad. Entre las variables intervinientes o mediadoras de la eficacia de los criterios de realidad del CBCA destacan la edad del evaluado (Vrij, Akehurst, Soukara, y Bull, 2004), la psicopatía (Lee, Klaver, y Hart, 2008), o trastornos de personalidad del grupo B, como el trastorno límite de la personalidad o el trastorno de la personalidad por inestabilidad emocional (Böhm, 2005; Böhm y Steller, 2008), pueden afectar al análisis de la credibilidad a través del CBCA. Por otro lado, Bensi et al. (2009) advirtieron que la falta de consistencia inter-estudios podía deberse al enfoque metodológico empleado en

los estudios experimentales. Generalmente, las investigaciones emplean diseños inter-sujeto lo que puede provocar que los resultados obtenidos se vean afectados por las diferencias individuales relativas a la conducta de engaño. Una estrategia para controlar estos efectos pasa por el uso de diseños intra-sujeto. Sometido a prueba los efectos del tipo de diseño, los resultados mostraron un mayor valor discriminativo entre declaraciones verdaderas e inventadas con diseños intra-sujeto.

En relación con el valor predictivo del CBCA, Vrij y Akehurst (1998), en una revisión de la literatura experimental, hallaron que éste oscilaba entre el 65% y el 85%, siendo mayor en la detección de declaraciones verdaderas que falsas. En una revisión posterior, Vrij (2000) obtuvo tasas de exactitud del 68% y del 76% en la clasificación de testimonios falsos y reales, respectivamente. Más recientemente, el mismo Vrij (2008) encontró unas tasas del 70,81% y del 70,47% para las declaraciones verdaderas y falsas, respectivamente. Sin embargo y a pesar de las limitaciones que pueda presentar el CBCA, en una revisión de la literatura,

Vrij (2005) apreció que el 92% de los estudios informaban sistemáticamente que las declaraciones de eventos verdaderos contienen más criterios de realidad que las inventadas.

Todo ello referido a la investigación experimental, pero cuando se trataba exclusivamente de la investigación con casos de campo, Vrij y Akehurst (1998) observaron un valor de clasificación correcta del 100% de los casos. En suma, la fiabilidad del sistema es mayor en contextos reales que en simulaciones.

Otra cuestión relevante es la existencia de diferencias en el poder discriminativo de los distintos criterios del CBCA. Sobre esto, DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, y Cooper (2003), en una amplia revisión sobre los diferentes indicadores de la mentira, observaron que los criterios más efectivos eran la cantidad de detalles, la estructura lógica, las correcciones espontáneas y la admisión de falta de memoria. Por su parte, Vrij (2005), en otra revisión, apuntó que eran la cantidad de detalles, la elaboración inestructurada, el engranaje contextual y la reproducción de conversaciones. Posteriormente, Lee et al. (2008), coincidiendo con Vrij en la bondad de los criterios cognitivos frente a los criterios motivacionales, hallaron que el engranaje contextual, la reproducción de conversaciones, las complicaciones inesperadas y la atribución del estado mental del autor del delito presentaban una mayor frecuencia dentro de las declaraciones reales. Como es fácilmente constatable, aunque existen criterios coincidentes entre los diversos estudios, también aparecen variaciones entre los mismos; algunas de éstas pueden atribuirse, claramente, al material experimental empleado en el estudio (p. e., que los criterios no estén presentes en el material que se visualiza para elaborar la historia que será objeto de evaluación) (Bensi et al., 2009).

Ante este estado de la literatura, nos hemos planteado llevar a cabo un estudio en el que poner a prueba el CBCA en casos de violencia de género. Para tal fin, se han recabado las declaraciones de 50 mujeres (25 casos reales y 25 falsos de violencia de género) a través de la entrevista cognitiva mejorada (EC; Fisher y Geiselman, 1992); posteriormente, los testimonios fueron sometidos a un análisis de contenido mediante el CBCA. En base a los resultados de la investigación se persigue satisfacer tres objetivos: en primer lugar, verificar si, de acuerdo con la literatura, las declaraciones reales contienen más criterios de realidad que las inventadas; en segundo término, establecer un criterio de decisión estricto, esto es, cuántos criterios del CBCA han de estar presentes en una declaración para que ésta sea considerada como real sin la comisión de falsos positivos; y, por último, conocer qué criterios del CBCA discriminan significativamente entre declaraciones reales e inventadas.

Método

Participantes

Las declaraciones sobre los hechos objeto de estudio se tomaron de 50 mujeres, 25 de ellas provenientes de casos reales acorde a una resolución judicial firme, con una edad media de 40,46 años ($DT = 10,46$), fluctuando entre los 19 y 64 años. Las otras 25, con historial de convivencia en pareja y que sostenían no haber sido objeto de violencia de género, tenían una edad promedio de 32,24 años ($DT = 13,71$), con un rango que oscilaba entre los 22 y los 72 años.

Diseño

La metodología de investigación empleada fue del tipo experimental con casos de archivo y de un contexto experimental. En concreto, se planificó un diseño para contrastar la memoria de casos reales de violencia de género, tomada de casos de archivo, con la memoria de declaraciones inventadas tomadas de víctimas irreales de violencia de género.

Procedimiento

Las declaraciones provenientes de casos reales fueron obtenidas de los archivos de la Unidad de Psicología Forense de la USC. Para ello, se consideraron declaraciones verdaderas todas aquellas en las que convergían dos criterios: la conformidad del acusado con la condena solicitada por la Fiscalía, esto es, la aceptación de los hechos imputados, y la existencia en el procedimiento de pruebas claras de culpabilidad (p.e., pruebas documentales, testimoniales, quebrantamientos de la orden de alejamiento) que llevaron a la condena del acusado. En ninguno de estos casos se había tomado el SVA y los criterios del CBCA como prueba de cargo. Por su parte, las declaraciones inventadas fueron recabadas entre mujeres mayores de edad que convivían o habían convivido en pareja. Para ello se les había instruido en lo que denominamos "instrucciones de engaño". Éstas consistían en pedirles que se pusiesen en el lugar de una mujer que presentaba una denuncia irreal de malos tratos y que iba a ser evaluada por ello en el contexto forense. La denuncia inventada de malos tratos se justificaba en la obtención de algún beneficio asociado (v. gr., guarda y custodia de los hijos, venganza, compensaciones económicas). Además, se les informaba de la relevancia de los resultados para la detección de las víctimas no reales (i.e., las secuelas indirectas en los hijos, condenas a inocentes). No se les dio ninguna instrucción de entrenamiento, únicamente que hicieran sus presentaciones creíbles y que se implicaran en la tarea (se les indicó a aquéllas que, por uno u otro motivo, no pudieran o quisieran

implicarse en la tarea que dejaran de participar en el estudio). Se les dio en torno a una semana para que planificaran la denuncia y evaluación. Transcurrida ésta, se inició el proceso de evaluación, consistente en la obtención de la declaración. El procedimiento seguido para la obtención de las declaraciones fue el de la entrevista cognitiva mejorada (Fisher y Geiselman, 1992). Los entrevistadores gozaban de experiencia forense (esto es, tenían práctica forense) y científica en este tipo de entrevistas (Arce, Fariña, y Vivero, 2007). En línea con las recomendaciones de Rogers (1997b) para la implementación de este tipo de diseños, se prestó especial atención a que las instrucciones fueran comprensibles (ya se habían mostrado como tales en estudios previos, Arce et al., 2007), específicas y contextualizadas para casos de violencia de género. Finalmente, se procedió a un *debriefing* en el que se prestó atención a si las participantes habían ejecutado la tarea de forma satisfactoria (recuerdo y comprensión de las instrucciones), a la motivación para participar, y a la preparación y grado de implicación en la tarea. De este modo, supimos que las participantes habían comprendido y ejecutado la tarea correctamente y que se habían implicado en la misma. Los resultados confirmaron la comprensión de la tarea e implicación de las participantes ya que crearon eventos narrativos constitutivos de violencia de género física y psicológica, incluyendo todas ellas criterios de realidad en sus declaraciones.

Instrumento de medida: La entrevista de hechos

Las declaraciones sobre los hechos que motivaban la denuncia fueron obtenidas por medio de la entrevista cognitiva (Fisher, Geiselman, y Amador, 1989). Ésta comprende cuatro técnicas generales de recuperación de memoria:

1. La primera técnica consiste en reconstruir mentalmente los contextos físicos y personales que existieron en el momento del crimen (o del hecho a recordar), esto es, la “reinstauración de contextos”. Esto implica que al testigo se le pida que trate de situarse mentalmente en el lugar del suceso teniendo en cuenta:
 - Elementos emocionales. Un ejemplo sería la sugerencia “trata de recordar cómo te sentías”.
 - Elementos secuenciales: “piensa en lo que estabas haciendo en ese momento”.
 - Características perceptuales: “Ponte de regreso en la escena del crimen y haz un dibujo de la habitación ¿Cómo olía? ¿Qué podías oír?”.

La razón que subyace a esta primera técnica es el “principio de codificación específica” de Tulving, que implica que la información contextual de un suceso se codifica junto con el evento y se conecta de una manera asociativa (Tulving y Thompson, 1973). Asimismo, la recuperación verbal

del suceso depende del grado en que los índices ambientales de la situación en la que se lleva a cabo el recuerdo se solapan con las propiedades previamente codificadas (Tulving, 1983). Diferentes investigaciones (p.e., Dando, Wilcock, y Milne, 2009; Davis, McMahon, y Greenwood, 2005; Emmett, Clifford, y Gwyer, 2003; Milne y Bull, 2002) han evidenciado que la activación de la información contextual que se procesa conjuntamente con el evento objeto de recuerdo, incrementa la cantidad de información correcta que se recupera sobre el mismo sin que, paralelamente, se produzca un aumento en el número de errores.

2. La segunda técnica, el “recuerdo libre”, consiste en pedirle al testigo que narre todo lo sucedido; que informe absolutamente de todo, incluso de los detalles que considere banales para la investigación, porque esos pueden llevar a otros, asociados entre sí en la memoria, que sí sean relevantes. Esta declaración, por tanto, es de tipo narrativo y al sujeto se le deja hablar, no se le interrumpe y no se realizan preguntas. Es preciso señalar que en todo momento ha de procurarse un ambiente propicio para la concentración del testigo, sin ruidos ni personas que lo distraigan, y es obvio que el entrevistador ha de ganarse la confianza del testigo para que su declaración sea lo más sincera y productiva posible.
3. La tercera técnica, el “cambio de perspectiva”, se asienta en la solicitud al testigo de que se ponga en el lugar de otro testigo del suceso, incluido el agresor, y que informe de lo que vio o hubiera visto desde esa perspectiva, si estuviera ocupando el lugar de esa persona. Esta técnica viene apoyada por los estudios de Bower (1967), quien advirtió que los sujetos al imaginarse los personajes de una historia, recordaban más detalles propios de la perspectiva del personaje con quien se han identificado, que de otros personajes. De esta manera se obtiene una segunda versión de la entrevista, desde una perspectiva diferente.
4. La última técnica consiste en la instrucción para que, desde diferentes puntos de partida, “recuerde en orden inverso”. En otras palabras, lo que se pretende es que el testigo narre los hechos desde un orden diferente a como se desarrollaron (v. gr., desde el final al principio, desde el medio hacia atrás) con el objetivo de recuperar pequeños detalles que pueden perderse al hacer una narración de los hechos siguiendo la secuencia temporal que éstos tuvieron. Esta técnica intenta reducir el efecto que los conocimientos previos, las expectativas y los esquemas producen en el recuerdo y, además, puede ser efectiva para recuperar detalles adicionales (Memon, Cronin, Eaves, y Bull, 1993). No en vano, como señalan Bower y Morrow (1990), se tiende a recordar el esquema o modelo mental que nos formamos de un evento,

más que el evento mismo. Esta técnica presenta una utilidad adicional para el estudio de la credibilidad del testimonio, ya que las declaraciones inventadas obtenidas en orden inverso contienen más indicios de engaño que las reales, lo que incrementa la capacidad de los evaluadores para detectar testimonios falsos (Vrij et al., 2008).

Además de esta versión estándar de la entrevista cognitiva, Fisher y Geiselman (1992) propusieron una mejorada, que fue la que se aplicó en nuestro caso. Ésta responde a una adaptación al contexto para la ejecución en un ámbito como el judicial. No obstante, la efectividad y el procedimiento en términos cognitivos es el mismo. Las siguientes fases resumen la estructura general mejorada de la misma.

- Fase 1. Presentaciones y personalización de la entrevista (presentación, usar nombre del entrevistado).
- Fase 2. Establecimiento de la comunicación (creación de atmósfera agradable, de confianza a través de la formulación de preguntas neutras).
- Fase 3. Explicación del propósito de la entrevista.
- Fase 4. Reinstauración de contexto.
- Fase 5. Recuerdo libre.
- Fase 6. Preparación para el interrogatorio (pedirle que se concentre intensamente, que diga lo que se le viene a la mente tal como llega, sin “fabricarlo”, que puede decir “no comprendo”, “no sé”, “no recuerdo”, etc., que active y contraste imágenes).
- Fase 7. Interrogatorio compatible con el testigo (cada testigo tiene una secuencia de memoria distinta del evento debiendo el interrogatorio ajustarse a esa secuencia).
- Fase 8. Recuerdo desde diferentes perspectivas.
- Fase 9. Recuerdo en orden inverso.
- Fase 10. Resumen (realizado por el entrevistador en función de lo que el entrevistado ha informado).
- Fase 11. Cierre (desactivación emocional y de tensiones en el entrevistado).

Por último, reseñar que la aplicación de la entrevista cognitiva no está limitada a la reproducción de un suceso de episodio único, ya que permite extender su utilización para recordar hechos que ocurren frecuentemente de una manera similar (Mantwill, Köhnken, y Ascherman, 1995), tal y como es característico de los casos reales de violencia de género. Además y para los propósitos de nuestro estudio, la entrevista cognitiva favorece la manifestación de los criterios de realidad (Vrij, 2005).

Análisis de los protocolos

Las declaraciones sobre los hechos objeto de la denuncia, grabadas en vídeo, fueron sometidas a un análisis de contenido en el que establecer la realidad de

las mismas por medio de las categorías de realidad del CBCA (ver Tabla 1). La unidad de análisis variaba en función de las categorías, siendo en unas la declaración en su totalidad (e.g., las características generales) o unidades lingüísticas como palabras o conjuntos de palabras (i.e., detalles), frases o conjuntos de frases (p.e., reproducción de conversaciones).

La unidad de análisis en todas las categorías fue el protocolo, registrándose las categorías como presentes o ausentes. Descartamos otras escalas de medida como el grado o intensidad de los criterios, no sólo porque la evaluación categórica se ajusta perfectamente al contexto forense, sino también porque las otras alternativas de medida abundan más en la carencia de objetividad del sistema al no poderse establecer puntos de corte sistemáticos ya que la productividad de la memoria está mediatizada por numerosos factores externos difícilmente controlables como la complejidad del evento, o la capacidad intelectual o narrativa de la persona evaluada. Sometido el sistema categorial metódico a un contraste de la consistencia interna obtuvimos un coeficiente *alpha de Cronbach* ($N = 50$) de ,789.

Entrenamiento de los codificadores

En el estudio participaron dos codificadores, ciegos en relación a las hipótesis de trabajo, con experiencia forense en la codificación de este tipo de material y con conocimientos de evaluación psicopatológica (Arce et al., 2007). Los codificadores fueron exhaustivamente entrenados en éste y otros sistemas de codificación. Para tal fin, tras presentarles y ejemplificarles cada categoría de análisis, en la ejecución con material que no iban a codificar posteriormente, se calculaba el índice de concordancia como instrumento de cotejo de la ejecución, lo que permitía, constatada inconsistencia, subsanar errores de codificación a través de la homogeneización de criterios.

Las definiciones de las categorías de análisis se elaboraron siguiendo la literatura existente (p.e., Steller y Köhnken, 1994). De este modo, los codificadores contaban como soporte para la codificación con un manual que recogía las definiciones de los diferentes criterios que componen el CBCA, así como ejemplos de cada uno de ellos.

Análisis de la fiabilidad de las entrevistas y codificaciones

Varios fueron los entrevistadores del estudio por lo que es preciso conocer si el factor entrevistador pudiera mediar la productividad de la entrevista. Para ello dividimos los protocolos de víctimas falsas y reales en dos grupos al azar. Si los protocolos son iguales no debería haber diferencias en las categorías de análisis. Los resultados del análisis de contenido de las entre-

vistas de hechos mostraron unos protocolos similares en las categorías de realidad, $F(1, 48) = 0,14$; $SC = 0,98$; ns , $\eta^2 = ,003$; $1-\beta = ,066$. 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 (p.e., Arce, Fariña, Carballal y Novo, 2006).

Asimismo, señalar que los dos codificadores dividieron el material entre ellos. La distribución del material a los codificadores fue aleatoria y combinando entrevistas de casos reales con entrevistas de casos falsos.

Por otro lado, para calcular la fiabilidad intra-jueces, los codificadores repitieron, una semana después del fin de la codificación original, en torno al 20% de las entrevistas que ellos mismos habían evaluado. Asimismo y para obtener la fiabilidad inter-jueces, cada codificador examinó alrededor del 20% de los protocolos que inicialmente había codificado el otro codificador.

Se consideran evaluaciones concordantes aquellas que superan el punto de corte de ,80 (Tversky, 1977), que es más restrictivo que los valores kappa. Contrastados nuestros resultados (véase la Tabla 2) podemos sostener que las evaluaciones son consistentes inter-, intra-evaluadores y en el tiempo. Además, los codificadores también fueron consistentes en otros contextos de evaluación (p.e., Arce et al., 2007). En suma, los datos obtenidos del análisis de contenido son fiables (Wicker, 1975).

Resultados

El contraste del estudio de las declaraciones a través del CBCA ha puesto de manifiesto que los criterios de

realidad del CBCA aparecen en mayor medida en las declaraciones verdaderas ($n = 316$) que en las falsas ($n = 241$), $\chi^2(1) = 10,10$; $p < ,001$. El estudio de la capacidad discriminativa entre declaraciones reales e inventadas de los criterios del CBCA (ver Tabla 3) mostró que los criterios estructura lógica, elaboración inestructurada, cantidad de detalles, complicaciones inesperadas, correcciones espontáneas, admisión de falta de memoria, planteamiento de dudas sobre el propio testimonio y autodesaprobación, discernían entre ambos tipos de declaraciones. En otras palabras, los criterios eficaces se concentran casi exclusivamente en las características generales de la declaración y en los contenidos referidos a la motivación. Por su parte, los criterios engranaje contextual, descripción de interacciones, detalles inusuales, detalles superfluos, incomprensión de detalles relatados con precisión, asociaciones externas relacionadas, relatos del estado mental subjetivo, atribución del estado mental del autor del delito, perdón al autor del delito y detalles característicos del delito, no diferenciaban entre ambos testimonios. En suma, los contenidos específicos (a excepción del criterio complicaciones inesperadas), las peculiaridades del contenido, los elementos específicos de la agresión y el criterio de contenido referido a la motivación, perdón al autor del delito, no son criterios de realidad válidos para casos de violencia de género. Así pues, del CBCA se derivan dos modelos de realidad: uno tomando todos los criterios de realidad (modelo estándar) y otro basado en aquellos con una capacidad discriminativa estadísticamente significativa entre declaraciones reales e inventadas (Modelo Ajustado).

Los anteriores resultados avalan estadísticamente la eficacia de los criterios de realidad del CBCA para diferenciar entre declaraciones reales e inventadas de violencia de género, así como qué criterios son produc-

Tabla 2. Índice de concordancia de las codificaciones de los criterios de realidad del CBCA

Variable	INTRA1	INTRA2	INTER
Estructura lógica	1	1	1
Elaboración inestructurada	1	1	1
Cantidad de detalles	1	1	1
Engranaje contextual	1	1	1
Descripción de interacciones	1	1	1
Reproducción de conversación	1	1	1
Complicaciones inesperadas incidente	1	,9	,95
Detalles inusuales	1	,9	,95
Detalles superfluos	,9	1	1
Incomprensión detalles relatados precisión	1	1	1
Asociaciones externas relacionadas	1	1	,95
Relatos del estado mental subjetivo	1	1	1
Atribución del estado mental del autor delito	1	1	1
Correcciones espontáneas	1	1	,95
Admisión de falta de memoria	1	1	,9
Plantear dudas sobre testimonio	1	1	,85
Auto-desaprobación	1	1	1
Perdón al autor del delito	1	,9	,95
Detalles característicos de la agresión	1	1	1

Tabla 3. Prueba de chi cuadrado para el estudio de la capacidad discriminativa entre declaraciones reales e inventadas de los criterios del CBCA

Variable	χ^2	p	phi	%dr	%di
CARACTERÍSTICAS GENERALES					
Estructura lógica	13,27	,001	-,562	100	52
Elaboración inestructurada	8,66	,003	-,468	96	56
Cantidad de detalles	3,72	,014	-,403	100	72
CONTENIDOS ESPECÍFICOS					
Engranaje contextual	—	—	—	100	100
Descripción de interacciones	—	—	—	100	100
Reproducción de conversaciones	—	—	—	100	100
Complicaciones inesperadas	7,29	,007	-,436	32	0
PECULIARIDADES DEL CONTENIDO					
Detalles inusuales	0	1	,000	36	36
Detalles superfluos	1,71	,192	-,246	96	80
Incomprensión detalles relatados	—	—	—	0	0
Asociaciones externas relacionadas	3,06	,080	-,288	76	48
Relatos estado mental subjetivo	—	—	—	96	96
Atribución estado mental autor delito	0,52	,470	,204	92	100
CONTENIDOS REFERENTES A LA MOTIVACIÓN					
Correcciones espontáneas	16,09	,001	-,608	72	12
Admisión de falta de memoria	5,15	,023	-,361	64	28
Plantear dudas sobre testimonio	9,52	,002	-,480	52	8
Auto-desaprobación	4,08	,043	-,327	56	24
Perdón al autor del delito	0	1	,000	12	12
ELEMENTOS ESPECÍFICOS DE LA AGRESIÓN					
Detalles característicos del delito	—	—	—	0	0

tivos en tal tarea, pero estos datos no se pueden generalizar directamente a la práctica forense porque ésta requiere de un criterio de decisión estricto y de una toma de decisiones diferente a la estadística. En concreto, estadísticamente se admiten unos márgenes de error en los denominados error tipo I (α) y tipo II (β), mientras que en la práctica forense el error tipo II o falsos positivos, esto es, la identificación de un caso inventado como real ha de ser cero (de hecho, en justicia se han acuñado muchos aforismos al respecto: es 10 veces más preferible absolver a un culpable que condenar a un inocente; in dubio pro reo; o la duda razonable ha de favorecer al acusado). Para contrastar la validez de este procedimiento para la práctica forense es preciso abordar un análisis de los efectos acumulativos de los criterios. Éste, que puede verse en la Tabla 4, pone de manifiesto que, para la correcta identificación de casos reales de violencia de género sin la comisión de ningún error tipo II, se han de verificar en la declaración más de 13 criterios de realidad en el Modelo Estándar del CBCA y más de 5 con el Modelo Ajustado. Ahora bien, el poder de correcta clasificación de las declaraciones verdaderas, sin identificar erróneamente ninguna inventada como real, es del 36% y del 48%, para el Modelo Estándar y Ajustado, respectivamente.

Tabla 4. Indicadores de realidad en el modelo del CBCA en las condiciones de verdad y mentira

Nº de criterios	Modelo estándar		Modelo ajustado	
	Real	Inventada	Real	Inventada
0				1(4%)
1				2(8%)
2				7(28%)
3			3(12%)	6(24%)
4			4(16%)	6(24%)
5		1(4%)	6(24%)	3(12%)
6		2(8%)	4(16%)	
7		2(8%)	4(16%)	
8		2(8%)	4(16%)	
9	1(4%)	3(12%)		
10	3(12%)	5(20%)		
11	4(16%)	6(24%)		
12	6(24%)	1(4%)		
13	2(8%)	3(12%)		
14	3(12%)			
15	3(12%)			
16	3(12%)			

Discusión

Previamente a la discusión de las implicaciones de estos resultados es necesario considerar las limitaciones que se deben tener presentes a la hora de generalizar los resultados. En primer lugar, el tamaño de la muestra es reducido por lo que los resultados pueden estar mediados, en parte, por éste. Así, aquellas categorías totalmente improductivas en la condición de casos reales puede que estén condicionadas por especificidad de los casos vividos. De este modo, podrían ser productivas si el abanico de casos fuera mayor. Segundo, a pesar de que se han tomado medidas con el objetivo de garantizar que las víctimas reales son tales, no existe certeza absoluta de ello. Igualmente, no existen garantías plenas de que las simuladoras nunca hubieran sido víctimas directas o indirectas de violencia de género. En tercer lugar, hemos asumido una equiparación entre prestar un testimonio inventado en condiciones reales y de laboratorio en recreaciones de alta fidelidad; sin embargo, no son totalmente similares. Cuarto, la casuística tomada es, exclusivamente, de violencia de género, por consiguiente, se han de tomar con gran cautela estos resultados a la hora de generalizarlos a otro tipo de casuísticas, incluida la violencia intrafamiliar que no sea de género. Con estas precisiones en mente, se pueden extraer una serie de conclusiones que den respuesta a los objetivos establecidos al inicio del estudio.

En línea con la literatura existente (v. gr., Gödert, Gamer, Rill, y Vossel, 2005; Vrij, 2005; Yuille, 1988), los criterios de realidad del CBCA aparecen en mayor medida en las declaraciones verdaderas que en las falsas. Esto es, los criterios de realidad cuentan con un aval estadístico a la hora de discriminar entre declaraciones reales a inventadas. No obstante, y en consonancia con la literatura previa (v.gr., Bensi et al., 2009; DePaulo et al., 2003; Vrij, 2005), no todos los criterios diferencian entre ambos tipos de declaraciones, lo que permite definir un modelo del CBCA estándar compuesto por todos los criterios del CBCA y otro ajustado constituido por aquellos criterios que diferencian significativamente entre declaraciones reales e inventadas. El rendimiento forense del modelo ajustado es superior al del estándar.

Dado que los resultados muestran, en consonancia con la literatura previa (Vrij, 2005), que las declaraciones inventadas pueden contener criterios de realidad y que el CBCA no define el número de criterios necesarios para establecer que una declaración es real, el procedimiento resultante vendría a ser para la práctica forense, como le definió ya en su momento Undeutsch (1967), *semi-objetivo* porque descansaría en el criterio subjetivo del forense. Sin embargo, el valor de prueba (Novo y Seijo, 2010) para la acusación (cuando el forense concluye que la declaración es real) otorgado por jueces y magistrados es del 93,3% y del 100% para la absolución (el forense concluye que no es creíble).

En consecuencia, el sistema es, de facto, *semi-objetivo*, pero se le confiere un valor de prueba prácticamente total, esto es, se le presupone total objetividad.

Si bien no se ha cuantificado un criterio de decisión estricto en cuanto al número de criterios necesarios para concluir que un testimonio es real, los psicólogos forenses requieren y notifican en sus informes periciales de varios criterios cuando soportan que el testimonio hace referencia a hechos reales. No contamos con estudios que refieran el número mínimo de criterios usados por los forenses, pero la literatura científica ha informado que podrían ser 5 (p.e., Landry y Brigham, 1992) ó 7 (e.g., Lamb, Sternberg, Esplin, Hershkowitz, Orbach, y Hovav, 1997; Steller, 1989) por lo que es de esperar que sean éstos. Con estos criterios, los falsos positivos oscilarían entre el 80% y el 96%, para más de 7 y 5 criterios, respectivamente. Resultados similares fueron hallados por Arce, Seijo y Rodríguez (2005) con muestras de adultos en casos de agresiones sexuales y amenazas. En suma, la utilidad del CBCA para la práctica forense sería nula al permitir que entre el 80 y el 96% de las declaraciones inventadas fueran clasificadas como reales.

En conclusión, los criterios de realidad discriminan entre declaraciones reales e inventadas de violencia de género, pero el sistema resultante del CBCA no es válido en su formulación original para la práctica forense en casos de violencia de género. Por su parte, los modelos forenses basados en criterios de decisión estrictos (esto es, sin la comisión de falsos positivos) que podrían derivarse no son muy efectivos al alcanzar una tasa forense de correcta clasificación inferior al 50% de los casos reales. Por ello, y dado que el sistema de criterios de realidad presenta un potencial para la discriminación entre declaraciones reales e inventadas de violencia de género, se han de abrir nuevas vías en las que se combinen los criterios del CBCA con los de otros sistemas como el RM (p.e., Vrij et al., 2004), o se confeccione un sistema *ad hoc* para casos de violencia de género (Arce y Fariña, 2009).

Referencias

- American Psychiatric Association (2002). *Manual diagnóstico y estadístico de los trastornos mentales* (IV Edición-Texto Revisado). Barcelona: Masson.
- Arce, R., y Fariña, F. (2009). Evaluación psicológico 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. Buena-Casal (Eds.), *Violencia de género. Tratado psicológico y legal*. Madrid: Biblioteca Nueva.
- 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.
- Arce, R., Fariña, F., Carballal, A., y Novo, M. (2009). Crea-

- ción 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. 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* (pp. 590-604). Cádiz: Asociación de Profesionales de la Psicología Social.
- Arce, R., Seijo, D., y Rodríguez, M. J. (2005). Falsas alarmas de los criterios de realidad en delitos de agresiones sexuales e intimidación en muestras de adultos. En J. Sobral, G. Serrano y J. Regueiro (Comps.), *Psicología jurídica, de la violencia y de género* (pp. 31-43). Madrid: Biblioteca Nueva.
- Arce, R. Vilariño, M., y Alonso, M. A. (2008, Diciembre). *Estudio de sentencias de violencia de género con menores implicados: Carga de la prueba y decisiones sobre menores*. I Congreso de Multidisciplinar sobre Interferencias Parentales tras la Ruptura de Pareja, Santiago de Compostela.
- 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.
- Blandon-Gitlin, I., Pezdek, K., Rogers, R., y Brodie, L. (2005). The effects of event familiarity on criterion-based content analysis ratings: An experimental study. *Law and Human Behavior*, 29, 187-197.
- Böhm, C. (2005). El peritaje psicológico sobre la credibilidad de declaraciones testimoniales en Alemania. Nuevos desafíos. En R. Arce, F. Fariña, y M. Novo (Eds.), *Psicología jurídica* (pp. 93-101). Santiago de Compostela: Consellería de Xustiza, Interior e Administración Local. Xunta de Galicia.
- Böhm, C., y Steller, M. (2008). Testimonio y trastorno límite de personalidad. En F. J. Rodríguez, C. Bringas, F. Fariña, R. Arce, y A. Bernardo (Eds.), *Psicología jurídica. Entorno judicial y delincuencia* (pp. 135-147). Oviedo: Ediciones de la Universidad de Oviedo.
- Bower, G. (1967). A multicomponent theory of memory trace. En K. W. Spence y J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 229-325). Nueva York: Academic Press.
- Bower, G. H., y Morrow, D. G. (1990). Mental models in narrative comprehension. *Science*, 247, 44-48.
- Centro Reina Sofía para el Estudio de la Violencia (2007). *Mujeres maltratadas por su pareja*. Recuperado el 16 de Marzo de 2010 de <http://www.centroreinasofia.es/paneldecontrol/est/pdf/EST007-3271.pdf>
- Centro Reina Sofía para el Estudio de la Violencia (2010). *Mujeres asesinadas por su pareja. España (2000-2009)*. Recuperado el 16 de Marzo de 2010 de <http://www.centroreinasofia.es/paneldecontrol/est/pdf/EST014-3293.pdf>
- Dando, C., Wilcock, R., y Milne, R. (2009). The cognitive interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology*, 23, 138-147.
- Davis, M. R., McMahan, M., y Greenwood, K. M. (2005). The efficacy of mnemonic components of the cognitive interview: Towards a shortened variant for time-critical investigations. *Applied Cognitive Psychology*, 19, 75-93.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., y Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129, 74-118.
- Emmett, D., Clifford, B. R., y Gwyer, P. (2003). An investigation of the interaction between cognitive style and context reinstatement on the memory performance of eyewitnesses. *Personality and Individual Differences*, 34, 1495-1508.
- Europa Press (2010, Febrero, 16). *Condenados casi 100.000 hombres por violencia de género en tres años y medio*. [Documento WWW]. URL <http://www.europapress.es/nacional/noticia-condenados-casi-100000-hombres-violencia-genero-tres-anos-medio-20090421214004.html>
- Fisher, R. P., y Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interview*. Springfield: Charles C. Thomas.
- Fisher, R. P., Geiselman, R. E., y Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and witness of crime. *Journal of Applied Psychology*, 74, 722-727.
- 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.
- Johnson, M. K., y Raye, C. L. (1981). Reality monitoring. *Psychological Review*, 88, 67-85.
- Köhnken, G., Schimossek, E., Aschermann, E., y Höfer, E. (1995). The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Cognitive Psychology*, 80, 671-684.
- Lamb, M. E., Sternberg, K. J., Esplin, P. W., Hershkowitz, I., Orbach, Y., y Hovav, M. (1997). Criterion-based content analysis: A field validation study. *Child, Abuse & Neglect*, 21, 255-264.
- Landry, K. L., y Brigham, J. C. (1992). The effect of training in criteria-based content analysis on the ability to detect deception in adults. *Law and Human Behavior*, 16, 663-676.
- Lee, Z., Klaver, J. R., y Hart, S. D. (2008). Psychopathy and verbal indicators of deception in offenders. *Psychology, Crime and Law*, 14, 73-84.
- Lorente, M. (2001). *Mi marido me pega lo normal*. Barcelona: Ares y Mares.
- 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.
- Matud, P., Bermúdez, M. P., y Padilla, V. (2009). Intervención psicológica con mujeres maltratadas por su pareja. En F. Fariña, R. Arce, y G. Buela-Casal (Ed.), *Violencia de Género. Tratado psicológico y legal* (pp.

- 193-207). Madrid: Biblioteca Nueva.
- Memon, A., Cronin, O., Eaves, R., y Bull, R. (1993). The cognitive interview and the child witness. En N. K. Clark y G. M. Stephenson (Eds.), *Issues in criminology and legal psychology: Vol. 20. Children, evidence and procedure* (pp. 3-10). Leicester: British Psychological Society.
- Milne, R., y Bull, R. (2002). Back to basics: A componential analysis of the original cognitive interview mnemonics with three age groups. *Applied Cognitive Psychology, 16*, 743-753.
- Ministerio de Igualdad (2008). *Evaluación de la aplicación de la ley orgánica 1/2004, de 28 de diciembre, de medidas de protección integral contra la violencia de género*. Recuperado el 1 de Abril de 2009 de http://www.migualdad.es/noticias/pdf/informe_ejecutivo-14_julio_2008_def.pdf
- Novo, M. (2000). *Heurísticos y evidencia en la toma de decisiones en la sala de justicia*. Tesis Doctoral, Universidad de Santiago de Compostela.
- Novo, M., y Seijo, D. (2010). Judicial judgement-making and legal criteria of testimonial credibility. *The European Journal of Psychology Applied to Legal Context, 2*, 91-115.
- Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., et al. (2004). Detecting deception in children: Event familiarity affects criterion-based content analysis ratings. *Journal of Applied Psychology, 89*, 119-126.
- Porter, S., y Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues in the interrogation context. *Law and Human Behavior, 20*, 443-458.
- Rogers, R. (1997a). Introduction. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 1-19). New York: Guilford Press.
- Rogers, R. (1997b). Researching dissimulation. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2ª ed., pp. 398-426). Nueva York: Guilford Press.
- 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: Kluwer Academic Publishers.
- Steller, M., y Boychuck, T. (1992). Children as witness in sexual abuse cases: Investigative interview and assessment techniques. En H. Dent y R. Flin (Eds.), *Children as witness* (pp. 47-71). Chichester: John Wiley and Sons.
- Steller, M., y Köhnken, G. (1994). Análisis de declaraciones basados en criterios. En D. C. Raskin (Ed.), *Métodos psicológicos en la investigación y pruebas criminales* (pp. 217-245). Bilbao: Desclée de Brouwer (Orig. 1989).
- Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon Press.
- Tulving, E., y Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 353-370.
- Tye, M. J. C., Amato, S. L., Honts, C.R., Deveitt, M. K., y Peters, D. (1999). The willingness of children to lie and the assessment of credibility in an ecologically relevant laboratory setting. *Applied Developmental Science, 3*, 92-109.
- Undeutsch, U. (1967). Beurteilung der glaubhaftigkeit von zeugenaussagen [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. (1988). The development of statement reality analysis. En J. Yuille (Ed.), *Credibility assessment* (pp. 101-119). Dordrecht: Kluwer Academic Publishers.
- Vilariño, M., Fariña, F., y Arce, R. (2009). Discriminating real victims from feigners of psychological injury in gender violence: Validating a protocol for forensic setting. *The European Journal of Psychology Applied to Legal Context, 1*, 221-243.
- Vrij, A. (2000). *Detecting lies and deceit*. Chichester: John Wiley and Sons.
- 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*. Chichester: John Wiley and Sons.
- Vrij, A., y Akehurst, L. (1998). Verbal communication and credibility: Statement validity. En A. Memon, A. Vrij, y R. Bull (Eds.), *Psychology and law. Truthfulness, accuracy and credibility* (pp. 3-31). Londres: McGraw-Hill.
- Vrij, A., Akehurst, L., Soukara, S., y Bull, R. (2002). Will the truth come out? The effect of deception, age, status, coaching and social skills on CBCA scores. *Law and Human Behavior, 26*, 261-283.
- Vrij, A., Akehurst, L., Soukara, S., 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., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., y Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior, 32*, 253-265.
- Wicker, A. W. (1975). An application of a multiple-trait-multimethod logic to the reliability of observational records. *Personality and Social Psychology Bulletin, 4*, 575-579.
- Yuille, J. C. (1988). The systematic assessment of children's testimony. *Canadian Psychology, 29*, 247-262.

Manuscrito Recibido: 10/01/2010

Revisión Recibida: 22/03/2010

Manuscrito Aceptado: 22/03/2010