Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones

Autores: Ramón Arce y Francisca Fariña.

Introducción

La evaluación de la credibilidad del testimonio ha sido a lo largo de la historia de la justicia una de las piedras angulares de las decisiones judiciales, ideándose diversos procedimientos para su estimación. Estos procedimientos se clasifican en correlatos del comunicador, indicios no verbales y paraverbales, indicios psicofisiológicos e indicios cognitivos (p. e., Arce y Fariña, 2006a; De Paulo Lindsay, Malone, Muhlenbruck, Charlton, y Cooper., 2003; Sporer, 1997; Vrij, 2008). Los correlatos del comunicador se relacionan con variables del sujeto, tales como características de personalidad o el tipo de población, que se vinculan con la mentira. Por ejemplo, bien conocida y establecida es la relación entre psicopatía y mentira, hasta el punto de que en el DSM-IV (American Psychiatric Association, 2002), el trastorno antisocial de la personalidad (psicopatía, sociopatía o trastorno disocial, como también se denomina) cuenta entre sus criterios diagnósticos con un patrón continuado de mentira repetida (Criterio A2). Asimismo, a los perpetradores del trastorno facticio por poderes se les atribuye como síntoma asociado la mentira patológica. También es bien sabido que a los niños se les confiere escasa credibilidad en el testimonio por motivos como la falta de desarrollo de la memoria, lo que se traduce en una memoria menos productiva (Memon, Meissner, y Freiser, 2010), y escasa fiabilidad (p,ej., alta sugestionabilidad, imaginabilidad, falsas memorias) (v.gr., Goodman, 2005; Judicial Studies Board, 2006). Como es obvio, estos rasgos no pueden aplicarse a una persona concreta en un momento dado. Ahora bien, judicialmente sí se observa que éstos median la fiabilidad conferida a los testigos (Bull, 1997). Más allá de la clasificación señalada anteriormente advertimos que cuestión distinta sería, si en vez de como rasgos, algunos de ellos fueran tomados como estados. Así, si bien un testigo con rasgos psicóticos puede ser fiable, no es menos cierto

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

que bajo un brote psicótico, no lo es. Del mismo modo, el engaño en términos de salud mental (simulación, disimulación o sobresimulación) ha de sospecharse en toda evaluación forense (American Psychological Association, 2002), esto es, es el contexto el que activa la probabilidad de engaño, no un rasgo.

Los indicios no verbales y paraverbales forman parte no sólo del acervo científico, sino también de la estimación social de la credibilidad del testimonio; es decir, de los mecanismos que usa el sujeto común en la estimación de la credibilidad. No en vano, todos somos, y estamos obligados, para ser adaptativos, detectores de mentiras, constituyendo los indicios no verbales y paraverbales los medios. La Tabla 1 recoge los indicios no verbales y paraverbales más manidos, así como la predicción modal de su valor predictivo del engaño (Vrij, 2000, 2008), tanto objetivo (científico), como subjetivo (uso social). Ahora bien, resulta que las predicciones científicas no son tan homogéneas como en esta recopilación de Vrij. De hecho, de los cuatro modelos teóricos explicativos de la relación entre indicadores no verbales y paraverbales, el modelo del *afecto* (relaciona la mentira con estados emocionales de miedo o de sentimiento de culpabilidad), el del *arousal* (vincula la mentira con un incremento en la actividad automática), el del *control* (los mentirosos intentan controlar algunos canales de comunicación para generar la impresión de credibilidad) y el *cognitivo* (la mentira demanda del sujeto un mayor esfuerzo cognitivo que la verdad, por lo que algunos indicadores incrementarán su tasa) no coinciden en las predicciones en ninguno de los indicadores 8véas en la Tabla 2 un resumen de las predicciones). Por ejemplo, las automanipulaciones se incrementarían con la mentira según las predicciones de los modelos del afecto y arousal, y decrecerían acorde a los modelos de control y cognitivo (Sporer y Schwandt, 2006, 2007). En suma, estos indicadores del engaño cometen el denominado error de idiosincrasia; es decir, que predicen una tendencia, y justamente la contraria (Ekman y O'Sullivan, 1994). En un meta-análisis sobre los indicadores paraverbales asociados al engaño, Sporer y Schwandt (2006) observaron que, en general, sólo el tono y la latencia de respuesta estaban ligados con la mentira, ambos con una relación positiva. En otro meta-análisis sobre los indicadores no verbales, Sporer y Schwandt (2007) hallaron que globalmente sólo el asentimiento/disentimiento, los movimientos de pies y piernas, y los movimientos de manos se relacionaban con la mentira, todos ellos negativamente. Sin embargo, ningún modelo explicativo da cuenta de estos resultados. Además, en ambos

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

meta-análisis se encontró que la dirección de la relación estaba condicionada por variables moderadoras (p. ej., motivación, contenido de la mentira, formato de la medida). En resumen, este estado de la literatura deja *en cuarentena* el valor de prueba de estos indicadores, porque no tienen un respaldo de modelo teórico alguno; no componen un sistema formal de relaciones, al no poder establecerse una correspondencia con el engaño en la mayoría de los indicadores; la predicción de la mentira está mediada por diferentes variables moderadoras; y porque no se puede extraer de ellos un criterio de decisión forense.

Tabla 1. Indicios paraverbales y no verbales y predicción objetiva (científica) y subjetiva (uso social) del engaño (Vrij, 2000).

Indicadores paraverbales	**Objetiva**	**Subjetiva**
Dudas del habla: interjecciones (expresiones como “ah”, “um”)	--	>
Oraciones incompletas, lapsus linguae, etc.	--	>
Tono de voz: cambios en el tono de voz, tales como subidas o bajadas	>	>
Tasa del habla: número de palabras habladas en un cierto período de tiempo	--	--
Latencia de la respuesta: tiempo de silencio entre la pregunta y la respuesta	>	--
Frecuencia de las pausas: frecuencia de períodos de silencio durante el habla	>	--
Duración de las pausas: longitud de los períodos de silencio durante el habla	--	>
Indicadores no verbales		
Mirada: mirar a la cara del interlocutor	--	<
Sonrisa: sonrisas y risas	--	--
Parpadeo: parpadeo de los ojos	--	>
Auto-manipulaciones: rascarse la cabeza, muñeca, etc.	--	>
Movimientos ilustradores: movimientos funcionales de brazos y manos dirigidos a modificar o suplementar lo que se está diciendo verbalmente	<	--
Movimientos de manos y dedos: movimientos no-funcionales de manos y dedos sin movimientos de brazos	<	>
Movimientos de piernas y pies	<	>
Movimientos de la cabeza: asentimiento y disentimiento con la cabeza	--	--
Movimientos del tronco: movimientos del tronco (generalmente		

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

acompañados con movimientos de la cabeza)	--	>
Cambios de posición: cambios en la postura corporal (generalmente acompañado de movimientos del tronco y piernas/pies)	--	>

Nota. Tendencia de interpretación empírica: > incremento con la mentira; < decrece con la mentira; -- no relacionada con la mentira; 1 cuando contar una mentira implica un gran esfuerzo de pensamiento.

Tabla 2. Predicciones de los indicadores no verbales y paraverbales de la mentira según los modelos teóricos (Sporer y Schwandt, 2006, 2007).

Indicador/modelo explicativo	**Arousal**	**Afecto**		**Control**	**Cognitivo**
		Miedo	**Culpable**		
Indicadores Paraverbales					
Duración del mensaje	?	<	?	< | >	<
Número de palabras	?	<	?	< | >	<
Tasa del habla	>	>	<	< | >	<
Interrupciones habla (um, ah,…)	>	>	?	<	>
Duración de las pausas	?	>	?	<	>
Tono de voz	>	>	? | <	?	?
Repeticiones de palabras o frases	? | >	>	?	?	>
Latencia de respuesta	? | <	?	>	<	>
Errores del habla (gramaticales)	>	?	?	? | <	>
Indicadores no Verbales					
Parpadeo	>	>	?	?	?
Contacto visual	?	<	<	>	<
Evitación mirada	?	>	>	<	>
Movimientos cabeza (inhibición)	>	>	<	<	<
Asentimiento/disentimiento	>	<	< | >	? | >	?
Sonrisas	?	<	< | >	? | >	<
Automanipulaciones	>	>	?	<	<
Movimientos de manos	>	>	?	<	<

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Mov. Ilustradores	?	<	<	< | >	<
Mov. de piernas y pies	>	>	?	<	<
Cambio de postura	>	?	?	<	?

Nota. > = incrementa con la mentira; < = decrece con la mentira; ? = indeterminado; | = predicciones contrapuestas en función de la asunción de unos prepuestos u otros de la aproximación.

Los indicadores fisiológicos estudiados en relación con la veracidad del testimonio incluyen el polígrafo, el registro de potenciales evocados relativos a eventos (PRE), medidas de tiempo de reacción (TR) y el análisis de la imagen obtenida mediante resonancia magnética funcional (IRMf). Éstos se basan en que la mentira se asocia a actividad automática (polígrafo), actividad cortical específica (PRES), a un mayor tiempo de reacción, y la localización en que se produce la actividad cerebral (IRMf). Sin embargo, no se han encontrado pruebas que validen estos supuestos, al tiempo que son admisibles otras hipótesis alternativas a las que predicen los modelos. El objeto de este capítulo no permite la entrada en profundidad en estas y otras técnicas menores. Remitimos al lector interesado a la exhaustiva revisión de Vrij (2008). Seguidamente examinaremos los indicios cognitivos.

Indicios cognitivos: Análisis de la realidad de las declaraciones

Del conjunto de técnicas científicas que se han postulado para la estimación de la credibilidad del testimonio, el estudio cognitivo, basado en la conocida como *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 (p. ej., Bensi, Gambetti, Nori, y Giusberti, 2009; Granhag, Strömwall, y Landström, 2006; Masip, Sporer, Garrido, y Herrero, 2005; Ruby y Brigham, 1997; Vrij, 2000, 2005, 2008). Así, Vrij (2005), en una revisión de la evidencia empírica, halló que el 92% de los estudios confirmaban esta hipótesis con las

categorías del CBCA (Criteria Based Content Analysis). 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. La potencia de los sistemas de análisis de la realidad de la declaración del testimonio derivados de este procedimiento es tal que Novo y Seijo (2010), en un análisis de sentencias en las que la verosimilitud del testimonio era el eje central de la decisión judicial, hallaron que la pericial psicológica basada en el análisis de la realidad del testimonio era la única prueba pericial en la que se apoyaban jueces y tribunales españoles. En concreto, en el 33% de los casos, siendo del 93,3% la tasa de condena si la pericial avalaba la realidad del testimonio, y del 100% de desestimación del caso si no la respaldaba. Esas diferencias entre memorias se concretaron en categorías de análisis (p. ej., elaboración inestructurada, concreción, correcciones espontáneas, información visual), que se agruparon en sistemas categoriales de análisis (v.gr., CBCA, SRA, RM). Inicialmente estos sistemas fueron estudiados y aplicados a la memoria de menores abusados sexualmente. No obstante, como quiera que la hipótesis Undeutsch también es aplicables a otras poblaciones, circunstancias y casos, se ha comprobado que estos sistemas son igualmente válidos para poblaciones de adultos, a testigos no víctimas, u otros casos como agresiones sexuales, amenazas o violencia de género (p. ej., Akehurst, Köhnken, y Höfer, 2001; Arce, Fariña, y Freire, 2002; Arce, Fariña, y Vivero, 2007; Bensi et al., 2009; Köhnken, Schimossek, Aschermann, y Höfer, 1995; Landry y Brigham, 1992; Porter y Yuille, 1995, 1996; Ruby y Brigham, 1997; Sporer, 1997; Steller y Köhnken, 1989/1994; Vilariño, Novo, y Seijo, 2011; Zaparniuk, Yuille, y Taylor, 1995). En suma, estos sistemas son aplicables a la memoria y no a poblaciones, casuísticas y circunstancias concretas, aunque los resultados experimentales no deben generalizarse directamente de unas a otras. No en vano y a modo de ejemplo, Arce, Fariña y Vilariño (2010) encontraron que los criterios del CBCA no eran válidos para casos de violencia de género, en tanto Willén y Strömwall (2011) observaron que el CBCA y el RM (Reality Monitoring) no eran efectivos en la discriminación entre confesiones verdaderas y falsas. Sobre la base de esta máxima, de la falta de universalidad, se han elaborado protocolos de evaluación forense que engloban otras medidas como el daño psicólogico o la capacidad para testimoniar, tal como el Sistema de Evaluación Global (Arce y Fariña, 2006a, 2006b) con adaptaciones a casuísticas

concretas como la violencia de género (Arce y Fariña, 2009) o el acoso laboral (Arce y Fariña, 2011). Repasaremos el procedimiento más ampliamente difundido y estudiado, el Statement Validity Analysis (SVA), nacido y desarrollado en Alemania, el Reality Monitoring (RM), adaptado al estudio del testimonio por Alonso-Quecuty, y, una propuesta española ajustada a nuestro ordenamiento legal y jurisprudencia, el Sistema de Evaluación Global (SEG).

Statement Validity Analysis (SVA)

Aunque se han documentado antecedentes en Alemania sobre el uso de categorías de análisis de contenido anteriores como recogen Undeutsch (1954) y una sentencia del Tribunal Supremo de la República Federal Alemana (*Bundesgerichtshof*, BGH) de 1954, fue Udo Undeutsch (1967) quien creó el primer conjunto homogéneo y amplio de criterios de realidad para aplicar a declaraciones de menores víctimas de abusos sexuales, el Análisis de la Realidad de las Declaraciones (SRA), conformado por dos conjuntos de criterios (ver Tabla 3), uno referido a la declaración y otro a la secuencia de las declaraciones, pero en términos operativos para la evaluación de la realidad se clasifican en criterios de realidad (i.e., criterios generales y manifestaciones especiales), cuya presencia se relaciona con realidad de la declaración, y criterios de validez de las declaraciones (i.e., criterios negativos o de control, criterios derivados de la secuencia de las declaraciones), que informan sobre la validez de la prueba. La aparición de criterios de los dos primeros factores indica que la declaración es verdadera, pero su ausencia no implica que sea falsa; en tanto que la detección de los criterios de control y de consistencia restaría valor de verdad a la declaración.

El SRA se enmarca en un procedimiento más amplio, el SVA, que establece todo el protocolo a seguir: estudio inicial de todo el procedimiento judicial (anteriores declaraciones del menor, de otros testigos, del encausado, pruebas documentales, etc.); obtención de la declaración del menor por medio de una entrevista estructurada (se graba para posteriores análisis); transcripción de la entrevista; aplicación de los criterios de realidad y validez; y evaluación de conjunto.

Tabla 3. Criterios del SRA.

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Criterios derivados de la declaración
a) Criterios generales, fundamentales:
✓ Anclaje, fijación espacio-temporal (concreción de la acción en un espacio y tiempo).
✓ Concreción (claridad, viveza).
✓ Riqueza de detalles (gran cantidad de detalles en la narración).
✓ Originalidad de las narraciones (frente a estereotipos o clichés).
✓ Consistencia interna (coherencia lógica y psicológica).
✓ Mención de detalles específicos de un tipo concreto de agresión sexual.
b) Manifestaciones especiales de los criterios anteriores:
✓ Referencia a detalles que exceden la capacidad del testigo (que van más allá de su imaginación o capacidad de comprensión).
✓ Referencia a experiencias subjetivas: sentimientos, emociones, pensamientos, miedos, etc.
✓ Mención a imprevistos o complicaciones inesperadas.
✓ Correcciones espontáneas, especificaciones y complementaciones durante la declaración.
✓ Auto desaprobación (declaración en contra de su interés).
c) Criterios negativos o de control:
✓ Carencia de consistencia interna (contradicciones).
✓ Carencia de consistencia con las leyes de la naturaleza o científicas.
✓ Carencia de consistencia externa (discrepancia con otros hechos o pruebas incontrovertibles).
Criterios derivados de las secuencias de declaraciones
✓ Carencia de persistencia (estabilidad en el tiempo y contextos).
✓ Declaración inconsistente con la anterior.

En 1989, Steller y Kohnken (1989/1994), sobre la base de la hipótesis y del sistema de categorías de Undeutsch (1967) y otros posteriores de Arntzen (1970) y Littmann y Szewczyk (1983) y de sus propias experiencias y reflexiones personales, confeccionaron un nuevo sistema de categorías de realidad compuesto por 19 categorías

de realidad, el Criteria Based Content Analysis (CBCA), estructurado en torno a 5 dimensiones: características generales, contenidos específicos, peculiaridades del contenido, contenidos referidos a la motivación y elementos específicos de la agresión (en la Tabla 4 se relacionan los criterios junto con una breve definición de los mismos por dimensiones). Al igual que las categorías de realidad de Undeutsch, el registro de estos criterios de contenido se interpretará en el sentido de que la declaración es real, en tanto que de su ausencia no puede desprenderse que sea falsa. Este registro puede tomarse en términos de presencia o ausencia, o puntuarse en cuanto a fuerza o grado en que aparecen en la declaración. Originalmente, este sistema se diseñó para la evaluación de declaraciones de niños en casos de abuso sexual.

Tabla 4. Criterios de realidad del CBCA.

a) Características generales:
✓ Estructura lógica (coherencia y consistencia interna).
✓ Elaboración inestructurada (presentación desorganizada).
✓ Cantidad de detalles (abundancia de detalles o hechos distintos).
b) Contenidos específicos:
✓ Engranaje contextual (ubicación de la narración en un espacio y tiempo).
✓ Descripción de interacciones (cadena de acciones entre el testigo y otros actores).
✓ Reproducción de conversación (réplica de conversaciones).
✓ Complicaciones inesperadas durante el incidente (por ejemplo, interrupción imprevista).
c) Peculiaridades del contenido:
✓ Detalles inusuales (detalles con baja probabilidad de ocurrencia).
✓ Detalles superfluos (detalles irrelevantes que no contribuyen significativamente a los hechos).
✓ Incomprensión de detalles relatados con precisión (explicitación de detalles que el menor no comprende pero realmente sí tienen sentido).

✓ Asociaciones externas relacionadas (inclusión de información externa a los hechos en sí pero relacionada con ellos, tal como en una agresión sexual recordar conversaciones anteriores sobre este tema).
✓ Relatos del estado mental subjetivo (referencias a sentimientos, emociones o cogniciones propias).
✓ Atribución del estado mental del autor del delito (referencias al estado mental del agresor y atribución de motivos).
d) Contenidos referentes a la motivación:
✓ Correcciones espontáneas (correcciones espontáneas o mejoras de la declaración).
✓ Admisión de falta de memoria (reconocimiento de lagunas de memoria).
✓ Plantear dudas sobre el propio testimonio.
✓ Auto-desaprobación (actitud crítica sobre su propia conducta).
✓ Perdón al autor del delito (la declaración de la víctima favorece al acusado, o evitación de más acusaciones).
e) Elementos específicos de la agresión:
✓ Detalles característicos de la ofensa (descripciones que contradicen las creencias habituales sobre el delito).

Aparejado a este nuevo sistema categorial también se ha ajustado el listado de validez del SVA con el objeto de completar una evaluación de realidad del CBCA porque ésta, por sí misma, no es prueba suficiente. Así y a modo de ejemplo, puede que el entrevistador haya dirigido la entrevista de tal manera que la contaminara (p.ej., preguntas engañosas, sugestivas). Por ello, el listado de validez tiene como objeto completar la evaluación del CBCA con el estudio y análisis de hipótesis alternativas. El listado más conocido, seguido y presentado de forma coetánea al CBCA, es el de Steller (Steller, 1989; Steller y Boychuk, 1992), que puede consultarse en la Tabla 5. Vrij (2008) apuntó varias limitaciones a este listado de validez: dificultades en la identificación de las categorías (v.gr., ¿cómo se identifica si un menor ha sido entrenado para prestar un testimonio falso?), dificultades en la medida (p.ej., ¿cómo es medida la

susceptibilidad a la sugestión?), indeterminación del impacto de cada categoría en la evaluación resultante del CBCA, carencia de justificación de algunas categorías (e.g., adecuación del afecto, consistencia inter-declaraciones, consistencia con otras pruebas, consistencia con las leyes de la naturaleza), el listado no es completo (p.ej., las declaraciones sobre eventos familiares para los entrevistados contienen más criterios de realidad del CBCA que aquellas sobre eventos que no le son familiares), uso inapropiado del listado, y dificultad en la detección de falsas memorias, y declaraciones falsas enmascaradas (ad exemplum, el testigo puede informar honestamente del delito denunciado, pero puede acusar, deliberadamente o no, a una persona inocente). Adicionalmente, ha de tenerse en mente que varias categorías se han confeccionado para menores. Sin embargo, y con independencia de que este listado de validez sea limitado, concreto para una población, con alguna categoría difusa en su definición y con problemas de medida, el estudio de las hipótesis alternativas, a semejanza del diagnóstico clínico diferencial, es ineludible pues el estudio de las categorías de realidad no resulta suficiente para validar una memoria como real de un hecho concreto denunciado y que, por ejemplo, podría confundir una memoria falsa enmascarada en una real, con una totalmente real.

En cuanto al criterio de decisión, los autores no han formulado uno específico, sino que se limitan a dar guías y a un modelo de decisión clínica. Así, Undeutsch (1967, 1989) propone una evaluación de conjunto, en la que cada criterio de realidad tenga un peso limitado en la formación de un juicio categórico sobre la realidad de la memoria (real v. no real) o del grado en que una declaración representa algo vivido por el testigo. Al respecto, aporta cuatro *guidelines* orientativas para la determinación de si la declaración describe un evento real o no:

- La intensidad o grado de las manifestaciones en los diferentes criterios.
- El número de detalles de la narración que se relacionan con un criterio (o más).
- Las capacidades del declarante para informar (edad, inteligencia, sugestión, etc.).
- Las características del evento narrativo (complejidad, relevancia, etc.).

Steller (1989) advierte del carácter subjetivo, comparándolo con el modelo de decisión clínica, del juicio sobre la validez de la declaración, proponiendo el ajuste del juicio a una escala probabilística (creíble, probablemente creíble, indeterminado,

probablemente increíble, increíble). Además, señala que es posible combinar el estudio de la realidad de la declaración con las categorías del CBCA con indicadores no verbales (ver en la introducción a este trabajo qué indicadores no verbales y paraverbales se asocian con la mentira). Asimismo, Köhnken (2004) también asume un modelo de decisión clínica en la formación global del juicio en el que se sigue un esquema decisional de verificación de hipótesis alternativas a la realidad del testimonio (p.ej., sugestión, instrucción del testimonio, enfermedad mental, fabricación parcial de la declaración); si alguna de éstas se acepta, se concluye que "no se puede confirmar que la declaración esté basada en una experiencia propia del testigo", esto es, que esté basada en hechos vividos. De nuevo, de la aceptación de una de estas hipótesis no se puede deducir que la declaración sea falsa. Si se descartan todas las hipótesis alternativas, se repara en los criterios de realidad registrados en la entrevista, siguiendo un modelo clínico en la formación del juicio sobre la base de las habilidades cognitivas del testigo. De no llegarse a un juicio de que la declaración está basada en hechos vividos, no es necesario proceder al estudio de las hipótesis alternativas porque no sería posible validar la realidad de la declaración.

Tabla 5. Listado de validez del SVA.

a) Características psicológicas:
✓ (In)adecuación del lenguaje y conocimientos. El testigo informa con un vocabulario y conocimientos impropios de su edad y desarrollo o que van más allá de lo que puede haber aprendido en el incidente.
✓ (In)adecuación del afecto. (In)adecuación de las expresiones afectivas mostradas por el testigo (generalmente por medio del comportamiento no verbal). Si éstas no son indicativas de victimación del delito puede indicar que la declaración ha sido inventada.
✓ Susceptibilidad a la sugestión. Observación de que el testigo es susceptible a la sugestión.
b) Características de la entrevista:
✓ Entrevista coercitiva, sugestiva o dirigida. Estudiar si el

entrevistador usó técnicas sugestivas, dirigidas o coercitivas.
✓ (In)adecuación global de la entrevista. Por ejemplo, remarcar al testigo que, en su caso, debe responder "No sé". Así, se evitan dos tipos de errores la admisión de falta de memoria o una respuesta inventada ante la presión de responder a la pregunta.
c) Motivación:
✓ Motivos para la denuncia. Estudio de los motivos para denunciar (p.ej., disputa de la custodia).
✓ Contexto en que se produjo la denuncia o revelación original. Por ejemplo, los hechos se descubrieron de forma casual o fueron por un interrogatorio de los padres, psicólogo u otro (considerar si tiene algún interés en la causa).
✓ Presiones para presentar una denuncia falsa. Existencia de indicios de que otros sugirieron, presionaron, entrenaron o coaccionaron al testigo para que presente una denuncia falsa o exagere ciertos elementos en una denuncia de un hecho verdadero.
d) Cuestiones de la investigación:
✓ (In)consistencia con las leyes de la naturaleza. Descripción de eventos que son imposibles.
✓ (In)consistencia con otras declaraciones. Falta de consistencia en elementos centrales, no periféricos, de la declaración del testigo con otras previas o de otros testigos.
✓ (In)consistencia con otras pruebas. Falta de consistencia entre elementos centrales de la declaración con otras pruebas fiables.

El Reality Monitoring (RM)

En 1992, Alonso-Quecuty llevó a cabo el primer experimento en que se sometía a prueba la validez del Reality Monitoring (RM) (Johnson y Raye, 1981) para diferenciar entre testimonios verdaderos y falsos, dando lugar a una línea de investigación prometedora que se ha mantenido hasta la actualidad. El RM se creó para discriminar entre memorias externas (esto es, perceptuales) de experiencias vividas de

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

memorias internas (esto es, imaginadas). El modelo asume que los eventos vividos contienen información sensorial, por estar activada la memoria sensorial, tal como visual o auditiva, y contextual (e.g., espacio, tiempo), e información idiosincrásica y semántica. Por el contrario, en las memorias externas se predice una mayor presencia de operaciones cognitivas. Basándose en esta aproximación, se han constituido listados de categorías de realidad a semejanza del CBCA. En la Tabla 6 se recogen las categorías creadas partir del original de Johnson y Raye, y, en la Tabla 7, un listado ampliado de criterios, fruto de un análisis factorial de Sporer (1997) del *Judgement of Memory Characteristics Questionnaire*, una adaptación para *juzgar la memoria de los relatos de otras personas* del cuestionario original del RM, el MCQ (Johnson, Foley, Suengas, y Raye, 1988).

Tabla 6. Categorías de contenido originales del RM (Otgaar, Candel, Memon, y Almerigogna, 2010).

Información visual (cualquier acción u cosa que fue vista en el evento).
Información temporal (cualquier información relacionada con el tiempo de los eventos).
Información espacial (cualquier información que se relaciona con la localización o posicionamiento de ítems).
Información auditiva (habla o información relacionada con sonidos).
Información cognitiva (suposiciones, pensamientos, razonamientos, y atribuciones de intención).
Información afectiva (información sobre emociones y sentimientos).

A pesar del soporte teórico que sustenta las categorías de realidad del RM y que el RM sería más fácil de usar y consume menos tiempo (Sporer, 1997; Vrij, Edward, Roberts, y Bull, 2000), a lo que añadimos una mayor facilidad y consistencia en la codificación (especialmente si se siguen las escalas de Sporer), los estudios científicos no apoyan, en general, la efectividad de este sistema categorial, a pesar de Vrij (2008) obtuvo, en una revisión de la literatura, una tasa clasificatoria global de verdad y mentira superior al azar, del 68.80%. Así, en la mayoría de las investigacionesahora bien sólo algunos criterios discriminan entre verdad y mentira; en

varios, la puntuación total tampoco discrimina entre verdad y mentira; algunas categorías son escasamente útiles para esta tarea (p.ej., reconstrucción de la historia, información espacial); e incluso se han obtenido resultados contarios a la predicción del modelo (e.g., operaciones cognitivas asociadas a verdad) (Masip et al., 2005; Vrij, 2008). Este sistema categorial presenta dos limitaciones adicionales para la práctica forense: no incluye un estudio de la validez, ni un modelo de decisión, simplemente se trata de una hipótesis científica que no puede extrapolarse al campo forense en las condiciones en que se encuentra su formulación.

Tabla 7. Criterios ampliados del RM de Sporer (1997, 2004).

Claridad (claridad, viveza en vez de vaguedad).
Información sensorial (tal como sonidos, gustos o detalles visuales).
Información espacial (lugares, ubicaciones).
Información temporal (ubicación del evento en el tiempo, descripción de secuencias de eventos).
Afecto (expresión de emociones y sentimientos sentidos durante el evento).
Reconstrucción de la historia (plausibilidad de reconstrucción del evento tras la información dada).
Realismo (plausibilidad, realismo y sentido de la historia).
Operaciones cognitivas (descripciones de inferencias hechas por otros durante el evento).

El Sistema de Evaluación Global (SEG)

En España, se ha formulado un protocolo de evaluación psicológico-forense que va más allá de la evaluación de la realidad del testimonio, al incluir la evaluación del daño psicológico (huella psicológica del crimen), y la capacidad para testimoniar. Éste se puede aplicar en su totalidad o en parte, según el mandato judicial, a menores, adultos o discapacitados, y a otros delitos diferentes de los abusos sexuales. El Sistema de Evaluación Global presenta una estructura general que puede verse en la Tabla 8.

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Tabla 8. Estructura general del SEG (Arce y Fariña, 2006a, 2006b).

Obtención de la declaración.
Repetición de la declaración.
Estudio de la motivación.
Análisis de la validez de las declaraciones.
Análisis de la realidad de las declaraciones.
Medida de las consecuencias clínicas del hecho traumático.
Análisis de la fiabilidad de las medidas.
Evaluación de la declaración de los actores implicados.
Análisis de las características psicológicas de los actores implicados.
Implicaciones para la presentación del informe.

Obtención de la declaración

Para que el procedimiento de evaluación psicológico-forense de las declaraciones sea productivo, fiable y válido requiere de unos instrumentos de obtención de la declaración que permitan los subsecuentes análisis de la realidad de las mismas. Por ello, las declaraciones han de obtenerse, según se trate de adultos con plenas capacidades cognitivas o con discapacidades, por los siguientes procedimientos: *Entrevista Cognitiva Mejorada* (Fisher y Geiselman, 1992), o la *Entrevista Forense a Discapacitados* (Arce, Novo, y Alfaro, 2000), y en el caso de menores por la *Entrevista Cognitiva Mejorada* o, acorde a la naturaleza de la demanda judicial, y a la edad y capacidades del menor, el *Memorandum of Good Practices* (Home Office and Department of Health, 1992). Por su parte, la entrevista sobre el daño psicológico se obtiene por medio de la *Entrevista Clínico-Forense* (Arce y Fariña, 2001). Para el análisis de pruebas practicadas con otros protocolos de obtención de la declaración, serían válidos otros, siempre y cuando se basen en el recuerdo libre y en técnicas de ayuda de recuperación de la información (serían totalmente inválidos para el contexto forense si la técnica contaminara la información; por ejemplo, con preguntas sugestivas o dirigidas).Ver en el capítulo de este mismo manual, *La entrevista psicológica forense a niños, adultos y discapacitados*, otros protocolos y una discusión de sus limitaciones y fuentes de contaminación que el forense ha de tener en cuenta.

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Repetición de la obtención de la declaración

La jurisprudencia española ha establecido la *persistencia en la incriminación* como un criterio básico exigible al testimonio para dotarlo de credibilidad. Ésta se define en la Sentencia del TS (Sala de lo Penal), de 28 septiembre 1988, RJ 7070 como que "ésta ha de ser prolongada, en el tiempo plural, sin ambigüedades ni contradicciones, con arreglo a los clásicos". Asimismo, desde una óptica científica, la consistencia en el tiempo de la medida es un indicador de la fiabilidad de la misma (e.g., en psicometría test-retest; en análisis de contenido la consistencia intra-sujeto). Al respecto, los procedimientos tradicionales de análisis de contenido de las declaraciones, tal como el SVA, obtienen una única declaración, sometiendo a prueba la consistencia en el tiempo de los contenidos de la declaración tomando, en su caso, otras declaraciones obrantes en el procedimiento (i.e., policiales, judiciales). Ahora bien, estas declaraciones no son válidas para el contexto forense porque no proporcionan toda la información precisa para el análisis de contenido de la realidad del evento; la información obtenida no es totalmente aséptica, es decir, libre de cualquier fuente de contaminación (ver el capítulo *La entrevista psicológica forense a niños, adultos y discapacitados* en este mismo manual). La no constatación de la persistencia en la incriminación por otro medio que no sean las declaraciones obrantes en el procedimiento, incluso ha llevado a que, en alguna ocasión, se desestimara el valor de prueba del SVA, frente al SEG, con esta motivación (v.gr., AP de Pontevedra, Sección 6ª, de 21 de enero de 2004). Además, los protocolos de entrevista anteriormente reseñados se pueden aplicar en series repetidas, sin que contaminen la información (Campos y Alonso-Quecuty, 1999; Memon, Wark, Bull, y Köhnken, 1997). Pero, ¿qué se entiende por consistencia en la narrativa de un evento en el contexto del análisis de las declaraciones? Sobre esto, Undeutsch (1967, p. 125) formuló la hipótesis de la centralidad/periferia del material que entra en contradicción. Así, señala que sólo es relevante la contradicción si afecta a detalles centrales para la acción de juicio. Esto es, la inconsistencia en la información periférica, o la omisión de cierta información, sólo es importante si es trascendente para la (re)construcción del evento. Esta hipótesis se completa con la hipótesis Trankell (1963/1972): la mentira deliberada es más consistente que lo vivido. El SEG, para dar cabida a las interferencias (teoría de la interferencia del olvido), a la entrada de nueva información (hipótesis constructiva del olvido) y a la curva

del olvido, establece el tiempo a transcurrir entre entrevista y entrevista en superior a una semana. Las hipótesis básicas que asume son tres. Primera, al ser un evento vital estresante el efecto del desuso será menor (referido al testimonio de víctima y agresor). Segunda, una teoría de racionalidad por parte del mentiroso que se plasma en que la mentira es planificada, aprendida y, por extensión, consistente en el tiempo, con lo que no estará, o, de estarlo, sería en pequeña medida, influenciada por interferencias e información post-suceso (hipótesis constructiva). Por ello, el SEG aconseja obtener la primera declaración en todas aquellas fases tomadas en formato de discurso libre sin ningún interrogatorio, porque a través de éste entraría información post-suceso que el entrevistado podría acomodar a la nueva reconstrucción. El interrogatorio se sugiere que se ejecute tras la obtención en discurso libre de la segunda declaración. Tercera, el entrevistado honesto, si se siguen las técnicas prescritas en los protocolos de entrevista mencionados y se dirige adecuadamente en la entrevista, narra imágenes e información sensorial, con lo que la descripción de los hechos aún siendo muy semejante, será de construcción diferente al no estar sustentada en esquemas episódicos (las técnicas practicadas en las entrevistas reducen los efectos en el recuerdo de los conocimientos previos, las expectativas y los esquemas o guiones). En suma, en formato de recuerdo libre la declaración verdadera será menos consistente y aunque el evento sea el mismo, la narración será significativamente distinta tanto en su recuperación como en el contenido (v. gr., omisiones, elicitación de nueva información distinta a los hechos, pero relacionada con ella, inconsistencia en información periférica, recuperación/pérdida de nueva información poco relevante para los hechos). Por su parte, el sujeto mentiroso narra historias aprendidas con lo que las repetirá en recuerdo libre básicamente iguales, guiado por un esquema episódico (hipótesis Trankell).

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 (e. g., procedimiento, sumario, diligencias); el contexto en el que se presenta la denuncia; y los motivos/intereses para presentar una denuncia falsa. Para este estudio es preciso disponer de todas las declaraciones del procedimiento judicial. Por ello, el forense ha de tener acceso a todo el procedimiento judicial. La utilidad se justifica por dos motivos. Primero, para

facilitar la entrevista. El conocimiento de los hechos denunciados y otros pormenores del caso facilitan la entrevista y, frente a la manifestación del desconocimiento de los hechos, el establecimiento de un buen clima con el entrevistado. Segundo, contrastar la persistencia (tiempo, hechos y contextos) y consistencia (en elementos centrales) entre las declaraciones policiales y judiciales con las forenses. Si éstas últimas son consistentes y persistentes entre sí, el valor de la falta de persistencia y consistencia con las judiciales y policiales es muy relativo, y el forense ha de manejarlo correctamente. De este modo, la carencia de persistencia en casos de violencia contra la mujer se observa con una cierta frecuencia sin que, por ello, la declaración sea inventada. Por ejemplo, en la declaración forense, a veces, se obtiene evidencia de que fue objeto de agresiones sexuales que no denunció al no ser consciente que sufrió las mismas. Asimismo, en el caso de menores, en ocasiones, no son éstos los que prestan la declaración de denuncia, sino los progenitores u otros. Esto puede llevar a que se registre falta de persistencia o consistencia entre las declaraciones originales y las forenses, que no es tal. Por su parte, la validez de los interrogatorios policiales y judiciales para el estudio de la persistencia en la incriminación no es definitiva porque este tipo de interrogatorios contaminan el testimonio (ver el capítulo *La entrevista psicológica forense a niños, adultos y discapacitados* en este mismo manual). Así pues, el forense ha de gestionar adecuadamente la falta de persistencia y consistencia con las declaraciones policiales y judiciales. Otro criterio, que jueces y tribunales tienen en consideración para otorgar credibilidad al testimonio del denunciante, es la ausencia de incredibilidad subjetiva. Ésta se define [Sentencia del TS (Sala de lo Penal), de 28 septiembre 1988, RJ 7070] como "Ausencia de incredibilidad subjetiva, derivada de las relaciones procesado-denunciante, que pudieran conducir a la deducción de la existencia de un móvil de resentimiento, enemistad u otro tipo que privase al testimonio de aptitud para generar este estado subjetivo de certidumbre en que la convicción judicial estriba esencialmente". Por ello, el SEG incluye, en este paso, el estudio de los potenciales motivos o intereses (v.gr., intereses económicos, venganza) para la presentación de una denuncia infundada o falsa. Novo y Seijo (2010) hallaron, en un análisis de sentencias judiciales, que el contexto en que se produce la denuncia original constituye un momento crítico para la estimación de ésta, siendo los indicadores de referencia la espontaneidad y contigüidad en la presentación de la denuncia.

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Obviamente, la concurrencia de motivos o intereses en la denuncia no presupone que ésta sea infundada o falsa, pero el forense ha de tenerlos siempre presentes (estudio de hipótesis alternativas).

Análisis de la validez de las declaraciones

Previamente al estudio de la validez y realidad de la declaración es necesario establecer si ésta es una prueba suficiente para tal estudio. Así, la simple acusación de un delito no es suficiente para llevar a cabo un estudio de la realidad, se necesitan unos hechos para poder realizar el análisis. Igualmente, en una declaración breve la probabilidad de que aparezcan criterios de realidad es muy limitada. De esta manera, concluir que una declaración breve no es real es inadecuado, porque puede ser consecuencia de motivos independientes a la realidad de la misma como la vergüenza, la falta de cooperación del entrevistado o la carencia de habilidad del entrevistador para obtener una declaración amplia. En apoyo de esta hipótesis encontramos que el criterio *gran cantidad de detalles* es el que más apoyo experimental ha obtenido en la literatura experimental en la discriminación entre declaraciones reales e maginadas (Vrij, 2008). La memoria de lo imaginado es más corta a fin de poder aprenderla de memoria. Los ponentes del CBCA Steller y Köhnken, también, aunque no lo sistematizan, apuntan la necesidad de disponer de un material amplio para el análisis de la realidad de la declaración (Köhnken, 2004; Raskin y Steller, 1989; Steller, 1989). Como ya se puso de manifiesto previamente, la Jurisprudencia del Tribunal Supremo, en el criterio persistencia en la incriminación, remarca que la declaración ha de ser *prolongada*. En esta línea, el SEG se inicia el proceso con el análisis de la declaración con el estudio de la suficiencia de la prueba. Para ello, se contrastan dos hipótesis. Primera, la declaración para que pueda someterse a un análisis de la realidad de la misma ha de tener suficiente amplitud, esto es, contener un evento narrativo completo de los hechos (excepcionalmente se puede admitir la falta de partes del evento si es consecuencia de un traumatismo craneal o pérdida de conciencia del testigo). Segunda, dado que estamos en un contexto forense, es necesario contrastar si se trata de un evento narrativo aprendido. Para verificar si la *Prueba es (in)suficiente*, el Sistema de Evaluación Global se sirve de las siguientes cuestiones ¿Supera la capacidad de memoria del testigo? Si la supera, entonces no se descarta la hipótesis de que estemos ante la creación de un evento memorizado ¿Contiene toda la

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

información necesaria de los hechos? Además, las declaraciones es preciso que sean prueba válida para proceder con el estudio de la realidad de éstas. 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* (con esto damos satisfacción al criterio legal de persistencia en la incriminación) el perito o forense ha de cotejar si se cumple cada uno de los criterios de validez que se relacionan en la Tabla 9. De incumplirse alguno la prueba es inválida [las categorías de análisis fueron obtenidas de un estudio de la jurisprudencia española de Arce, Seijo, y Novo (2009) de modo que jueces y tribunales consideran una prueba inválida si se incumple alguno de estos criterios; es decir, las categorías de validez se "basan en la evidencia jurisprudencial"].

Tabla 9. Criterios de validez del SEG.

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, hechos y 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, es decir, son físicamente imposibles?).

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 que, de la desestimación de la prueba, por ser insuficiente o inválida, no se puede inferir que sea falsa.

Análisis de la realidad de las declaraciones

Los protocolos de análisis de contenido basados en categorías que discriminan entre memoria de lo percibido (real) y lo imaginado (no real) fueron concebidos, en

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

principio, para el testimonio de menores víctimas de agresiones sexuales pero, como ya advertimos anteriormente, son igualmente efectivos con adultos' en secuencias de medidas, y en casuísticas más allá de los abusos sexuales. En la adaptación a estos nuevos contextos, obviamente, no todas las categorías son productivas. Así, Landry y Brigham (1992) limitan el uso a 14 categorías con adultos, porque advirtieron que tres de ellas (incomprensión de detalles relatados con precisión; perdón al autor del delito; y detalles característicos de la agresión) sólo eran aplicables a menores, mientras que otras dos (elaboración inestructurada y asociaciones externas relacionadas) no eran productivas. Sin embargo, se ha encontrado que la categoría *perdón al autor del delito* era productiva, en declaraciones de adultos, tanto en agresiones sexuales como en amenazas (Arce, Fariña, y Freire, 2002), en tanto en un estudio se halló que las categorías *detalles característicos de la agresión* e *incomprensión de detalles relatados con precisión* eran productivas (Arce, Fariña, y Vivero, 2007), en tanto en otro no (Arce, Fariña, y Vilariño, 2010). En consecuencia, todas las categorías pueden ser productivas en otros casos diferentes a los abusos sexuales y en poblaciones de adultos. En la limitación del número de categorías en otros contextos diferentes a los de abusos sexuales a menores, no se han considerado variables que median su productividad como la casuística, las particularidades de la acción a examinar o el perfil sociodemográfico del entrevistado. Por ello, un sistema universal, esto es, aplicable a todo tipo de casos y poblaciones (incluidos casos de abusos sexuales a menores), debería incluir inicialmente todas las categorías de realidad. De no ser productivas o aplicables a poblaciones y casos específicos, el propio procedimiento de análisis de contenido no las registraría por lo que, en éstos, su peso sería nulo, pero medible en otros. A su vez, la combinación de los diferentes sistemas categoriales es posible y efectiva, ya que pueden sumar sus efectos (e.g., Arce, Fariña, y Freire, 2002; Arce, Fariña, y Vivero, 2007; Spörer, 2004; Vrij, 2008; Vrij, Edward, y Bull, 2001; Vilariño, Novo, y Seijo, 2010). En concreto, la combinación del *Criteria Based Content Analysis* (CBCA) (Steller y Köhnken, 1989/1994) y los criterios ampliados del *Reality Monitoring* (RM) (Spörer, 1997) mejora ligeramente la fiabilidad del sistema (en simulaciones), producto de añadir al *CBCA* los criterios *información perceptual* y *operaciones cognitivas* del *RM* (Vrij, 2000, 2008). De tal modo que el RM no debe entenderse como una alternativa al CBCA, sino como complementario de éste (Sporer, 2004). De este estado

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

de la cuestión se desprende que es factible un sistema combinado de todos los criterios de realidad de los diferentes sistemas propuestos en la literatura, y que se pueden definir criterios propios de delitos y poblaciones específicas. Para la creación de un sistema combinado, partimos de las categorías de realidad (se excluyeron las de validez) de los sistemas de Arntzen (1970), Szewczyk (1973), Trankell (1963/1972), del *CBCA* (Steller y Köhnken, 1989/1994), *RM* ampliado (Spörer, 1997, 2004) y del *Statement Reality Analysis* (SRA) (Undeutsch, 1967, 1988). Como quiera que estos sistemas se han creado a partir de la experiencia de los investigadores, esto es, *top-down*, completamos el procedimiento con la búsqueda de categorías a partir del análisis de contenido de declaraciones reales, *bottom-up*, por medio de la técnica de *aproximaciones sucesivas*. Por este medio, identificamos varias categorías de realidad para casos de violencia de género (ver Arce y Fariña, 2009) y una universal *síntomas sutiles del daño psíquico*, esto es, la aparición espontánea en la declaración de sintomatología clínica propia de la victimación no accesible a la simulación (ver cómo se identificaron y una relación de estos síntomas en Arce, Fariña, Carballal, y Novo, 2006, 2009; Arce, Pampillón, y Fariña, 2002). A fin de garantizar una independencia de las categorías, tal y como requiere un sistema metódico, éstas fueron evaluadas en el grado de solapamiento por jueces en tarea *thurstone*. Partiendo de las categorías del *CBCA*, hallamos que las categorías *concreción* (claridad, viveza en vez de vaguedad) y *originalidad de las narraciones* (frente a estereotipos) del *SRA* eran complementarias y distintas de las originales; y que la categoría del *CBCA correcciones espontáneas* era parte de una más amplia del *SRA*, *correcciones espontáneas, especificaciones y complementaciones de la declaración*. Adicionalmente, encontramos que categorías del RM y de CBCA se sobreponían de modo que la mera adicción de categorías de ambos sistemas llevaría acarreado un error de medida, al registrarse la misma contingencia en dos o más categorías (duplicidad de medidas). Por ello, las categoría *estructura lógica* del CBCA pasó a incluir las categorías *reconstrucción de la historia* y *realismo* del RM; *engranaje contextual* del CBCA, las categorías *información espacial y temporal* del RM; y la categoría *concreción* del SRA, la *claridad* del RM. Asimismo, Sporer (1997, 2004) obtuvo correlaciones significativas entre estas categorías del RM y del CBCA y estructuras factoriales que avalan esta recategorización. De este modo, resultó un sistema categorial compuesto por 24 categorías que se relacionan en la Tabla 10.

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Este sistema metódico se mostró productivo, fiable y válido (Vilariño, Novo, y Seijo, 2010) y significativamente más efectivo que el CBCA en casos de violencia de género (Arce, Fariña, y Vilariño, 2010). En otros tipo de casuísticas y poblaciones, cuando menos, debería ser de igual eficacia que el CBCA, al conformarse sobre la base de las categorías de éste. Es una tarea de futuro contrastar la validez incrementada de las categorías de realidad del SEG en relación a las del CBCA en otras casuísticas y poblaciones.

Tabla 10. 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)*.
2.- Elaboración inestructurada*.
3.- Cantidad de detalles*.
4.- Concreción (SRA, incluye claridad del RM)*.
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).
10.- Operaciones cognitivas (RM, codificación inversa. Se toma la definición de Vrij y sus colegas (ver Vrij, 2008) que amplía la definición original de las descripciones de inferencias hechas por el testigo, a las descripciones de inferencias realizadas por el testigo en la descripción del evento. Esta definición se ha mostrado más efectiva que la original y con resultados conforme a las predicciones del modelo).
PECULIARIDADES DEL CONTENIDO. Se incluyen aquí aquellas características de una declaración que aumentan su concreción o viveza.
11.- Detalles inusuales.

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

12.- Detalles superfluos.
13.- Incomprensión de detalles relatados con precisión.
14.- Originalidad de las expresiones (SRA)*.
15.- Asociaciones externas relacionadas.
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 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 (p.ej., no es consciente de determinados tipos de agresiones, tal como, cuando se le pregunta por las relaciones sexuales describe relaciones sin consentimiento pero no es consciente de que eso es un delito).
DAÑO PSICOLÓGICO DERIVADO DE LA ACCIÓN OBJETO DE JUICIO.
24.- Síntomas clínicos sutiles que forman parte de la huella psíquica.

Nota. * Criterios que discriminaron significativamente entre declaraciones reales y fabricadas en casos de violencia de género. En las declaraciones inventadas se han llegado a registrar hasta 5 de estos criterios por lo que para concluir que una declaración está basada en hechos vividos se requiere la presencia de más de 5 de estos criterios (Vilariño, Novo, y Seijo, 2010).

Medida de las consecuencias clínicas del hecho traumático

Las consecuencias para la víctima de un delito, la victimización, presenta cinco tipos de lesiones (United Nations, 1988): física, mental, sufrimiento emocional, pérdida o daño material y pérdida o menoscabo en los derechos. La lesión mental y el sufrimiento emocional son la denominada huella psicológica. Ésta se obtiene a través de la medida de los efectos de un acto delictivo en la salud mental que se ha relacionado tanto con síntomas

internalizados (e.g.., depresión, ansiedad, quejas somáticas, inadaptación social) como con externalizados, esto es, problemas de conducta (v. gr., agresividad, bullying, delincuencia), ya que puede producir ambos tipos de huella, o una interacción de ambas (p. ej., Edleson, 1999). No obstante, en la práctica forense se recurre a los internalizantes porque la victimización no presenta una relación causal directa con los problemas conductuales, sino de factor de riesgo (Baldry y Winkel, 2004), con lo que no se cumpliría un requisito legal imprescindible para la estimación de la prueba (el establecimiento de una relación causa-efecto). Se ha tomado el Trastorno de Estrés Postraumático, y bajo determinadas condiciones el Adaptativo (ver Arce y Fariña, 2009, 2011 para casos excepcionales), como el trastorno de referencia de la huella psicológica, al comprobarse sistemáticamente que estaba relacionado con la exposición a la victimización de delitos como agresiones sexuales (Breslau, Davis, Andreski y Peterson, 1991; Resnick, Kilpatrick, Dansky, Saunders, y Best, 1993), o violencia (Bargai, Ben-Shakhar y Shalev, 2007). Este trastorno suele aparecer asociado a otros, coocurrencia cifrada en los estudios de comorbilidad entre el 80 y el 85% de los casos (Creamer, Burguess, y Mcfarlane, 2001; Kessler, Sonnega, Hughes, y Nelson, 1995), siendo la depresión el que mayor índice de comorbilidad presenta, entre el 50 y el 60% (Blanchard, Hickling, Freidenberger, Malta, Kuhn, y 2004; O'Donnell, Creamer, y Pattison, 2004). Sin embargo, cuando se encuentra depresión u otro trastorno en ausencia de Trastorno de Estrés Postraumático, no se puede considerar una secuela del hecho traumático (O'Donnell, Creamer, Bryant, Schnyder, y Shalev, 2006). Identificada la naturaleza de la huella psicológica, el forense ha de tener presente dos premisas propias de la evaluación en este contexto: la sospecha de simulación (American Psychiatric Association, 2002), y la necesidad de establecimiento de una relación causa-efecto entre los hechos denunciados y la huella psicológica observada. Para ello se precisa de una aproximación multimétodo (Arbisi, 2005), mostrándose como las más efectiva aquella que combina entrevista con una medida psicométrica (Gothard, Rogers, y Sewell, 1995). A este respecto, la entrevista clínico-forense de Arce y Fariña (2001) posibilita identificar la causa de la sintomatología informada, no así la entrevista clínica estándar ni la evaluación psicométrica. Con estas premisas en mente, creamos y validamos un protocolo para la evaluación forense de la huella psíquica (Arce, Fariña, Carballal, y Novo, 2006, 2009; Arce, Pampillón y Fariña, 2002). Éste consiste en la adopción de una aproximación

multimétodo (MMPI-2[1], entrevista clínico-forense, y observación y registro conductual). El protocolo se basa en los siguientes criterios positivos de simulación: falta de consistencia inter-medidas, puntuaciones invalidantes en las escalas originales F y K de control de la validez del MMPI-2, la configuración F-K, un *perfil en V invertida* y la detección de estrategias de simulación en la *entrevista clínico-forense*. Para concluir que se verifica una huella psicológica se ha de constatar ésta inter-medidas y no hallar más de dos criterios positivos de simulación. En todo caso, el forense o perito psicólogo ha de tener presente las siguientes limitaciones para la redacción y defensa del informe: no toda agresión produce un Trastorno de Estrés Postraumático y la no presencia de un Trastorno de Estrés Postraumático no implica que la agresión no sea cierta, sino que no ha dejado, en su caso, huella psíquica. Asimismo, si se observan más de dos criterios de simulación el perito forense habrá de proceder al control de los falsos positivos (verdaderos enfermos informados como simuladores) mediante el uso de los criterios positivos, el registro de sintomatología no accesible a la simulación y el Modelo de Decisión Clínica de Cunnien (1997). Este protocolo presenta ligeras variaciones en los criterios de decisión según el tipo de caso que pueden consultarse en Arce et al. (2006, 2009) y Arce, Pampillón y Fariña (2002).

Análisis de la fiabilidad de las medidas

El sistema categorial metódico, como acabamos de demostrar, garantiza la disponibilidad de un instrumento de medida fiable y válido (Weick, 1985), pero no así de una medida fiable, tal y como se demanda en el contexto forense. Por ello, los sistemas de evaluación SVA y RM son, a la luz de la práctica forense, procedimientos de medida semi-objetivos (p. e., Arntzen, 1970; Undeutsch, 1967, 1988) o de decisión clínica (Köhnken, 2004; Steller, 1989). Para contrastar la fiabilidad de la medida concreta, el *SEG* incluye el sometimiento a prueba de la consistencia inter- e intra-evaluadores e inter-contextos (Wicker, 1975). Para ello se precisan dos evaluadores entrenados en la evaluación categorial, debiendo haber sido, al menos uno de ellos, efectivo y consistente en otros

1 El MMPI-2 puede sustituirte por el SCL-90-R a costa de une ligera merma en la capacidad de discriminación entre respuestas honestas y simuladas. El lector interesado puede encontrar el protocolo en Vilariño, Fariña y Arce (2009).

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

contextos previos, o sea, en pericias anteriores, con otros evaluadores. Los evaluadores ejecutarán la tarea por separado. De este modo se obtiene una medida de la consistencia inter-evaluadores e inter-contextos. Como herramienta estadística de análisis de la consistencia inter-evaluadores, el *SEG* toma el *Índice de Concordancia*, que es más restrictivo que los valores kappa, siendo el punto de corte ,80 (Tversky, 1977). La consistencia inter-e intra-medidas (la consistencia intra-medidas se obtiene de las medidas psicométricas y de entrevista clínico forense que contienen indicadores de engaño, esto es, (di)simulación que conforman los indicadores intra-medida) viene de la mano de los indicadores de la consistencia interna de las medidas (p. e., las escalas de validez de las pruebas psicométricas, plausibilidad de la declaración, estudio de las estrategias de simulación en la entrevista clínica, criterios de realidad y validez en las declaraciones); y de la consistencia entre distintas medidas (v. gr., concordancia entre medidas psicométricas y entrevista clínica, consistencia entre lo observado y lo manifestado, entre declaraciones, complementación o no [verbigracia, unas presentan indicios de veracidad y otras no, o no son prueba válida o suficiente] de las declaraciones y evaluaciones obtenidas del denunciante y denunciado). Los datos de medida de la consistencia se tomarán en unas hojas de registro creadas ad hoc, que incluyen todas las categorías de análisis de contenido con dos métodos de codificación: presencia de la categoría (e.g., las que se refieren a la declaración en su totalidad, esto es, las características generales) o codificación de las expresiones que dieron lugar a la codificación de la categoría (i.e., las categorías que se relacionan con partes específicas del testimonio). En el primer tipo de variables, la consistencia se mide sobre la base de la codificación del evaluador, presencia de la categoría vs. no presencia de la categoría, mientras que en la segunda no basta con el registro de la categoría, sino que es preciso verificar la correspondencia de las medidas. En otras palabras, el hecho de que se registre una determinada categoría del mismo modo por los dos evaluadores, no implica que las medidas se estén refiriendo exactamente a las mismas incidencias. Por ello, en la hoja de registro el evaluador habrá de anotar, en estas categorías, las expresiones que motivan la misma.

Evaluación de la declaración de los actores implicados

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Si bien en un principio la técnica fue creada para la evaluación del testimonio del denunciante o demandante, el mismo procedimiento de análisis de contenido de las declaraciones también es susceptible de aplicación al denunciado o demandado. Dos serían los potenciales campos de actuación en esta contingencia. Primero, la evaluación de las confesiones. Segundo, la evaluación de la declaración en la que el encausado presenta una versión distinta de los hechos al denunciante. En la evaluación de confesiones, la puntuación total de los criterios del CBCA y del RM (lo que lo haría extensivo a los del SEG) no discriminan entre confesiones reales y fabricadas, aunque sí alguno de los criterios (Willén y Strömwall, 2011). En estas circunstancias, adolece totalmente de valor de prueba concluyente. Por su parte, la evaluación de la declaración del encausado que contiene una versión de los hechos exculpatoria, debería tener la misma consideración que la del denunciante. Así, en ocasiones los hechos declarados por denunciante y denunciado son los mismos en parte o hasta cierto punto. Por ejemplo, una denuncia de agresión sexual que se inicia con una relación consentida y en la que el objeto de la decisión judicial va a estribar en si a partir de un determinado momento la relación fue o no consentida o se interrumpió. Obviamente, en estos casos ambas declaraciones son de interés forense por ser complementarias y, en ocasiones, los juzgados no sólo las autorizan, sino que las demandan. Aunque no existe literatura al respecto, pues esto no es aplicable en un sistema de adversarios, al no poder acceder el perito a los dos partes al mismo tiempo, se pueden generalizar los resultados de los hallazgos con el denunciante, pues de no ser real la acusación, la víctima sería el denunciado. Con esta doble evaluación del testimonio, cuando se demande o autorice judicialmente, obtendremos una estimación de la validación convergente de las declaraciones, lo que redunda en una mayor fiabilidad del sistema.

Análisis de las características psicológicas de los actores implicados

El estudio de las características psicológicas del encausado puede ser de suma trascendencia ya que en él podemos encontrar las claves explicativas de la agresión o cualquier enfermedad mental con implicaciones jurídicas relevantes. De ser el caso (en ocasiones el juzgado demanda que se evalúe el estado mental del encausado), se procederá con el protocolo de la evaluación de la salud mental con control de la (di)simulación (véase Arce y Fariña, 2006c; Arce, Pampillón, y Fariña, 2002). Adicionalment, y en respuesta al criterio jurisprudencial de verosimilitud [Sentencia del

TS (Sala de lo Penal), de 28 septiembre 1988, RJ 7070], esta evaluación tendrá por objeto informar sobre la ausencia o presencia de problemas psicológicos que afecten o no a la credibilidad del denunciante y, excepcionalmente, del encausado (Novo y Seijo, 2010). Además, se evalúan las capacidades cognitivas [esto es, medida psicométrica de la capacidad de razonamiento-inteligencia, capacidad de expresión y narrativa, y, en su caso, una exploración neuropsicológica. Ver en Arce y Fariña (2007) las medidas e instrumentos], del denunciante y del denunciado, cuando sea requerida una evaluación de éste, a fin de conocer la capacidad para testimoniar y ser objeto de una evaluación psicológico-forense. Por último, se evaluará cualquier otra área psicológica que se demande desde el juzgado, tal como la cuantificación del daño causado [ver en Arce y Fariña (2007) un método de cuantificación y de evaluación forense] o evaluación del riesgo de violencia/peligrosidad [ver en Andrés-Pueyo y Echeburúa (2010) instrumentos y aplicación]. Por lo que se refiere a la demanda de perfiles (p.e., maltratador, víctima, pedófilo), el perito forense ha de tener presente que no son prueba de ello (la presencia de características propias del perfil no implica causalidad o realidad), sólo aportan información adicional.

Implicaciones para la presentación del informe

El sistema de evaluación forense prescrito en el SVA en torno a 5 categorías de respuesta, no es operativo pues crea incertidumbre en el decisor. En esta línea, el Tribunal Supremo [Sentencia de 29 de octubre de 1981, RA 3902] estableció la necesidad de que en el informe se concrete en una seguridad plena, no en una escala de probabilidad. Aun así, toda medida tiene un margen de error que el científico debe estimar y el forense conocer y gestionar. La concreción del margen de error de la prueba es una exigencia que el Tribunal Supremo de los Estados Unidos, en el caso *Daubert v. Merrell Dow Pharmaceuticals, Inc*, instituyó que toda prueba debe cumplir para considerarla científica. Por ello, en los estudios científicos los forenses han de restar especial atención a los márgenes de error de la medida. Ahora bien, en el caso que nos ocupa, la pericial psicológica de la realidad del testimonio, los márgenes de error científico y forense no son directamente intercambiables. Así, en términos científicos el criterio de decisión se fija en función de la capacidad global de la técnica para discriminar entre realidad y engaño, de modo que se cuantifica la eficacia global

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

(clasificaciones correctas + rechazos correctos/2) o se fija un punto de corte en las escalas de control de la simulación (Curva ROC). En otras palabras, se equiparan los errores tipo I y II, cuando en la pericial el error tipo II (falsos positivos, esto es, falsas aceptaciones como reales de casos falsos o inventados) ha de ser 0. Bajo esta contingencia, el SEG dejando el error tipo II en 0, no clasifica como casos reales aproximadamente un 10%, es decir, no puede garantizar la realidad del daño o del testimonio en el 10% de los casos reales (con los modelos del RM y CBCA esta tasa está en torno al 30%, según se desprende la revisión de Vrij de 2008, que no incluye estudios de campo). Todo ello sin la comisión de error tipo I (con los modelos del RM y del CBCA esta tasa está en torno al 35% y 30%, respectivamente, pero en los estudios con el CBCA no se contrastan las hipótesis alternativas a la realidad de la declaración que completa al CBCA con el listado de validez del SVA, por lo que habría que descontar al 30% de falsas alarmas, la tasa desconocida de aceptación de hipótesis alternativas entre las declaraciones falsas informadas por el CBCA como verdaderas). Las categorías de decisión del SEG sobre la credibilidad del testimonio son: *declaración (muy) probablemente real*; *declaración carente de criterios de realidad; declaración o prueba inválida* y, en su caso, *indeterminado* (también puede referirse como *prueba insuficiente*). Si la declaración contiene criterios de realidad para avalar ésta se concluirá que la acción es muy probablemente real. Por su parte, no se concluirá que la declaración es falsa si no se observan suficientes criterios de realidad, sino que carece de criterios de realidad (de la no presencia de estos criterios no se deriva que sea falsa ya que puede deberse a otros motivos). Si las declaraciones incumplen alguno de los criterios de validez, se definirán como inválidas y como indeterminadas o prueba insuficiente si carecen de la amplitud necesaria para un análisis de la realidad o son incompletas. Por lo que se refiere al daño, se concluirá si se ha constatado, si es compatible con la victimización de un delito (trastorno por estrés postraumático), si éste es consecuencia del delito denunciado (relación causa-efecto), y si (no) se han registrado indicios sistemáticos de simulación.

Conclusiones finales

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Sobre la base de todo lo anterior, se puede concretar que el análisis cognitivo de las declaraciones es el procedimiento más productivo en la discriminación entre declaraciones verdaderas e inventadas. Así lo reconocen nuestros juzgados en sus demandas a peritos y forenses. Como quiera que todos estos avances científicos han de trasladarse en la práctica forense a un caso concreto con unos peritos o forenses concretos, se han ido definiendo protocolos para afrontar las limitaciones que median la fiabilidad y validez de la evaluación. El primer punto crítico estriba en la entrevista. Si el entrevistador contamina los contenidos de la entrevista, tal como con preguntas sugestivas o coercitivas, la prueba podría carecer de valor legal por ello. Por el contrario, si el entrevistador no obtiene toda la información del caso, o bien se pierde definitivamente, o bien si se obtiene posteriormente podría incurrirse en falta de persistencia en el testimonio, lo que invalidaría la prueba. En consecuencia, la figura del entrevistador es clave y, por ello, ha de contar con las habilidades necesarias para dirigir una entrevista forense para el análisis de contenido de la declaración y para la evaluación de la huella psicológica o salud mental. Como puede constatarse en el capítulo de este mismo manual, *La entrevista psicológica forense a niños, adultos y discapacitados*, la entrevista de hechos estándar o policial y la entrevista clínica estándar, no son válidas para los propósitos forenses. Aun es más, algunas de sus técnicas son contrarias a la buena praxis forense. Así pues, el perito o forense que conduzca la entrevista ha de tener profundos conocimientos de la Psicología de las Declaraciones y Confesiones y haber desarrollado las habilidades que requieren las entrevistas forenses que se adquieren con un exhaustivo entrenamiento (y supervisión) a fin de controlar la tendencia natural en este tipo de profesionales a usar las técnicas de entrevista estándar. En segundo lugar, la fiabilidad y potencia de los sistemas de análisis de la realidad del testimonio y de la evaluación de la huella psicológica/salud mental, no garantizan la fiabilidad y validez de una evaluación forense específica. Por tanto, han de seguirse los procedimientos descritos en el SEG para ello. Tercero, tanto el SEG como el SVA, no así el RM, apuntan esquemas decisionales a seguir por peritos y forenses en sus evaluaciones a lo que el SEG aporta un criterio de decisión para el número de criterios. No obstante, en ningún caso se informa de criterio alguno para la toma de decisión sobre la presencia de un determinado criterio. Estos descansan en el evaluador; en nuestro caso, el perito o forense. Para la formación de estos juicios, el perito o

forense ha de tener un amplio conocimiento de Psicología del Testimonio (v.gr., memoria de testigos), Evaluación Psicológica (p.ej., evaluación de la capacidad para testimoniar), Psicología del Desarrollo, y haber cubierto un plan de entrenamiento y experiencia compartida (o supervisión) con otros expertos. Cuarto, la evaluación de la huella psicológica y de la salud mental requiere conocimientos en Evaluación Clínico-Forense, Psicopatología Forense y Neuropsicología Forense y haber completado un plan de entrenamiento y experiencia compartida (o supervisión) con otros expertos. En resumen, el perito o forense que lleve a cabo estas evaluaciones ha de ser un psicólogo forense con formación, entrenamiento y experiencia (o práctica supervisada) en este tipo de evaluaciones, no debiendo llevarse a cabo por psicólogos especializados en otros perfiles profesionales que no sea el forense o por otros profesionales.

Referencias

Akehurst, L., Köhnken, G., y Höfer, E. (2001). Content credibility of accounts derived from live and video presentations. *Legal and Criminological Psychology, 6*, 65-83.

Alonso-Quecuty, M. L. (1992). Deception detection and reality monitoring: A new answer to an old question? En F. Lösell, D. Bender, y T. Bliesener (Eds.), *Psychology and law. International perspectives*. (pp. 328-332). Berlín: Walter de Gruyter.

American Psychiatric Association. (2002). *DSM-IV-TR. Manual diagnóstico y estadístico de los trastornos mentales*. Barcelona: Masson.

Andrés-Pueyo, A., y Echeburúa, E. (2010). Valoración del riesgo de violencia: Instrumentos disponibles e indicaciones de aplicación. *Psicothema, 22*, 403-409.

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., 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.

Arce, R., y Fariña, F. (2006a). Psicología del testimonio y evaluación cognitiva de la veracidad de testimonios y declaraciones. En J. C. Sierra, E. M. Jiménez y G. Buela-Casal (Coords.), *Psicología forense: Manual de técnicas y aplicaciones* (pp. 563-601). Madrid: Biblioteca Nueva.

Arce, R., y Fariña, F. (2006b). 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: Consejo General de Poder Judicial.

Arce, R., y Fariña, F. (2006c). Propuesta de un protocolo válido y fiable para la evaluación psicológico-forense de la imputabilidad. En R. Arce, F. Fariña, E. Alfaro, C. Civera, y F. Tortosa. (Eds.), *Psicología jurídica. Evaluación e intervención* (pp. 59-65). Valencia: Servicio de Publicaciones de la Diputación de Valencia.

Arce, R., y Fariña, F. (2007). Cómo evaluar el daño moral consecuencia de accidentes de tráfico: Validación de un protocolo de medida. *Papeles del Psicólogo, 28*, 205-210

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. (2011). Evaluación forense del acoso moral en el trabajo (mobbing) mediante el Sistema de Evaluación Global. En M. Matos, R. Abrunhosa, y C. Machado (Coords.), *Manual de psicologia forense: Contextos, práticas e desafios* (pp. 375-398). Braga, Portugal: Psiquilibrios Ediçoes.

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). 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., y Freire, M. J. (2002). Contrastando la generalización de los métodos empíricos de detección del engaño. *Psicologia: Teoria, Investigação e Prática, 7*, 71-86.

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.

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. I, pp. 590-604). Cádiz: Asociación de Profesionales de la Psicología Social.

Arce, R., Novo, M., y Alfaro, E. (2000). La obtención de la declaración en menores y discapacitados. En A. Ovejero, M. V. Moral, y P. Vivas (Eds.), *Aplicaciones en psicología social* (pp. 147-151). Madrid: Biblioteca Nueva.

Arce, R., Pampillón, M. C., y Fariña, F. (2002). Desarrollo y evaluación de un procedimiento empírico para la detección de la simulación de enajenación mental en el contexto legal. *Anuario de Psicología, 33*, 385-408.

Arce, R., Seijo, A., y Novo, M. (2009). Validez del testimonio: Un estudio comparativo de los criterios legales y empíricos. *Anuario de Psicología Jurídica, 19*, 5-13

Arntzen, F. (1970). *Psychologier der Zeugenaussage* [*Psicología del testimonio*]. Goettinger: Hogrefe.

Baldry, A. C., y Winkel, F. W. (2004). Mental and physical health of Italian youngsters directly and indirectly victimized at school and at home. *The International Journal of Forensic Mental Health, 3*, 77-91.

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.

Blanchard, E. B., Hickling, E. J., Freidenberger, B. M., Malta, L. S., Kuhn, E., y Sykes, M. A. (2004). Two studies of psychiatric morbidity among motor vehicle accident survivors 1 year after the crash. *Behaviour Research and Therapy, 42*, 569-583.

Breslau, N., Davis, G., Andreski, P., y Peterson, E. (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives of General Psychiatry, 48*, 216-222.

Bull, R. (1997). Entrevistas a niños testigos. En F. Fariña, y R. Arce (Eds.), *Psicología e investigación judicial* (pp. 19-38). Madrid: Fundación Universidad Empresa.

Campos, L., y Alonso-Quecuty, M. L. (1999). The cognitive interview: Much more than simply "try again". *Psychology, Crime and Law, 5*(1-2), 47-59.

Creamer, M., Burguess, P., y Mcfarlane, A. C. (2001). Posttraumatic stress disorder: Findings from the Australian national survey of mental health and well-being. *Psychological Medicine, 31*, 1237-1247.

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: Guilford Press.

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.

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

Ekman, M., y O'Sullivan, M. (1994). Riesgos en la detección del engaño. En D. C. Raskin (Ed.), *Métodos psicológicos en la investigación y pruebas criminales* (pp. 253-280). Bilbao: Desclée de Brouwer (Orig. 1989).

Fisher, R. P., y Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview.* Springfield, IL: Charles C. Thomas.

Goodman, G. S. (2005). Wailion babies in her wake. *American Psychologist, 60*, 872-881.

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

Granhag, P. A., Strömwall, L. A., y Landström, S. (2006). Children recalling an event repeatedly: Effects on RM and CBCA scores. *Legal and Criminological Psychology, 11*, 81-98.

Home Office and The Department of Health. (1992). *Memorandum of good practice on video recorded interviews with child witnesses for criminal proceedings*. Londres: HMSO.

Judicial Studies Board. (2006). *Evidence*. Recuperado de http://www.jsboard.co.uk/criminal_law/cbb/mf_04.htm#19

Johnson, M. K., Foley, M. A., Suengas, A. G., y Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General, 117*, 371-376.

Johnson, M. K., y Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67-85.

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

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, UK. Cambridge University Press.

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.

Landry, K. L., y Brigham, J. C. (1992). The effect of training in criteria-based content analysis on the ability to of detect deception in adults. *Law and Human Behavior, 16*, 663-676.

Littmann, E., y Szewczyk, H. (1983). Zu einigen kriterien und ergebnissen forensisch-psychologischer glaubwürdigkeitsbegutachtung von sexuell misbrauchten kindern und jugendlichen [Sobre unos criterios y resultados de las evaluaciones psicológico forenses de credibilidad de niños y adolescentes víctimas de abuso sexual]. *Forensia, 4*, 55-72.

Masip, J., Sporer, S. L., Garrido, E., y Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology,*

Crime & Law, 11, 99-122.

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

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.

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.

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: Springer.

O'Donnell, M. L., Creamer, M., y Pattison, P. (2004). PTSD and depression following trauma: Understanding comorbidity. *American Journal Psychiatry, 161*, 1-7.

Otgaard, H., Candel, I., Memon, A., y Almerigogna, J. (2010). Differentiating between children's true and false memories using reality monitoring criteria. *Psychology, Crime & Law, 16*, 555-566.

Porter, S., y Yuille, J. C. (1995). Credibility assessment of criminal suspects through statement analysis. *Psychology, Crime & Law, 1*, 319-331.

Porter, S., y Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior, 20,* 443-458.

Raskin, D. C., y Steller, M. (1989). Assessing credibility of allegations of child sexual abuse: Polygraph examinations and statement analysis. En H. Wegener, F. Lösel, y J. Haisch (Eds.), *Criminal behaviour and the justice system: Psychological perspectives* (pp. 290-302). Heidelberg: Springer-Verlag.

Resnick, H. S., Kilpatrick, D. G., Dansky, B. S., Saunders, B. E., y Best, C. L. (1993). Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology, 61,* 984-991.

Ruby, C. L., y Brigham, J. C. (1997). The usefulness of the criteria-based content analysis technique in distinguishing between truthful and fabricated allegations. *Psychology, Public Policy, and Law, 3*, 705-737.

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.

Sporer, S. L. (2004). Reality monitoring and detection of deception. En A. Granhag y L. A. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 64-102). Cambridge, UK. Cambridge University Press.

Sporer, S. L., y Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20*, 421-446.

Sporer, S. L., y Schwandt, B. (2007). Moderators of nonverbal indicators of deception. A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13*, 1-34.

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 Boychuck, T. (1992). Children as witness in sexual abuse cases: Investigative interview and assessment techniques. En H. Dent y R. Flin (Eds.), *Children as witnesses* (pp. 47-73). Chichester, UK: John Wiley and Sons.

Steller, M., y Köhnken, G. (1989/1994). 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).

Szewczyk, H. (1973). Kriterien der beurteilung kindlicher zeugenaussagen [Criterios para la evaluación de la validez de declaraciones de niños]. *Probleme und Ergebnisse der Psychologie, 46*, 46-66.

Trankell, A. (1963/1972). *Reliability of evidence. Methods for analyzing and assessing witness statements*. Estocolmo: Beckman (Orig. 1963).

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

Undeutsch, U. (1954). Die Entwicklung der gerichtspsychologischen Gutachtertätigkeit [El desarrollo de la actividad psicológico forense]. En A. Wellek (Ed.), *Bericht über den 19. Kongreß der Deutschen Gesellschaft für Psychologie in Köln 1953* (pp. 132-154). Göttingen: Hogrefe.

Undeutsch, U. (1967). Beurteilung der glaubhaftigkeit von zeugenaussagen [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: Kluwer Academic Publishers.

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

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. (2000). *Detecting lies and deceit.* Chichester, UK: 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* (2a. ed.). Chichester, UK: John Wiley and Sons.

Vrij, A., Edward, K., y Bull, R. (2001). Stereotypical verbal and non-verbal responses while deceiving others. *Personality and Social Psychology Bulletin, 27*, 899-909.

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.

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: Lawrence Erlbaum Associates.

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

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

Arce, R., y Fariña, F. (2013). Psicología forense experimental. Testigos y testimonio. Evaluación cognitiva de la veracidad de testimonios y declaraciones. En S. Delgado (Dir. Tratado), y S. Delgado, y J. M. Maza (Coords. Vol.), *Tratado de medicina legal y ciencias forenses: Vol. V. Psiquiatría legal y forense* (pp. 21-46). Barcelona: Bosch. **ISBN: 978-84-9790-694-4.**

Avance de publicación online. Recuperado el 25 de Agosto de 2011, de http://onlinelibrary.wiley.com/doi/10.1111/j.2044-8333.2011.02018.x

Zaparniuk, J., Yuille, J. C., y Taylor, S. (1995). Assessing the credibility of true and false statements. *International Journal of Law and Psychiatry*, *18*, 343-352.