

Artículo

Simulación y Validez de la Información de Síntomas Psicopatológicos en España: Conceptos, Métodos y Desafíos

Esteban Puente-López¹, David Pina², Julia C. Daugherty³, Miguel Pérez-García⁴ y Thomas Merten⁵

¹Departamento de Psicología, Universidad de Valladolid, España

²Departamento de Ciencias de la Educación, Universidad de la Rioja, España

³Departamento de Psicología, Université Clermont Auvergne, Francia

⁴Centro de Investigación Mente, Cerebro y Comportamiento (CIMCYC), Universidad de Granada, España

⁵Departamento de Neurología, Vivantes Klinikum im Friedrichshain, Berlin, Alemania

ARTICLE INFO

Recibido: 04-03-2024

Aceptado: 21-05-2024

Palabras clave:

Psicología forense

Exageración de síntomas

Validez de los síntomas auto-informados

Test de validez de síntomas

Evaluación psicológica forense

RESUMEN

Antecedentes/objetivo: El área de simulación de síntomas ha avanzado en los últimos años en España. No obstante, aún se necesita un impulso para alcanzar los recursos y estándares internacionales. Resulta de gran importancia disponer de guías y revisiones donde se incorporen recomendaciones basadas en la evidencia reciente, pero adaptadas al contexto español. Nuestro objetivo es realizar una revisión de la literatura de simulación de síntomas, y conceptos relacionados, en el contexto psicológico-forense español. Además, se ofrecerá una guía de trabajo actualizada para los profesionales del contexto. **Método:** Se desarrolló una revisión bibliográfica de la literatura nacional e internacional. **Resultados:** Se han identificado una serie de limitaciones en la terminología del área que se han tratado de solventar proponiendo un sistema de conceptos coherente con la literatura internacional. De la misma forma, se ha propuesto un sistema de evaluación completo, desde formulación de hipótesis, hasta componentes y planteamiento de conclusiones, orientado a la práctica profesional forense. **Conclusiones:** El estado de la literatura española evidencia carencias que dificultan alcanzar los estándares de las guías de buenas prácticas. Es importante fomentar la adaptación de nuevos test, y el desarrollo de nuevas investigaciones con los ya adaptados, en nuevos contextos y poblaciones de interés.

Malingering and Validity of Reported Psychopathological Symptoms in Spain: Concepts, Methods, and Challenges

ABSTRACT

Background/Objective: The area of malingering has advanced in recent years in Spain. However, a boost is still needed to reach and harness international resources and standards. It is of great importance to have guidelines and reviews that can incorporate and adapt recent evidence to the Spanish context. Our aim is to review the literature on malingering and related concepts, in the Spanish forensic-psychological context. In addition, an updated working guide for professionals in the context will be offered. **Method:** A bibliographic review of the national and international literature was carried out. **Results:** A series of limitations in terminology have been identified, for which we propose several suggestions that are coherent with the international literature. Similarly, we propose a complete evaluation system, from hypothesis formulation to components and conclusions, oriented towards forensic professional practice. **Conclusions:** The current state of Spanish literature in this area shows shortcomings that make it difficult to reach the standards of good practice. It is important to promote the adaptation of new tests and the development of research with those that have already been adapted, in new contexts and populations of interest.

Keywords:

Forensic psychology

Symptom exaggeration

Validity of self-reported symptoms

Asymptom validity test

Forensic psychological evaluation

Introducción

Las evaluaciones psicológicas realizadas tanto en el contexto clínico como en el forense dependen en gran medida de la presentación que las personas valoradas realicen de la presunta condición padecida, de su apertura y precisión al responder, y de su predisposición a realizar un esfuerzo sincero y continuado (Sherman et al., 2020). Por este motivo, las conclusiones obtenidas a través del proceso de evaluación psicológica son vulnerables a una posible alteración de los estilos de respuesta, tanto intencionada como no intencionada (Rogers et al., 2010). El/la profesional debe considerar que existe la posibilidad de que las personas evaluadas no ofrezcan una presentación de síntomas honesta, influidos por metas o motivos ajenos al diagnóstico o la condición padecida (Merten y Merckelbach, 2020).

Esta posibilidad es de obligado estudio en el contexto de evaluación forense porque, a diferencia de la evaluación clínica, el mandato judicial, en consonancia con el principio de presunción de inocencia (asuntos penales) o la carga de la prueba (asuntos civiles), implica que el perito ha de garantizar que los síntomas clínicos informados por la persona evaluada son ciertos (Arce y Fariña, 2015). Así, en el análisis de un posible cuadro clínico, como por ejemplo en un caso de estudio del daño psicológico a raíz de un evento traumático, será necesario desarrollar un proceso de diagnóstico diferencial de simulación donde se descarte una presentación intencionalmente distorsionada, o completamente inventada, de un cuadro psicopatológico relacionado con traumas.

Cuando la respuesta distorsionada se produce de manera intencional y voluntaria, y está orientada a la obtención de beneficios externos (legales o financieros), recibe el nombre de simulación de síntomas. El estudio y análisis de la simulación de síntomas ha sido objeto de interés tanto profesional como científico desde hace más de dos décadas (Merten et al., 2013, 2022; Sweet et al., 2021). España es uno de los países europeos que más ha avanzado en este campo en los últimos años. No obstante, a pesar de los esfuerzos y de los avances sustanciales realizados en el país, en la actualidad parece haber “más preguntas abiertas que respuestas” en el área (Merten et al., 2022, pág. 119). Si bien la disciplina avanzó significativamente en España entre los años 2005 y 2013, en la última década la producción científica en el área de validez de síntomas y simulación en psicología forense se ha reducido. Los manuales de referencia de la temática son el de González-Ordi et al. (2012) que, si bien su lectura es altamente aconsejable, fue publicado hace ya más de una década, y el de Arce y Fariña (2020), que desarrolla una técnica (en adelante técnica Uforense), científica y judicialmente validada, de evaluación forense del daño psicológico y de la salud mental con un diagnóstico diferencial de simulación (Arce et al., 2002, 2006, 2009, 2015; Fariña et al., 2014; Vilariño et al., 2009, 2013). Ahora bien, ésta última no es de aplicación general en los juzgados del país, ni en la formación en psicología forense de postgrado. En consecuencia, el contenido que se consume en parte de la formación universitaria, sobre todo cuando el docente no está especializado en la temática, no refleja los avances en la disciplina. De la misma forma, las técnicas habituales en la formación y de uso frecuente se han mostrado científicamente no válidas y están siendo desestimadas como prueba válida en juzgados y tribunales (Fariña et al., 2014).

Así, el objetivo que se persigue en el presente trabajo es hacer una revisión narrativa general del área de simulación de síntomas y validez

de los síntomas psicopatológicos auto-informados en el contexto psicológico-forense español. De manera secundaria, se pretende ofrecer una propuesta de guía de trabajo para los profesionales de la psicología forense que incorpore los estándares validados.

Simulación de Síntomas y Otros Conceptos Relevantes

Según la definición del Manual Estadístico Diagnóstico de los Trastornos Mentales, la simulación (en inglés *malingering*) es la “producción intencionada de síntomas físicos o psicológicos desproporcionados o falsos, motivados por incentivos externos, como evitar el servicio militar o el trabajo, obtener una compensación económica, evadir responsabilidades criminales u obtener fármacos” (American Psychiatric Association, 2023, pág. 835). Además, especifica que se debe sospechar (no diagnosticar) simulación cuando se observe la combinación (2 o más) de: 1) Contexto médico legal de la presentación, 2) discrepancia entre las quejas o discapacidad mencionada por el individuo y los hallazgos y observaciones objetivas del profesional, 3) falta de cooperación durante el proceso de valoración y para cumplir las pautas de tratamiento prescritas, 4) presencia de un trastorno de la personalidad antisocial.

Esta definición es, probablemente, una de las más utilizadas en la bibliografía nacional e internacional. A pesar de ello, ha sido objeto de varias revisiones, y los criterios de clasificación propuestos han sido ampliamente criticados. El motivo es que, tal y como están formulados, resultan de muy poca utilidad en un proceso de diagnóstico diferencial, dado que confunden características distintivas con frecuentes o comunes (Merten y Merckelbach, 2020). De hecho, la tasa de falsos positivos aplicando estos criterios es del 80% (Rogers y Vitacco, 2002), es decir, rinden peor que el azar, por lo que no deben usarse.

En la bibliografía internacional se han propuesto múltiples definiciones. Por ejemplo, Merten y Merckelbach (2020) señalan que la simulación es la presentación intencionadamente falsa (invención, exageración flagrante o cualquier otra tergiversación flagrante) de síntomas o problemas de salud con el objetivo principal de obtener una recompensa externa. La recompensa externa puede ser un incentivo económico (por ejemplo, una indemnización), pero también puede estar relacionada con la obtención de medicamentos, la evitación de un deber formal (por ejemplo, una baja por enfermedad, el servicio militar, la realización de un examen), la responsabilidad legal o un castigo, o la obtención de otra ventaja material o inmaterial. A éstas, en el contexto forense, hay que añadir cualquier causa que pueda motivar una denuncia falsa (p.ej., la venganza o resentimiento; Arce, 2017). El especificador *externo* indica que la motivación para la manifestación de una enfermedad falsa no está relacionada principalmente con la gestión del estrés, un conflicto interno (interpersonal), o el rol de enfermo. De esta manera, para establecer la presencia de simulación se debe disponer de evidencia convincente de que: (1) se trata de un acto totalmente intencional; (2) existe un incentivo externo sustancial subyacente a este comportamiento, y (3) no se identifican otros factores o condiciones que podrían explicar completamente la presentación distorsionada observada (Dandachi-FitzGerald et al., 2022). Además, ha de tener en cuenta que la presencia de dicho incentivo no es suficiente, debiendo examinarse también la percepción del/la examinado/a sobre el/ella mismo/a y valorar si es lo suficientemente motivador como para justificar el uso del término simulación (Rogers, 2018a).

Tradicionalmente, se han propuesto diferentes modelos que diferencian tipos de simulación. Así [Gamer \(1965\)](#), extendido por [Resnick \(1997\)](#), y referido al trastorno por estrés postraumático, divide la simulación en: a) parcial, que implica la exageración de síntomas existentes o de un trastorno ya superado; b) pura, que presupone la invención completa; y c) falsa imputación, que conlleva la atribución errónea de la sintomatología padecida a una causa que la persona sabe que no está relacionada con el síntoma. También pueden encontrarse trabajos españoles que siguen modelos menos utilizados en la bibliografía internacional. [Arce y Fariña \(2012\)](#), [Calcedo \(2000\)](#) o [Echeburúa et al. \(2011\)](#) diferencian entre simulación y sobresimulación. Simulación haría referencia a la invención total de sintomatología por parte de un sujeto sano, que ha de sospecharse en las evaluaciones forenses penales. Sobresimulación, por otro lado, implicaría la amplificación o magnificación de daño existente, que ha de sospecharse en las evaluaciones forenses civiles de compensación y reparación de daños ([Arce, 2017](#)). Dichas propuestas son de utilidad para evitar que los peritos adopten una interpretación simplista y extrema del mismo, es decir, que consideren que el único tipo de simulación posible es la simulación pura ([Wooley y Rogers, 2015](#)). No obstante, algunos de los subtipos planteados resultan altamente complicados de estudiar. Por ejemplo, la sobresimulación requiere tener la certeza de que la persona evaluada presenta una patología o alteración previa, lo cual resulta complejo en la gran mayoría de los diagnósticos psicopatológicos. Tras el pertinente proceso de evaluación se podrá determinar si la persona está simulando la sintomatología presentada o no, pero los medios y herramientas de valoración disponibles hoy en día no permiten establecer con precisión hasta qué punto se está exagerando (sobresimulación) algo ya existente, o se está fabricando en su totalidad. Por este motivo, con el fin de operativizar el concepto de simulación, y facilitar la toma de decisiones, no es recomendable la división en categorías, en tanto que en el contexto judicial (evaluación forense) ambas son simulación y se etiquetan como tales. En todo caso, el forense ha de conocer qué tipo de simulación se ha de sospechar en cada caso. Así, el término simulación integraría tanto la invención como la exageración flagrante, equiparándolo al concepto *malingering* usado en la bibliografía inglesa.

En una línea similar, los criterios originales de [Slick et al. \(1999\)](#) para simulación neuropsicológica recomendaban dividir el resultado de simulación en grados de certeza (probable, posible y definitiva), pero en la revisión de [Sherman et al. \(2020\)](#) eliminaron dicha categorización dado que resultaba compleja para la práctica profesional. En lugar de tres tipos diferentes de simulación, los criterios actuales utilizan un proceso de decisión binario (presente/no presente), y a continuación incluye especificadores relativos al tipo de simulación (de síntomas psicológicos, neuropsicológicos, etc.). Este procedimiento facilita el proceso de organización de información, formación de conclusiones y presentación de hallazgos.

Otras Explicaciones Para una Presentación Anómala

Una de las cuestiones más importantes a considerar es que no toda exageración, o ‘indicador’ de exageración, es sinónimo de simulación. Tradicionalmente, el foco de la investigación estaba centrado en el concepto de simulación, y gran parte de las investigaciones iniciales desarrolladas en el área han presentado la simulación como la única

explicación posible a una presentación de síntomas anómala. En la práctica profesional, se tiende a pensar que cuando una persona exagera sus síntomas está simulando ([Merckelbach et al., 2019](#)). En España, gran parte de la bibliografía científica, especialmente la publicada antes del 2013, se ha desarrollado en torno al concepto de simulación, y se presenta prácticamente como la única explicación. Esto ha llevado a que gran parte de las propuestas de evaluación estén centradas exclusivamente en este constructo, y, por ello, la mayoría de los peritos creen que las herramientas y técnicas desarrolladas miden simulación. Por ejemplo, los test psicométricos utilizados, como el Inventario Estructurado de Simulación de Síntomas (SIMS, [Widows y Smith, 2005](#)) o el Test de Simulación de Problemas de Memoria (TOMM, [Tombaugh, 1996](#)), son entendidos en la práctica profesional como ‘test de simulación’, y se considera que la obtención de resultados por encima o por debajo de los puntos de corte establecidos es un indicador (diagnóstico) de simulación, a pesar de la advertencia expresa de los manuales contra esta práctica (p.ej., [González-Ordi y Santamaría, 2009](#)). Esto se trata de un error conceptual, ya que ningún test psicométrico tiene la capacidad de determinar la presencia de simulación de síntomas ([Rogers, 2021](#)). O, dicho de otra manera, un resultado positivo en el SIMS, o en el TOMM, no es sinónimo de simulación, dado que dichos test no miden ese constructo. Realmente los instrumentos psicométricos advierten de anormalidad (concepto estadístico) en los protocolos de repuestas, siendo posible varias hipótesis explicativas, además de simulación, como una actitud no cooperativa, daño grave, estado de confusión, o protocolo inválido ([Arce, 2018](#); [Graham, 2011](#); [Greene, 2011](#); [Merckelbach et al., 2019](#)). Una interpretación en términos de falsedad (por ejemplo, simulación) requiere inferencias adicionales sobre la intencionalidad y los motivos detrás de estos comportamientos. En el ámbito de evaluación forense la hipótesis alternativa explicativa de obligado examen de anormalidad en las respuestas es la de daño grave, tal que con frecuencia casos genuinos de daño en la evaluación forense son clasificados por estos test como respuestas anormales ([Arce, 2018](#)). Por su parte, en la evaluación clínica las hipótesis con una mayor prevalencia son las de un patrón de respuesta desatenta ([Ward y Meade, 2022](#)) o los motivos facticios ([Chafetz et al., 2020](#)).

Así pues, la simulación de síntomas no es la única fuente de distorsión de respuesta que debe evaluarse en el contexto forense, pero ¿qué diferencia hay entre, la citada simulación, y otras posibles explicaciones? Para poder responder correctamente esta pregunta es necesario conocer tres conceptos adicionales: evaluación de la validez de la información de síntomas psicopatológicos¹, el sesgo de respuesta negativa y el fingimiento. Imaginemos, por ejemplo, un caso de valoración forense de daño psicológico. El perito desarrollará un proceso de evaluación psicológica donde el evaluado manifestará su sintomatología, ya sea verbalmente —en el proceso de entrevista— o por escrito —en los test psicométricos. La hipótesis principal que se barajará será la de que el paciente tiene un cuadro clínico genuino (p.ej., estrés postraumático) que corresponde a la sintomatología presentada. Adicionalmente, se plantearán una serie de hipótesis alternativas, entre las que se encuentran que la

¹ De ahora en adelante, información de síntomas siempre hará referencia a síntomas psicopatológicos.

sintomatología presentada no se concrete completamente en el cuadro clínico genuino de interés forense (los subsíndromes no son prueba forense), y que por ello sea prueba válida. Si se concreta el cuadro clínico de interés forense, es obligado someter a análisis la validez de los síntomas informados (diagnóstico diferencial de simulación). Las fuentes de información que se utiliza para evaluar dicha validez son: la (in)consistencia inter-medidas; en el proceso de evaluación, los resultados de las medidas de validez de la entrevista y los resultados de las medidas de validez de los test psicométricos administrados, que en las evaluaciones psicológicas reciben el nombre de test de validez de síntomas (*Symptom Validity Test*, SVT) (Arce, 2018; Giromini et al., 2022).

Cuando en un SVT, como el SIMS, se obtiene un resultado positivo (supera el punto de corte establecido o se registra un indicador de simulación), se ha obtenido evidencia de una anomalía relacionada con la forma de cumplimentar el test (p. ej., aquiescencia, estilo de respuesta aleatoria) o un *sesgo de respuesta negativa* (SRN). El SRN se refiere a la tendencia a responder un test generando una presentación menos saludable o más patológica de la que se podría esperar en base al nivel de ajuste de la persona evaluada (Giromini et al., 2022). Este concepto no incorpora variables motivacionales como la intencionalidad de la simulación, si no que se refiere exclusivamente al estilo de respuesta con que el evaluado ha respondido al test o entrevista. El SRN indicará que la información de síntomas de la persona evaluada es inválida, y la sintomatología podría no atribuirse exclusivamente a un cuadro clínico genuino. En este caso, se dispone de evidencia que confirma la hipótesis de invalidez de los síntomas informados. Autores, como Arce (2018), señalan que este resultado es un error del modelo y genera sospecha de invalidez debido a que cabe entre otras hipótesis, como que sea un caso grave.

Una vez determinada dicha invalidez, se deberá tratar de explicar el origen y el trasfondo motivacional que la subyace. En lo relativo al origen, puede dividirse, a grandes rasgos, en intencionales o no intencionales. Entre las explicaciones ‘no intencionales’ se encuentran aquellas que no están bajo el control de la persona evaluada, y ocurren de manera inconsciente. Por ejemplo, la existencia de un trastorno somatomorfo o disociativo (no como causa si no como condición en la que se producen las distorsiones), la influencia de la secuencia de los test aplicados, la falta de información sobre los síntomas, un patrón de respuesta a los instrumentos inatento (Merckelbach et al., 2019), determinados rasgos de la personalidad o un patrón de actuación defensivo (Rogers, 2018b). Por otro lado, las causas ‘intencionales’ se dividen acorde al tipo de incentivo y el trasfondo motivacional que motiva la presentación inválida. Si el incentivo es de origen psicológico, o interno, como los derivados de la asunción del rol de enfermo (atención al paciente, procedimientos médicos, diagnóstico y tratamiento de los problemas de salud), búsqueda de justicia, represalias o expiación, o deseo de reconocimiento, se tendrá que explorar la posibilidad de un trastorno facticio (Merten y Henry, 2022; Merten y Merckelbach, 2020). Por otro lado, si la presentación inválida está orientada a la obtención de algún incentivo externo, como una baja por enfermedad, una compensación económica, o eludir la responsabilidad legal, se considerará que la persona evaluada está simulando (Merten, 2018).

No obstante, como explican Merten y Merckelbach (2020), la mera presencia de un posible incentivo, ya sea interno o externo, no desempeña automáticamente un papel reforzador en la presentación

de los síntomas de los pacientes. Atribuir la responsabilidad de la presentación falsa a un único motivador puede ser una tarea altamente compleja, e incluso imposible cuando se identifican tanto factores externos como los internos (Merten y Rogers, 2017). En este caso, no será posible establecer la presencia de simulación o de trastorno facticio, y se utilizará el término ‘fingimiento’ (en inglés *feigning*). Fingir implica la exageración deliberada de sintomatología psicológica o física sin hacer ninguna atribución sobre la intencionalidad de la misma (Rogers y Bender, 2013). El término fingimiento ha sido sustituido en algunas publicaciones españolas por ‘exageración de síntomas’, como en González-Ordí y Santamaría (2012). No obstante, recomendamos el uso de fingimiento, dado que la propia definición ya incorpora el término exageración: “(la exageración de síntomas) podría definirse como aquel estilo de respuesta en el que el sujeto busca deliberadamente presentarse como deteriorado o alterado, exagerando o fabricando...”, pág. 192), y puede resultar confuso. De la misma forma, fingimiento es una traducción más precisa del término usado en la bibliografía internacional (*feigning*).

En la Figura 1 se ofrece un esquema de los conceptos expuestos, y en la Tabla 1 se presenta un glosario de referencia rápida con los principales conceptos expuestos.

Instrumentos Para la Evaluación de la Validez de la Información de Síntomas

Instrumentos Psicométricos: Test de Validez de Síntomas

El uso de instrumentos psicométricos, en especial los autoinformes, es uno de los métodos más utilizados para evaluar la validez de los síntomas informados (Aparcero et al., 2021). Los SVTs tienen como objetivo evaluar la validez de los síntomas manifestados o auto-informados (Giromini et al., 2022). Los SVTs son pruebas que se utilizan específicamente en procesos de evaluación de sintomatología psicológica/psiquiátrica. Como se ha indicado con anterioridad, un resultado positivo en un SVT advierte acerca de la posible falta de validez de las respuestas. Los SVTs pueden ser de dos tipos, independientes (*freestanding*) o integrados (*embedded*). Los SVTs independientes son test psicométricos auto-informados desarrollados para medir específicamente SRN, es decir, no miden otros constructos como personalidad patológica, severidad de sintomatología ansioso-depresiva, etc. En la actualidad se han adaptado a población española el SIMS (González-Ordí y Santamaría, 2009) y el Inventario de Problemas-29 (IOP-29; Puente-López, Pina, Rambaud-Quiñones et al., 2023). Por otro lado, se entiende por SVTs integrados las escalas diseñadas para medir SRN que están incorporadas en un test psicométrico que miden psicopatología. Por ejemplo, las escalas e índices del Inventario Multifásico de Personalidad de Minnesota-2 F, Fb, Fp, F-K, Fs, FBS, Ds y RBS (MMPI-2; Butcher et al., 2019); las escalas F-r, Fp-r, FBS-r Fs y RBS del Inventario Multifásico de Personalidad de Minnesota-2 Reestructurado (MMPI-2-RF; Ben-Porath y Tellegen, 2015), las escalas de validez (infrecuencia e impresión negativa) del Inventario de Evaluación de la Personalidad (PAI; Morey; 1991, 2007), las escalas de validez (sinceridad y simulación) del Inventario Clínico Multiaxial de Millon (MCMI; Millon et al., 2015), o la escala de distorsión negativa del Cuestionario de Impacto del Trauma (Crespo et al., 2020).

Figura 1
Diagrama de Flujo del Sistema Recomendado Para la Evaluación de la Validez de la Información de Síntomas

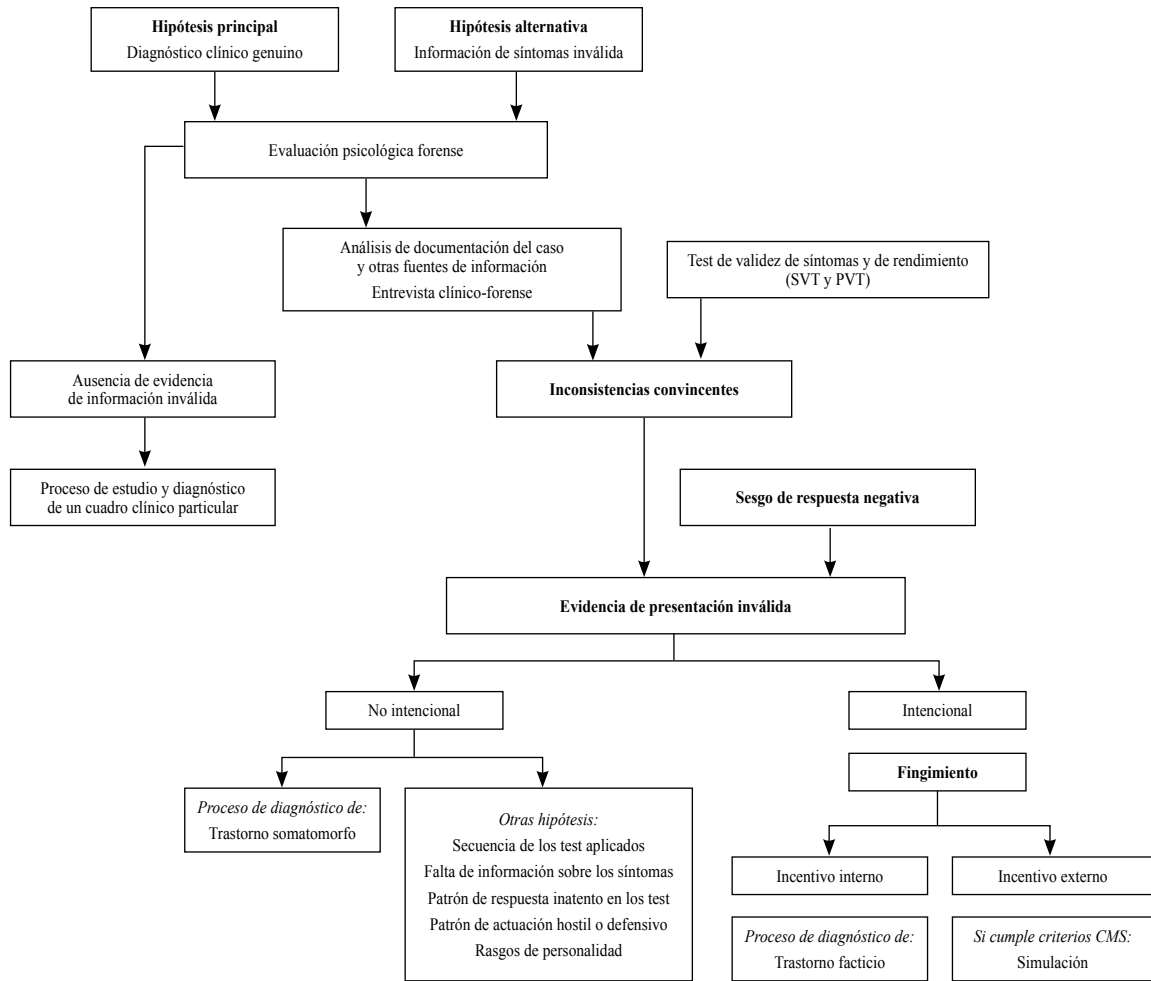


Tabla 1
Glosario de Conceptos Relacionados con la Validez de la Información de Síntomas

Concepto	Equivalencia ing.	Definición
Validez de la información de síntomas psicopatológicos	<i>Validity of self-reported symptoms; validity of symptom presentation</i>	La sintomatología referida por la persona valorada se puede atribuir completamente a un cuadro psicopatológico particular. La invalidez implica que se ha identificado la presencia de aspectos anómalos o inconsistentes en la información obtenida.
Sesgo de respuesta negativa	<i>Negative response bias</i>	Tendencia a responder un test generando una presentación (ya sea de síntomas o rendimiento) menos saludable o más patológica de la que se podría esperar en base al nivel de ajuste de la persona evaluada.
Fingimiento	<i>Feigning</i>	Fabricación o exageración flagrante intencional de sintomatología psicológica o física, sin hacer ninguna atribución sobre el trasfondo motivacional de la misma.
Simulación	<i>Malingering</i>	Fabricación o exageración flagrante intencional de sintomatología psicológica o física motivada por la obtención de un incentivo externo.
Test de validez de síntomas	<i>Symptom validity test (SVT)</i>	Instrumentos psicométricos diseñados para evaluar sesgo de respuesta negativa, producido por sobreinforme (<i>overreporting</i>) de síntomas psicopatológicos o por alteraciones en la forma de cumplimentar el test.
Test de validez del rendimiento	<i>Performance validity test (PVT)</i>	Instrumentos neuropsicológicos para evaluar sesgo de respuesta negativa, específicamente bajo rendimiento (<i>under-performance</i>).

Además de los SVTs, existe otro tipo de test diseñados en el ámbito de la neuropsicología para evaluar si el rendimiento cognitivo de un paciente es válido o inválido (es decir, patrones que no se ajustan al historial clínico del paciente o perfiles

neurocognitivos atípicos) (Greher y Wodushek, 2017). Este tipo de test se conocen como test de validez del rendimiento (en inglés *Performance Validity Test*, PVT), y en España el más usado es la adaptación española del Test de Simulación de Problemas de

Memoria (TOMM; Vilar-López et al., 2011). Dentro de los límites del presente trabajo no se profundizará en su uso y características particulares. Para extender información en este sentido pueden consultarse publicaciones como el artículo de Sweet et al. (2021). Únicamente mencionamos los PVTs para diferenciarlos de los SVTs, ya que, al igual que ocurre con el SIMS, PVTs como el TOMM son confundidos con test que evalúan simulación, y habitualmente son integrados dentro de una misma categoría, cuando evalúan constructos distintos.

Técnicas de Entrevista

Tanto para la evaluación psicológico-forense en general, como para el estudio de la validez de la información de síntomas, las técnicas de entrevista desempeñan un papel fundamental. Las entrevistas son el medio del diagnóstico de la salud mental y adicionalmente complementan la evaluación psicométrica del SRN. En la actualidad, las tres herramientas de entrevista más utilizadas en el campo de la validez de síntomas son la Entrevista Estructurada de Síntomas Informados (SIRS-2; Rogers et al., 2010), el Test Miller de Evaluación Forense de Síntomas (M-FAST; Miller, 2001) y la entrevista clínico-forense de (Arce y Fariña, 2001; Vilariño et al., 2013).

El M-FAST y la SIRS-2 son dos entrevistas estructuradas, compuestas por 25 y 172 ítems, respectivamente. Ambas se desarrollaron para evaluar una presentación inválida de sintomatología psicopatológica y estudiar la forma en que es probable que se produzca; por ejemplo, exageración de la gravedad de los síntomas frente a fabricación de la sintomatología (Rogers et al., 2010). Si bien cuentan con un amplio respaldo científico internacional, en España la investigación sobre su rendimiento ha sido limitada y, hoy en día, sólo se dispone de dos estudios preliminares de adaptación (Zaldívar-Basurto et al., 2008a; Zaldívar-Basurto et al., 2008b). En todo caso, estas entrevistas miden invalidez en las respuestas en general (miden la invalidez independientemente, *freestanding*), no en el contexto de la medida clínica o forense.

Por su parte, la entrevista clínico-forense (Arce y Fariña, 2001), es una medida enmascarada de la invalidez en el auto-informe de la salud mental y el estado emocional, que consiste en una entrevista narrativa (no estructurada) basada en dos técnicas: narración en discurso libre del estado de salud mental y estado emocional, y reinstauración de contextos (Vilariño et al., 2013). Dicha entrevista somete a los evaluados a “una tarea de conocimiento; no contamina la prueba con preguntas del entrevistador; posibilita el establecimiento de un nexo causal entre los hechos a enjuiciar y el daño psicológico, y contiene una herramienta estandarizada para el diagnóstico diferencial de simulación” (Arce et al., 2015, pág. 74). Esta entrevista mide, además, la sospecha de simulación en el contexto propio de la evaluación forense (y también clínica si fuera preciso, aunque no se planteó para este contexto). Esta entrevista se ha mostrado fiable y válida (validez predictiva, discriminante y convergente), tanto para la evaluación de la salud mental y estado emocional ($\alpha = .85$) como de la invalidez de los contenidos de lo manifestado ($\alpha = .744$) (Vilariño et al., 2013).

Criterios de Clasificación

Con la finalidad de disponer de un método organizado y sistemático mediante el cual interpretar los resultados de la

evaluación y llegar a una conclusión particular, Slick et al. (1999) propusieron los criterios para la simulación de disfunciones neurocognitivas (*Malingered Neurocognitive Dysfunction*, MND). Estos criterios, desarrollados para el ámbito de la neuropsicología forense, fueron la primera propuesta que integró la información científica acumulada hasta la fecha en un sistema de clasificación de la simulación comprensible y bien razonado (Boone, 2007). La propuesta de Slick et al. (1999) ha tenido un gran impacto en el área de la validez de síntomas y rendimiento, y su desarrollo ha influido tanto en la práctica profesional como en la investigación.

Los criterios MND siguen un sistema similar al de los criterios diagnósticos de manuales de clasificación, como el DSM. En su propuesta original, se consideraba la existencia de un incentivo externo significativo (Criterio A), evidencia de exageración o fabricación de disfunción cognitiva en pruebas neuropsicológicas (Criterio B), evidencia procedente de autoinformes en apoyo de un diagnóstico de simulación (Criterio C) y que los comportamientos que cumplan los criterios B o C no se expliquen totalmente por factores psiquiátricos, neurológicos o del desarrollo (Criterio D). Los criterios B y C contienen subcriterios que operativizan qué se entiende por evidencia. Así, por ejemplo, el criterio B indica que la evidencia podía ser un SRN definitivo (criterio B1) cuando la ejecución está por debajo del azar ($p < .05$) en una o más medidas cognitivas de elección forzosa (elegir entre dos estímulos diferentes); un SRN probable (criterio B2) cuando las respuestas en uno o más test o indicadores psicométricos empíricamente bien validados y diseñados para evaluar la exageración o fabricación de déficits cognitivos, es consistente con el engaño; discrepancia entre los datos de los test y los patrones de funcionamiento cerebral (criterio B3); discrepancia entre los datos de uno o más test neuropsicológicos que evalúan un mismo dominio y la conducta observada (criterio B4); discrepancia entre los test neuropsicológicos de un dominio y los informes colaterales fiables (criterio B5); y discrepancia entre los test neuropsicológicos y la historia clínica documentada (criterio B6). En función de los criterios y subcriterios que se identifican, se clasifica el protocolo en simulación definitiva, probable y posible (ver, para una descripción en español detallada de estos criterios, González-Ordi et al., 2012).

Recientemente, Sherman et al. (2020) publicaron una actualización de los criterios MND, los cuales pasaron a llamarse criterios multidimensionales para la simulación neurocognitiva, somática y psiquiátrica (criterios CMS). Como exponen los autores, esta actualización del modelo incorpora el conocimiento científico acumulado en los últimos 20 años, teniendo como principal finalidad simplificar las categorías diagnósticas; clarificar conceptos ambiguos como incentivo externo e incoherencia convincente; actualizar y redefinir el número de SVTs o PVTs que tienen que fallarse para alcanzar los criterios; abordar la cuestión de los falsos positivos en la administración de los SVTs y PVTs; y definir claramente los criterios de exclusión basados en la investigación sobre simulación en neuropsicología (Sherman et al., 2020). Además de la actualización de los criterios, la principal novedad del nuevo modelo es que ya no crea tres categorías de simulación (definitiva, probable o posible), sino que establece la presencia o ausencia de simulación de forma dicotómica, y, además, permite diferenciar entre cuatro especificadores de simulación: de disfunción neurocognitiva, de síntomas somáticos, de síntomas psiquiátricos/psicológicos, y de presentación mixta.

Esta nueva clasificación supone una mejoría significativa para el uso de los criterios en el ámbito forense. De hecho, a la hora de establecer una conclusión en un informe pericial, el perito puede encontrar una dificultad considerable para exponer qué diferencia hay entre posible y probable. Asimismo, juzgados y tribunales esperan una clasificación exacta, no probable, expresando el margen de error (Arce, 2017). El nuevo sistema permite alcanzar una definición unitaria de simulación a través de un sistema de clasificación dicotómico (presente/no presente) (Sherman et al., 2020). De la misma forma, abre la posibilidad de utilizar los criterios en otras áreas diferentes a la neuropsicología. En particular, la inclusión de la categoría de síntomas psiquiátricos/psicológicos permite su uso en psicología forense, dado que los criterios de decisión propuestos (presencia de incentivo externo, incoherencias en la presentación de los síntomas observados en la entrevista y/o en los test psicométricos y discrepancias marcadas entre datos externos y datos recogidos en la evaluación) son coherentes con los objetivos de evaluación seguidos en dicha área.

El Proceso de Evaluación de la Simulación y la Validez de la Información de Síntomas

Aspectos Previos a Considerar

A la hora de plantear el sistema de evaluación a utilizar, el perito tendrá que seguir los mismos principios generales por los que se rige la evaluación en psicología forense, es decir, se deberán emplear métodos y técnicas basadas en el conocimiento científico. En este sentido, a día de hoy, no se ha alcanzado un consenso en el contexto psicológico-forense español sobre un estándar de calidad de la evidencia científica a utilizar en los procesos forenses. Los criterios propuestos en la bibliografía, como los de Arce (2017), no son de obligado cumplimiento. No obstante, propuestas como las del citado autor son de utilidad para el correcto desarrollo de la práctica profesional, ya que ofrecen una guía de qué (y qué no), debería aplicarse en un contexto en el que se exige el máximo rigor profesional. Así, atendiendo a los criterios tanto nacionales (Arce, 2017) como internacionales (Rogers et al., 2023; Sweet et al., 2021) proponemos que las técnicas administradas en el contexto forense, destinadas a evaluar la validez de los síntomas, deban cumplir principalmente tres estándares: 1) que se disponga de bibliografía científica que respalde su aplicabilidad en el contexto de interés, 2) que la evidencia sea admisible en un proceso judicial y 3) que la batería psicométrica se configure atendiendo a criterios clínicos y estadísticos.

Respecto al *primer estándar*, todo/a profesional deberá asegurarse que la instrumentación o técnica ha sido adaptada a la población y tarea de interés. No se pueden ‘tomar prestados’ instrumentos de medida y asumir que sus propiedades psicométricas, especialmente la precisión de clasificación, va a mantenerse de un contexto a otro (Rogers et al., 2019). Por ejemplo, se indicaba en el apartado anterior que Zaldívar-Basurto et al. (2008a) y Zaldívar-Basurto et al. (2008b) habían iniciado el proceso de adaptación de la versión española la SIRS-2 y el M-FAST con dos breves estudios preliminares, y desde entonces dichos test no han recibido soporte empírico adicional en el país. Los estudios desarrollados por los citados autores comprenden el primer paso en el estudio

del funcionamiento de un SVT/PVT en un contexto particular, y son necesarios para hacer una valoración preliminar orientada a la toma de decisiones científica. No obstante, el diseño utilizado (diseño de simulación, ver González-Ordi et al., 2012) presenta una validez externa limitada, dado que se ha desarrollado con una muestra de estudiantes, y los hallazgos no pueden extrapolarse a un contexto aplicado en el que se toman decisiones importantes. Por este motivo, aunque a nivel internacional se trate de dos entrevistas que cuentan con amplia aceptación y apoyo empírico, no se puede determinar cuál es la precisión de clasificación del test, o validez diagnóstica, en el contexto forense español, y su administración en el mismo resulta desaconsejable.

Sobre el *segundo estándar*, deberá cumplir criterios de admisibilidad de la evidencia científica, como por ejemplo los criterios Daubert (Daubert vs. Merrell Dow Pharmaceuticals, 1993), los cuales determinan las características que debe presentar una prueba para ser admitida como científica en la Sala de Justicia (Arce, 2017; Neal et al., 2019). De entre los criterios Daubert, resulta especialmente relevante el relativo a la necesidad de conocer la potencial tasa de error de la técnica o instrumento administrado. La estimación de la tasa de error es fundamental en el contexto forense para determinar el peso y rendimiento de la evidencia (Dror y Scurich, 2020). Como exponen Puente-López, Pina et al. (2023) en psicología forense se ha propuesto habitualmente evaluar la tasa de error con lo que se conoce como índices de precisión de la clasificación, que son principalmente la sensibilidad y la especificidad (Greve y Bianchini, 2004; Langleben y Moriarty, 2013). Siguiendo las definiciones de Lange y Lippa (2017), la sensibilidad es la ratio de auténticos positivos de un test o medida, y responde a la pregunta de qué porcentaje del grupo de personas que tienen una condición concreta (por ejemplo, información de síntomas inválida), han sido correctamente clasificados por el instrumento. Por otro lado, la especificidad es la ratio de auténticos negativos de un test o medida, y responde a la pregunta de qué porcentaje del grupo control (sin la condición estudiada) han sido clasificados correctamente por el instrumento utilizado. En el caso de que la clasificación de la persona por el test sea errónea, el error puede ser un falso negativo (protocolo inválido clasificado como válido) o un falso positivo (protocolo válido clasificado como inválido).

La tasa de falsos negativos y falsos positivos se calcula restando a 1 el valor de la sensibilidad y especificidad, respectivamente, y se expresa en un porcentaje. Por ejemplo, recientemente, Puente-López, Pina, López-Nicolás et al. (2023) ejecutaron un meta-análisis sobre la capacidad de clasificación del IOP-29 con los diferentes puntos de corte propuestos por los autores. Los resultados mostraron que, para un punto de corte más liberal, la sensibilidad del IOP-29 es de .92 y la especificidad es de .82, lo que implica una tasa de falsos negativos del 8% (.08), y una tasa de falsos positivos del 18% (.18). Esto quiere decir que aproximadamente 2 de cada diez personas evaluadas que tengan una presentación válida serán catalogadas como que tienen una presentación inválida. Para evitar cometer un falso positivo, los autores proponen utilizar un punto de corte más conservador, que alcanza una especificidad de .97, lo que supone una tasa de falsos positivos del 3% (.03). En la doctrina Daubert no se establece qué es una tasa de error ‘aceptable’. Las directrices internacionales establecen que la tasa de falsos positivos para los test de evaluación de validez de síntomas

debe situarse entre el 5% y el 10% (Sweet et al., 2021). No obstante, en el contexto forense español también se considera que la tasa de falsos negativos debe ser lo más cercana posible a cero, dado que el caso contrario contravendría el principio de presunción de inocencia en causas penales o el de la carga de la prueba en causas civiles para la compensación y reparación de daños (Arce, 2017; Arce y Fariña, 2015).

En cuanto al *tercer estándar*, debe tenerse en consideración que los SVTs no pueden aplicarse en solitario, y deben estar integrados dentro de una batería psicométrica completa que incluya medidas de psicopatología adicionales a las de validez. En este sentido, una cuestión de gran importancia es cuántos SVTs incorporar en la batería psicométrica, ya que resulta desaconsejable combinar instrumentos psicométricos de forma arbitraria y sin un apoyo empírico que respalde la toma de decisiones. A pesar de que esta línea de investigación aún está en desarrollo, se ha alcanzado cierto consenso con que un único resultado positivo no es un indicador fiable de invalidez debido a la necesidad de baja tasa de falsos positivos y negativos ya expuesta, por lo que administrar un único test (por ejemplo, solo el SIMS) no es suficiente (Arce, 2018; Giromini et al., 2022; Sweet et al., 2021). En neuropsicología se ha estimado que, para identificar con un grado de precisión adecuado la invalidez del rendimiento, se necesitan al menos dos resultados positivos en hasta 9 instrumentos de medida (Sweet et al., 2021). Aunque esta pregunta se ha tratado con mucha menos frecuencia en el estudio validez de síntomas, la evidencia reciente parece apuntar a que el criterio de al menos dos resultados positivos también resulta válido en los SVTs (Arce, 2018; Chafetz, 2022; Giromini et al., 2022; Puente-López, Pina, Rambaud-Quñones et al., 2023). En las técnicas de evaluación Uforense se requieren, en línea con los hallazgos empíricos, al menos, 3 criterios positivos en protocolos consistentes inter-medidas para validar la psicopatología informada por la persona evaluada.

El Proceso de Evaluación

A continuación, se ofrece una propuesta, basada en la evidencia nacional e internacional, de cómo desarrollar el proceso de evaluación de la validez de síntomas. Las presentes pautas están supeditadas al estado de la bibliografía de la temática en España, de manera que, conforme avance el conocimiento científico, podrán ser ajustadas teniendo en cuenta en todo momento los estándares indicados en el apartado anterior.

El proceso de evaluación de la validez de la información de síntomas deberá seguir ineludiblemente una aproximación multi-medida (evaluación de la consistencia de la manifestado por el individuo, criterio judicial de persistencia) y multi-método, que combine entrevista con medidas psicométricas y someta a la persona evaluada a una tarea de conocimiento y de reconocimiento (Arce, 2018; Graham, 2011; Santamaría et al., 2020; Vilariño et al., 2009; 2013). Este enfoque está regido por el principio de consistencia, mediante el cual el/la profesional analiza la coherencia entre los diferentes hallazgos obtenidos en la evaluación forense (p. ej., anamnesis, informe libre de diagnóstico, datos educativos, observación del comportamiento, información de síntomas de la persona evaluada, resultados de los test psicométricos, etc.). El perito podrá identificar inconsistencias entre: a) los antecedentes informados por la persona evaluada y los

antecedentes documentados; b) síntomas informados por la persona evaluada que no concuerdan con los patrones clínicos conocidos; y c) síntomas informados por la persona evaluada que no concuerdan con el comportamiento observado (Bush et al., 2005; Santamaría et al., 2020).

Las principales herramientas que el/la profesional podrá utilizar en el proceso, además de la documentación previa relativa al caso, serán las descritas en el apartado segundo, es decir la entrevista especializada, la observación y los test de validez de síntomas. Esto permite integrar las herramientas en el proceso de valoración forense habitual, evitando la necesidad de extender en exceso el tiempo de evaluación. Teniendo en consideración que el proceso se orienta a estudiar las dos hipótesis expuestas en el primer apartado (diagnóstico genuino e invalidez de síntomas/presentación), la evaluación deberá desarrollar los siguientes puntos:

- a) Análisis de la información específica del caso. Como ya se exponía en el manual de González-Ordi et al. (2012) y en el capítulo de Santamaría et al. (2020), la evaluación de la validez de síntomas y la simulación no puede hacerse a partir de una única fuente de información, como un test psicométrico. Se trata de un proceso centrado en la convergencia de múltiples fuentes de información que provienen de diversas áreas de conocimiento y profesionales. Así, además de la información obtenida en el proceso de evaluación forense, se deberá recopilar todos los datos disponibles sobre el caso, como el historial clínico, evidencia médica previa, evaluaciones psicológicas previas, declaraciones anteriores, información sociodemográfica general, etc. Esta información permite al profesional asentar una línea base de trabajo con la que poder contrastar la información obtenida durante la evaluación.
- b) La obtención de evidencia de SRN. Habitualmente es el primer paso a realizar en el proceso de evaluación de la invalidez de la información de síntomas. Dicha evidencia se obtendrá mediante el estudio de los resultados de los cuestionarios psicométricos (SVTs) administrados. Para configurar la batería psicométrica de evaluación se tendrá que incorporar al menos dos SVTs (p. ej., SIMS y MMPI-2), y la presencia de SRN se confirmará con, al menos, dos resultados positivos en los test administrados (Giromini et al., 2022; Sweet et al., 2021). En la técnica Uforense se confirmará con 3 o más resultados positivos en las medidas de validez, en protocolos consistentes inter-medidas (Arce et al., 2015). A la hora de expresar los resultados en un informe pericial, deberá considerarse que la presencia de SRN es un indicador de presentación inválida, pero no permite realizar ninguna inferencia sobre la intencionalidad de la misma, por lo que no basta para establecer la presencia de simulación. De la misma forma, no obtener evidencia de SRN no es indicativo de presentación válida (Chafetz, 2022), es decir, no es evidencia de que la sintomatología de la persona evaluada es genuina, y dicha conclusión únicamente podrá alcanzarse mediante el estudio completo de toda la información del caso.
- c) El estudio de la coherencia, intencionalidad e hipótesis explicativas. Éste se desarrollará principalmente a través de una técnica de entrevista adaptada al contexto forense, como la entrevista clínico-forense. La entrevista perseguirá tres objetivos: obtener información sobre el paciente en las dimensiones biopsicosocial, estudiar la información

contradictoria directamente, ya sea para disolver las contradicciones aparentes o para confirmarlas y analizar la intencionalidad del SRN (si procede). En este paso, es recomendable que el/la profesional haya administrado y corregido los test psicométricos de la batería configurada para el caso y haya elaborado un listado de inconsistencias a estudiar, de manera que pueda desarrollar la entrevista teniendo en consideración esta información. La entrevista se desarrollará de la forma habitual en psicología forense (Arce y Fariña, 2015; Arce et al., 2015), incorporando:

- 1) Registro de incoherencias convincentes. Se tratan de observaciones que pueden realizarse en el proceso de entrevista que suponen una prueba clara y contundente de presentación inválida. Por ejemplo, el patrón de síntomas que presenta el paciente resulta incoherente o inconsistente con el conocimiento científico disponible. Específicamente, se deberá prestar atención a las siguientes incoherencias: presentación de síntomas sutiles, improbables, raros u obvios, combinación atípica de síntomas, severidad de síntomas anómala, presentación de síntomas inconsistente y agrupación indiscriminada de síntomas (Rogers, 2018b).
- 2) Evaluación de las incoherencias. Las inconsistencias e incoherencias observadas en el proceso de evaluación, tanto en la entrevista como en la valoración psicométrica, deberán ser estudiadas en profundidad ya sea para disolverlas o para confirmarlas. Resulta recomendable utilizar en este proceso el enfoque que siguen las entrevistas investigativas, es decir, no acusatorio y exploratorio (Meissner et al., 2023; Powel et al., 2005), dado que reduce la resistencia a la evaluación.
- 3) Valoración de la intencionalidad. En el caso de que se obtenga evidencia de presentación inválida, ya sea mediante el registro de incoherencias y/o presencia de SRN, deberá evaluarse la intencionalidad/no intencionalidad de la misma. En primer lugar, se estudiarán hipótesis explicativas relacionadas con la no intencionalidad, como que el resultado haya sido producido por las variables relativas a la evaluación (p. ej., cansancio ante una batería excesivamente extensa), por falta de información, o información errónea, sobre los síntomas, por ciertos rasgos de personalidad, o por un patrón de respuesta inatento (Merckelbach et al., 2019). De la misma forma, se deberá hacer un diagnóstico diferencial de un posible trastorno somatomorfo o disociativo (Merten y Merckelbach, 2020).
- 4) Una vez descartadas las hipótesis explicativas de la no intencionalidad, la valoración de la intencionalidad se realizará mediante un estudio pormenorizado de la presencia e impacto de incentivos externos o internos sustanciales. En el primer caso, se tendrá que evaluar en profundidad la situación socioeconómica, los factores sociales, financieros y del contexto familiar, la historia laboral, los conflictos actuales y los asuntos legales. En el segundo se tendrá que estudiar variables relacionadas con la asunción del rol de enfermo, y realizar un diagnóstico diferencial de trastorno facticio. En esta fase resulta importante que el/la profesional

considere que la evaluación de la intencionalidad en formato dicotómico no siempre será posible, al igual que en muchos casos será difícil aclarar el verdadero trasfondo motivacional que subyace la presentación inválida (Merten y Merckelbach, 2020).

- d) Determinación de la simulación. Para ello se deberá llegar a una conclusión respaldada por los hallazgos obtenidos durante el proceso de valoración propuesto. Hoy en día no se dispone en España de un criterio de decisión específico para combinar los citados hallazgos, lo que conlleva el riesgo de llegar a una conclusión de validez/invalides de síntomas arbitraria, poco fundamentada o errónea. Por este motivo, recomendamos que la presencia/ausencia de riesgo de simulación se establezca con los criterios multidimensionales para la simulación neuropsicológica de Sherman et al. (2020) expuestos con anterioridad, dado que ofrecen un método estructurado, sistemático y replicable para la toma de decisiones forense.

Para establecer una conclusión con el sistema de Sherman et al. (2020), el forense valorará si se cumplen los criterios diagnósticos de simulación a partir de los hallazgos obtenidos durante el sistema de evaluación propuesto. Específicamente, se evaluará la presencia de un incentivo externo (criterio A), de invalidez de los síntomas informados (ya sea identificada en la entrevista o con la presencia de SRN, criterio B), de discrepancias marcadas (criterio C) y se realizará un diagnóstico diferencial con condiciones psiquiátricas que puedan explicar la presentación (trastornos somatomorfos, facticios, etc., criterio D). Dentro del criterio B se podrán encontrar cuatro opciones (presentación neurocognitiva, psiquiátrica, somática y mixta), y deberá elegirse la que corresponda al área de trabajo en particular. Específicamente, para la evaluación de la validez de síntomas de una condición psicopatológica (p. ej., trastorno de estrés postraumático, trastorno adaptativo, trastornos del espectro de la esquizofrenia y psicóticos) se valorarán los subcriterios B3 de presentación psiquiátrica.

En el Anexo 1 se presentan los criterios CMS de Sherman et al. (2020) traducidos al castellano. En la traducción se han realizado dos modificaciones para adaptarlos correctamente al contexto forense español. La primera ha sido en el número de SVTs necesarios para determinar SRN (criterio B3b). La publicación original indica que la evidencia psicométrica puede ser 'una o más puntuaciones de pruebas de validez de síntomas'. Esto implica que podría considerarse que se cumple el citado criterio con un positivo en un solo SVT como, por ejemplo, el SIMS, siempre y cuando tenga una tasa aceptable de falsos positivos y puntos de corte estimados empíricamente en la población de interés. Este planteamiento puede resultar problemático para el contexto forense español, dado que ninguno de los SVTs disponibles en la bibliografía cumple el criterio de tasa de falsos positivos inferior al 10%. El SIMS o el MMPI-2 tienen una muy elevada sensibilidad, en especial el primero, dado que es un test de screening, pero una tasa elevada de falsos positivos (entre el 20% y el 30%), por lo que su uso en solitario resulta desaconsejable.

Como se exponía con anterioridad, resulta necesario alcanzar una tasa de falsos negativos cercana a 0 y de falsos positivos no superior al 10%, lo que resulta imposible con la administración de un solo test. Por este motivo, se recomienda el uso de al menos dos SVTs, y obtener al menos dos resultados positivos, para considerar que se cumple el citado criterio B3b. Esta propuesta resulta consistente

con los hallazgos recientes de estudios del área en España, como la adaptación al castellano del IOP-29, donde se observó que su uso junto al SIMS disminuía la tasa de falsos positivos y generaba una mayor validez incremental (Puentes-López et al., 2023). No obstante, establecer el número óptimo de SVTs que deben incorporarse en la batería psicométrica es una cuestión que aún debe ser investigada con mayor profundidad, por lo que nuestra propuesta deberá ser adaptada en función del conocimiento científico futuro.

La segunda modificación ha sido que, para considerar la existencia de simulación, deben cumplirse los criterios B3a (inconsistencias identificadas en el proceso de entrevista forense) y B3b (evidencia psicométrica de invalidez). La publicación original indica que debe cumplirse uno de los dos criterios, no obstante, en la bibliografía española se hace hincapié en la necesidad de utilizar tanto las técnicas de entrevista como psicométricas para establecer una conclusión diagnóstica en general. Específicamente en el área de validez de síntomas, estudios como el de Arce et al. (2015) evidencian que el uso conjunto de dos instrumentos resulta beneficioso dado incrementa la validez y ayuda a controlar el número de falsos positivos y negativos, crítico en la evaluación forense.

Cómo Reflejar los Hallazgos en el Informe Pericial

A la hora de plasmar los resultados de la evaluación de la validez de la información de síntomas, el perito deberá ser especialmente cuidadoso con la forma de expresarse, y tendrá que ser claro en lo relativo a las posibles limitaciones identificadas durante el proceso. Para establecer la presencia de riesgo de simulación el perito, una vez identificada invalidez de la información de síntomas, debe garantizar que se han cumplido los criterios CMS de Sherman et al. (2020) para simulación psiquiátrica y los establecidos por el sistema Uforense. El informe pericial deberá incluir una justificación pormenorizada de por qué se considera que cada criterio se cumple o no se cumple. En este caso, deberá especificarse claramente que la sintomatología presentada no puede atribuirse en su totalidad a un cuadro clínico genuino, ya que se ha identificado un intento intencional de presentar o exagerar síntomas falsos aparentemente motivado por la obtención de un incentivo externo determinado.

El profesional debe recordar en todo momento que, para establecer la presencia de simulación, deben cumplirse necesariamente los criterios CMS, o los del sistema Uforense. En el caso de que no sea posible establecer la intencionalidad o identificar un incentivo externo, no se podrá hablar de simulación. En este sentido, se ha venido utilizando en la práctica pericial profesional española la frase ‘indicios compatibles con riesgo de simulación’, para aquellos casos en los que no se ha conseguido identificar el trasfondo motivacional subyacente. Esta frase resulta parcialmente correcta pero imprecisa, dado que dichos indicios (generalmente SRN) también son compatibles con otras hipótesis, como las detalladas en apartados anteriores. De la misma forma, también resulta desaconsejable, debido a que la mención del término ‘simulación’ tiene connotaciones, y consecuencias, negativas.

En el caso de que el perito haya identificado evidencia que indica una información de síntomas inválida (p. ej., SRN), pero no pueda hacer ninguna atribución a la intencionalidad o a la orientación a un incentivo en particular, deberán estudiarse las hipótesis alternativas expuestas en apartados anteriores (p.ej., trastornos facticios). Si

no es posible llegar a una conclusión o diagnóstico en particular, deberá reflejarse en las conclusiones que se ha identificado que la información de síntomas de la persona evaluada es inválida, lo que impide atribuir los síntomas informados a un cuadro clínico en particular. Deberán exponerse también las limitaciones que hayan impedido alcanzar una conclusión de simulación (p. ej., falta de información que permita establecer con certeza el sistema motivacional subyacente).

Consideraciones Finales

La metodología expuesta en la presente publicación, así como los tres estándares científicos aconsejados, resumen las principales recomendaciones de los expertos internacionales en la temática (Arce, 2017, 2018; Santamaría et al., 2020; Sweet et al., 2021), y tienen una amplia aceptación en países como Estados Unidos o Alemania. No obstante, debe tenerse en consideración que su correcta aplicación depende de los medios disponibles en el contexto en particular. Específicamente en España, debemos advertir que la evidencia y los recursos disponibles son limitados, en comparación con la bibliografía internacional, y cumplir los citados estándares puede ser una tarea compleja, o incluso inalcanzable. Esta dificultad es más notable en la evaluación de la validez de síntomas, donde se dispone de un menor número de instrumentos desarrollados y adaptados, en comparación con la evaluación neuropsicológica de la validez del rendimiento. Por este motivo, recomendamos encarecidamente que los/as profesionales sigan un enfoque conservador centrado en minimizar el riesgo de cometer falsos positivos y negativos.

Si bien no hay duda de que la disciplina ha avanzado en la última década, aún queda un largo camino por recorrer para integrar en la bibliografía española los principales avances internacionales. Entre ellos, resulta de vital importancia fomentar tanto la adaptación de nuevos test de validez de síntomas y rendimiento (como, por ejemplo, el *Self-Report Symptom Inventory*, SRSI, Merten et al., 2022) como el desarrollo de nuevas investigaciones con los ya adaptados, en nuevos contextos y poblaciones de interés e incluso con diseños de investigación con mayor validez externa.

Referencias

- American Psychiatric Association. (2023). *DSM-5-TR manual diagnóstico y estadístico de los trastornos mentales: Texto revisado*. Editorial Panamericana.
- Aparcero, M., Picard, E. H., Nijdam-Jones, A., y Rosenfeld, B. (2021). The impact of coaching on feigned psychiatric and medical symptoms: A meta-analysis using the MMPI-2. *Psychological Assessment*, 33(8), 729-745. <https://doi.org/10.1037/pas0001016>
- Arce, R. (2017). Análisis de contenido de las declaraciones de testigos: Evaluación de la validez científica y judicial de la hipótesis y la prueba. *Acción Psicológica*, 14(2), 171-190. <https://dx.doi.org/10.5944/ap.14.1.21347>
- Arce, R. (2018). Evaluación del daño psicológico: Psicometría, entrevista y técnica forense. En E. Carbonell, D. Pineda, y M. Novo (Eds.), *Psicología jurídica: Ciencia y profesión. Colección Psicología y Ley*, 15 (pp. 235-244). Sociedad Española de Psicología Jurídica y Forense.
- 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. (2012). La entrevista psicológica forense a niños, adultos y discapacitados. S. Delgado, F. Bandrés, y A. Tejerina (Eds.), *Tratado de medicina legal y ciencias forenses: Vol. V. Pediatría legal y forense. Violencia. Víctimas* (pp. 795-817). Bosch.
- Arce, R., y Fariña, F. (2015). Evaluación psicológico-forense de la credibilidad y daño psíquico mediante el Sistema de Evaluación Global. En P. Rivas y G.L. Barrios (Eds.), *Violencia de género: Perspectiva multidisciplinar y práctica forense* (pp. 411-441). Thomson Aranzadi.
- Arce, R., y Fariña, F. (2020). *Técnicas y protocolos de periciales del Sistema de Evaluación Global*. Andavira.
- Arce, R., Fariña, F., Carballeda, 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(2), 241-247.
- Arce, R., Fariña, F., Carballeda, 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(2), 278-283.
- 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(3), 385-408.
- Arce, R., Fariña, F., y Vilariño, M. (2015). Daño psicológico en casos de víctimas de violencia de género: Estudio comparativo de las evaluaciones forenses. *Revista Iberoamericana de Psicología y Salud*, 6(2), 72-80. <https://doi.org/10.1016/j.riips.2015.04.002>
- Ben-Porath, Y. S., y Tellegen, A. (2015). *Inventario Multifásico de Personalidad de Minnesota-2 Reestructurado*. TEA Ediciones.
- Boone, K. B. (2007). A reconsideration of the Slick et al. (1999) criteria for malingered neurocognitive dysfunction. En K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 29-49). The Guilford Press.
- Bush, S. S., Ruff, R. M., Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., Reynolds, C. R., y Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity NAN policy & planning committee. *Archives of Clinical Neuropsychology*, 20(4), 419-426. <https://doi.org/10.1016/j.acn.2005.02.002>
- Butcher, J. N., J. R. Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., y Kaemmer, B. (2019). *MMPI-2. Inventario Multifásico de Personalidad de Minnesota-2*. TEA Ediciones.
- Calcedo, A. (2000). Evaluación forense. En J. Bobes, M. Bousoño, A. Calcedo y M. P. González (Eds.), *Trastorno de estrés postraumático* (pp. 287-299). Masson.
- Chafetz M. D. (2022). Deception is different: Negative validity test findings do not provide “evidence” for “good effort”. *The Clinical Neuropsychologist*, 36(6), 1244-1264. <https://doi.org/10.1080/13854046.2020.1840633>
- Chafetz, M. D., Bauer, R. M., y Haley, P. S. (2020). The other face of illness-deception: Diagnostic criteria for factitious disorder with proposed standards for clinical practice and research. *Clinical Neuropsychologist*, 34(3), 454-476. <https://doi.org/10.1080/13854046.2019.1663265>
- Crespo, M., González-Ordí, H., Gómez-Gutiérrez, M. M., y Santamaría, P. (2020). *Cuestionario de Impacto del Trauma (CIT)*. TEA Ediciones.
- Dandachi-FitzGerald, B., Merckelbach, H., y Merten, T. (2022). Cry for help as a root cause of poor symptom validity: A critical note. *Applied Neuropsychology: Adult*. Advance online publication. <https://doi.org/10.1080/23279095.2022.2040025>
- Daubert vs. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786 (1993).
- Dror, I. E., y Scurich, N. (2020). (Mis)use of scientific measurements in forensic science. *Forensic Science International*, 2, 333-338. <https://doi.org/10.1016/j.fsisyn.2020.08.006>
- Echeburúa, E., Muñoz, J. M., y Loinaz, I. (2011). La evaluación psicológica forense frente a la evaluación clínica: Propuestas y retos de futuro. *International Journal of Clinical and Health Psychology*, 11(1), 141-159.
- Fariña, F., Arce, R., Vilariño, M., & Novo, M. (2014). Assessment of the standard forensic procedure for the evaluation of psychological injury in intimate-partner violence. *The Spanish Journal of psychology*, 17, E32. <https://doi.org/10.1017/sjp.2014.30>
- Garner, H. H. (1965). Malingering. *Illinois Medical Journal*, 128, 318-319.
- Giromini, L., Young, G., y Sellbom, M. (2022). Assessing negative response bias using self-report measures: New articles, new issues. *Psychological Injury and Law*, 15(1), 1-21. <https://doi.org/10.1007/s12207-022-09444-2>
- González-Ordí, H., y Santamaría, P. (2009). *Adaptación española del Inventario Estructurado de Simulación de Síntomas – SIMS*. TEA Ediciones.
- González-Ordí, H., Santamaría, P., y Capilla, P. (Eds.) (2012). *Estrategias de detección de la simulación. Un manual clínico multidisciplinar*. TEA Ediciones.
- Graham, J. R. (2011). *MMPI-2: Assessing personality and psychopathology* (5a ed.). Oxford University Press.
- Greene, R. L. (2011). *The MMPI-2/MMPI-2-RF: An interpretive manual*. Allyn & Bacon.
- Greher, M. R., y Wodushek, T. R. (2017). Performance validity testing in neuropsychology: Scientific basis and clinical application - A brief review. *Journal of Psychiatric Practice*, 23(2), 134-140. <https://doi.org/10.1097/PRA.0000000000000218>
- Greve, K. W., y Bianchini, K. J. (2004). Setting empirical cut-offs on psychometric indicators of negative response bias: A methodological commentary with recommendations. *Archives of Clinical Neuropsychology*, 19(4), 533-541. <https://doi.org/10.1016/j.acn.2003.08.002>
- Lange, R. T., y Lippa, S. M. (2017). Sensitivity and specificity should never be interpreted in isolation without consideration of other clinical utility metrics. *Clinical Neuropsychologist*, 31(6-7), 1015-1028. <https://doi.org/10.1080/13854046.2017.1335438>
- Langleben, D. D., y Moriarty, J. C. (2013). Using brain imaging for lie detection: Where science, law and research policy collide. *Psychology, Public Policy, and Law*, 19(2), 222-234. <https://doi.org/10.1037/a0028841>
- Meissner, C. A., Kleinman, S. M., Mindthoff, A., Phillips, E. P., y Rothweiler, J. N. (2023). Investigative interviewing: A review of the literature and a model of science-based practice. En D. DeMatteo y K. C. Scherr (Eds.), *The Oxford handbook of psychology and law* (pp. 582-603). Oxford University Press.
- Merckelbach, H., Dandachi-FitzGerald, B., van Helvoort, D., Jellicic, M., y Otgaar, H. (2019). When patients overreport symptoms: More than just malingering. *Current Directions in Psychological Science*, 28(3), 321-326. <https://doi.org/10.1177/0963721419837681>
- Merten, T. (2018). False symptom claims and symptom validity assessment. En H. Otgaar y M. Howe (Eds.), *Finding the truth in the courtroom? Problems with deception, lies, and memories*. Oxford University Press.
- Merten, T., y Henry, M. (2022). Symptom and performance validity assessment in older adults and patients with dementia and claimed dementia. En S. S. Bush y B. P. Yochim (Eds.), *A handbook of geriatric neuropsychology: Practice Essentials* (2a ed., pp. 286-303). Routledge.

- Merten, T., Dandachi-FitzGerald, B., Hall, V., Bodner, T., Giromini, L., Lehrner, J., González-Ordí, H., Santamaría, P., Schmand, B., y Di Stefano, G. (2022). Symptom and performance validity assessment in European countries: An update. *Psychological Injury and Law*, 15(2), 116-127. <https://doi.org/10.1007/s12207-021-09436-8>
- Merten, T., Dandachi-FitzGerald, B., Hall, V., Schmand, B. A., Santamaría, P., y González-Ordí, H. (2013). Symptom validity assessment in European countries: Development and state of the art. *Clinica y Salud*, 24(3), 129-138. [https://doi.org/10.1016/S1130-5274\(13\)70014-8](https://doi.org/10.1016/S1130-5274(13)70014-8)
- Merten, T., y Merckelbach, H. (2020). Factitious disorders and malingering. En J. R. Geddes, N. C. Andreasen, y G. M. Goodwin (Eds.), *New Oxford textbook of psychiatry* (3a ed., pp. 1342-1349). Oxford University Press.
- Merten, T., y Rogers, R. (2017). An international perspective on feigned mental disabilities: Conceptual issues and continuing controversies. *Behavioral Sciences & the Law*, 35(2), 97-112. <https://doi.org/10.1002/bsl.2274>
- Miller, H. A. (2001). *Miller Forensic Assessment of Symptoms Test: Professional manual*. Psychological Assessment Resources.
- Millon, T., Grossman, S., y Millon, C. (2015). *Millon Clinical Multiaxial Inventory-IV*. Pearson Assessments.
- Morey, L. C. (1991). *The Personality Assessment Inventory. Professional manual*. Psychological Assessment Resources, Inc.
- Morey, L. C. (2007). *The Personality Assessment Inventory professional manual* (2a ed.). Psychological Assessment Resources.
- Neal, T. M. S., Slobogin, C., Saks, M. J., Faigman, D. L., y Geisinger, K. F. (2019). Psychological assessments in legal contexts: Are courts keeping “Junk Science” out of the courtroom? *Psychological Science in the Public Interest*, 20(3), 135-164. <https://doi.org/10.1177/1529100619888860>
- Powell, M. B., Fisher, R. P., y Wright, R. (2005). Investigative interviewing. En N. Brewer y K. D. Williams (Eds.), *Psychology and law: An empirical perspective* (pp. 11-42). The Guilford Press.
- Puente-López, E., Pina, D., y Arce, R. (2023). Misunderstandings and misconceptions in the application of nonverbal communication in the Spanish legal-forensic context. *Acción Psicológica*, 20(2), 43-70. <https://doi.org/10.5944/ap.20.2.39334>
- Puente-López, E., Pina, D., López-Nicolás, R., Iguacel, I., y Arce, R. (2023). The Inventory of Problems - 29 (IOP-29): A systematic review and bivariate diagnostic test accuracy meta-analysis. *Psychological Assessment*, 35(4), 339-352. <https://doi.org/10.1037/pas0001209>
- Puente-López, E., Pina, D., Rambaud-Quiñones, P., Ruiz-Hernández, J. A., Nieto-Cañaveras, M. D., Shura, R. D., Alcazar-Crevillén, A., y Martínez-Jarreta, B. (2023). Classification accuracy and resistance to coaching of the Spanish version of the Inventory of Problems -29 and the Inventory of Problems - Memory: A simulation study with mTBI patients. *Clinical Neuropsychologist*. Advance online publication. <https://doi.org/10.1080/13854046.2023.2249171>
- Resnick, P. J. (1997). Malingering of posttraumatic disorders. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2a ed., pp. 84-103). Guilford Press.
- Rogers R. (2018a). An introduction to response styles. En R. Rogers y S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4a ed., pp. 3-17). The Guilford Press.
- Rogers, R. (2018b). Current status of clinical methods. En R. Rogers y S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4a ed., pp. 571-519). The Guilford Press.
- Rogers, R. (2021). Professional pitfalls in malingering determinations. *The Journal of the American Academy of Psychiatry and the Law*, 49(3), 296-299. <https://doi.org/10.29158/JAAPL.210073-21>
- Rogers, R., y Bender, S. D. (2013). Evaluation of malingering and related response styles. En R. K. Otto y I. B. Weiner (Eds.), *Handbook of psychology: Forensic psychology* (pp. 517-540). John Wiley & Sons, Inc.
- Rogers, R., Donnelly, J. W. II, y Correa, A. A. (2019). Translated measures in forensic evaluations with specific applications to feigned mental disorders. *Psychological Injury and Law*, 12(3-4), 191-203. <https://doi.org/10.1007/s12207-019-09362-w>
- Rogers, R., Tazi, K. Y., y Drogin, E. Y. (2023). Forensic assessment instruments: Their reliability and applicability to criminal forensic issues. *Behavioral Sciences & the Law*, 41(5), 415-431. <https://doi.org/10.1002/bsl.2613>
- Rogers, R., Sewell, K. W., y Gillard, N. D. (2010). *Structured Interview of Reported Symptoms 2nd Edition (SIRS-2): Professional manual*. Psychological Assessment Resources.
- Rogers, R., y Vitacco, M. J. (2002). Forensic assessment of malingering and related response styles. En B. Van Dorsten (Ed.), *Forensic psychology: From classroom to courtroom* (pp. 83-104). Kluwer Academic
- Santamaría, P., González-Ordí, H., y Capilla, P. (2020). La evaluación de la simulación de sintomatología. En M. A. Carrasco y F. Rubio-Garay (Coord.). *Psicología jurídica y forense, volumen I: Aspectos psicológicos y legales básicos* (pp. 355-409). Sanz y Torres.
- Sherman, E. M. S., Slick, D. J., y Iverson, G. L. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology*, 35(6), 735-764. <https://doi.org/10.1093/arclin/aaa019>
- Slick, D. J., Sherman, E. M. S., y Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, 13(4), 545-561. [https://doi.org/10.1076/1385-4046\(199911\)13:04;1-Y;FT545](https://doi.org/10.1076/1385-4046(199911)13:04;1-Y;FT545)
- Sweet, J. J., Heilbronner, R. L., Morgan, J. E., Larrabee, G. J., Rohling, M. L., Boone, K. B., Kirkwood, M. W., Schroeder, R. W., Suhr, J. A., y Conference Participants (2021). American Academy of Clinical Neuropsychology (AACN) 2021 consensus statement on validity assessment: Update of the 2009 AACN consensus conference statement on neuropsychological assessment of effort, response bias, and malingering. *Clinical Neuropsychologist*, 35(6), 1053-1106. <https://doi.org/10.1080/13854046.2021.1896036>
- Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM)*. Multi-Health Systems, Inc.
- Vilar-López, R., Pérez-García, M., y Puente, A. (2011). *Adaptación española del Test de Simulación de Problemas de Memoria TOMM*. TEA Ediciones.
- Vilarriño, M., Arce, R., y Fariña, F. (2013). Forensic-clinical interview: Reliability and validity for the evaluation of psychological injury. *European Journal of Psychology Applied to Legal Context*, 5(1), 1-21.
- Vilarriño, M., Fariña, F., y Arce, R. (2009). Discriminating real victims from feigners of psychological injury in gender violence: Validating a protocol for forensic setting. *European Journal of Psychology Applied to Legal Context*, 1(2), 221-243.
- Ward, M. K., y Meade, A. W. (2023). Dealing with careless responding in survey data: Prevention, identification, and recommended best practices. *Annual Review of Psychology*, 74, 577-596. <https://doi.org/10.1146/annurev-psych-040422-045007>
- Widows, M. R., y Smith, G. P. (2005). *SIMS: Structured Inventory of Malingered Symptomatology. Professional manual*. Psychological Assessment Resources.

Wooley, C. N., y Rogers, R. (2015). The effectiveness of the personality assessment inventory with feigned PTSD: An initial investigation of Resnick's model of malingering. *Assessment*, 22(4), 449-458. <https://doi.org/10.1177/1073191114552076>

Zaldívar Basurto, F., García Montes, J. M., López Ríos, F., Molina, A., y Santiago, E. (2008a). Adaptación y validación al castellano de la evaluación forense de Miller del Test de Síntomas (M-Fast). En F. J. Rodríguez, C. Bringas, F. Fariña, R. Arce y A. Bernardo (Eds.).

Psicología jurídica. Familia y victimología (pp. 461-468). Ediciones de la Universidad de Oviedo.

Zaldívar Basurto, F., García Montes, J. M., Santiago, E. Molina, A., y López Ríos, F. (2008b). Adaptación y validación al castellano de la Entrevista Estructurada de Síntomas Informados (SIRS). En F. J. Rodríguez, C. Bringas, F. Fariña, R. Arce y A. Bernardo (Eds.). *Psicología jurídica. Familia y victimología* (pp. 469-477). Ediciones de la Universidad de Oviedo.

Financiación: El presente trabajo no recibió financiación específica de agencias del sector público, comercial o de organismos no gubernamentales.

Declaración de Consentimiento Informado: No aplicable.

Declaración de Disponibilidad de los Datos: No aplicable.

Conflicto de Intereses: Los autores declaran que no hay conflicto de intereses.

Anexo I

Criterios Multidimensionales Para la Simulación Neurocognitiva, Somática y Psiquiátrica (Sherman et al., 2020)

La simulación es la invención o exageración voluntaria de síntomas neurocognitivos, somáticos o psiquiátricos con el fin de obtener beneficios materiales y servicios o evitar deberes formales, responsabilidades o resultados indeseables. Está indicado por pruebas claras y convincentes basadas en los cuatro criterios enumerados a continuación (Criterios A-D).

A. Presencia de un incentivo externo

Un incentivo externo claramente identificable y sustancial para fingir o exagerar déficits o síntomas está presente en el momento del examen. Los incentivos externos para fingir incluyen el acceso a un resultado deseable, como un acuerdo económico, un pago por incapacidad, la sustitución del salario, la asistencia social; el acceso a servicios o adaptaciones en entornos comunitarios, académicos o laborales, o el acceso a medicación recetada.

Los incentivos externos también pueden incluir la evitación de un resultado indeseable, como los relacionados con procesos penales (por ejemplo, evitar ser considerado competente para ser juzgado o evitar una sentencia penal), el servicio militar (por ejemplo, evitar el despliegue) o entornos laborales o escolares (por ejemplo, evitar la libertad condicional, la suspensión, la expulsión o el despido). Evitar un resultado indeseable en el contexto del engaño también puede ser adaptativo (por ejemplo, fingir una enfermedad para ser liberado en una situación de rehenes o para evitar ser devuelto a una situación de maltrato). Los incentivos externos para el fingimiento también pueden incluir evitar tener que cumplir con deberes y responsabilidades más básicos, como evitar el trabajo, la escuela o los exámenes, o las responsabilidades domésticas.

Los tipos de evaluaciones asociadas a los incentivos externos del fingimiento incluyen las relacionadas con litigios por lesiones personales, la determinación de las prestaciones por incapacidad y la indemnización de los trabajadores, la elegibilidad para los servicios sociales, los procedimientos penales, las evaluaciones militares y las evaluaciones para diagnósticos clínicos específicos que se asocian a incentivos externos, como los de lesión cerebral, discapacidad intelectual, dolor crónico y afecciones relacionadas, síntomas médicos o neurológicos inexplicables, TDAH y problemas de aprendizaje, entre otros.

B. Presentación no válida en la exploración que indica fingimiento o invención

En la exploración del examinado, existen incoherencias convincentes que indican exageración deliberada o invención de los déficits o síntomas o pruebas psicométricas de exageración o invención de los déficits o síntomas en las pruebas de validez del rendimiento (PVT) o las pruebas de validez de los síntomas (SVT).

Para cumplir los criterios de Presentación Inválida en Examen Indicativa de invención o exageración, el examinado debe presentar uno o más de los siguientes criterios.

1. *Presentación neurocognitiva no válida. Deben estar presentes uno o más de los criterios a, b o c.*
 - a. Durante la evaluación se observan o documentan una o más incoherencias convincentes relativas a déficits o síntomas cognitivos.
 - b. Puntuaciones no válidas en las PVT. Pruebas psicométricas de rendimiento de pruebas cognitivas no válidas basadas en (a) el uso de al menos dos o más PVT que, solas o combinadas, tengan una tasa baja de falsos positivos (es decir, .10), mientras que (b) se tenga en cuenta la relación entre las puntuaciones PVT fallidas y el número total de PVT administradas, (c) se minimice la redundancia de PVT y (d) se utilicen puntos de corte de PVT que hayan sido validados en estudios clínicos. La obtención de una PVT en el rango significativamente por debajo de la probabilidad también cumpliría este criterio (es decir, un rendimiento significativamente por debajo de la probabilidad en pruebas de elección forzada basadas en la teoría de la probabilidad binomial).
 - c. Pruebas psicométricas de síntomas cognitivos exagerados en las SVT auto-informadas. Pruebas psicométricas de síntomas cognitivos exagerados mediante SVT que, solas o combinadas, tienen una tasa baja de falsos positivos (es decir, .10). Por ejemplo, una o más puntuaciones de SVT que midan principalmente síntomas cognitivos fingidos o exagerados en el rango no válido utilizando pruebas con una tasa de falsos positivos aceptable; pruebas que proporcionen información no redundante; y pruebas que tengan puntos de corte que hayan sido validados en estudios clínicos cumplirían este criterio.
2. *Presentación inválida de síntomas somáticos. Deben estar presentes una o ambas de a o b.*
 - a. Se observan o documentan durante la evaluación una o más inconsistencias convincentes relativas a síntomas somáticos.
 - b. Evidencia psicométrica de síntomas somáticos exagerados en SVT auto-informados. Pruebas psicométricas de síntomas somáticos exagerados en las SVT que, solas o combinadas, tienen una tasa baja de falsos positivos (es decir, .10). Por ejemplo, una o más puntuaciones de SVT que midan principalmente síntomas somáticos fingidos o exagerados en el intervalo no válido utilizando (a) SVT con una tasa aceptable de falsos positivos, (b) SVT que proporcionen información no redundante y (c) SVT que tengan puntos de corte que hayan sido validados en estudios clínicos cumplirían este criterio.
3. *Presentación psiquiátrica no válida. Deben estar presentes una o ambas de a o b.*
 - a. Se observan o documentan durante la evaluación una o más inconsistencias convincentes relativas a síntomas psiquiátricos.
 - b. Evidencia psicométrica de síntomas psiquiátricos exagerados en las SVT auto-informadas. Evidencia psicométrica de informe de síntomas exagerados utilizando SVT que solos o en combinación tienen una tasa baja de falsos positivos (es decir, .10). Por ejemplo, una o más puntuaciones de pruebas de validez de síntomas que midan principalmente síntomas psiquiátricos fingidos o exagerados en el rango no válido utilizando pruebas con una tasa aceptable de falsos positivos; pruebas que proporcionen información no redundante; y pruebas que tengan puntos de corte que hayan sido validados en estudios clínicos cumplirían este criterio.
4. *Presentación de síntomas mixtos no válida. Evidencia de inconsistencia convincente y/o evidencia psicométrica de resultados PVT o SVT inválidos o exagerados en dos o más de los dominios cognitivos, somáticos o cognitivos. Por ejemplo, lo siguiente satisfaría este criterio:*
 - Dos o más incoherencias convincentes entre dominios (es decir, dos o más de B1a, B2a y B3a).
 - Pruebas psicométricas en más de un dominio (es decir, dos o más entre B1b, B1c, B2b y B3b).
 - Una o más inconsistencias convincentes combinadas con evidencia psicométrica de déficits o síntomas inválidos o exagerados en uno o más dominios (es decir, una o más inconsistencias convincentes con uno o más de cualquiera de B1b, B1c, B2b o B3b).

C. Discrepancias marcadas

Existen una o más discrepancias marcadas entre los datos obtenidos de la prueba/informe de síntomas y los tipos de pruebas, según se indica a continuación:

1. *Historia natural y patogénesis de la afección*
La información obtenida por autoinforme o mediante pruebas o escalas es marcadamente discrepante de los modelos actualmente aceptados de funcionamiento neurológico, médico o psiquiátrico normal y anormal, de un modo que sugiere invención o exageración de déficits o síntomas.
2. *Registros y otros medios*
La información obtenida mediante autoinforme o a través de pruebas o escalas es marcadamente incoherente con los registros u otros antecedentes documentados (p. ej., audio, vídeo, medios sociales) de un modo que sugiere invención o exageración de déficits o síntomas.
3. *Informe de informante colateral fiable*
La información obtenida por autoinforme o mediante pruebas o escalas es notablemente discrepante del nivel de funcionamiento cotidiano descrito por al menos un informante colateral fiable con un interés mínimo en el resultado de la evaluación, de forma que sugiere invención o exageración de la disfunción.

D. Las conductas que cumplen el Criterio B no se explican completamente por otra condición de desarrollo, médica o psiquiátrica

Se asume que las conductas que cumplen el Criterio B reflejan un intento informado, racional y volitivo de adquirir o lograr resultados tal y como se definen en el Criterio A y no pueden explicarse completamente por condiciones significativas de desarrollo, médicas o psiquiátricas que resulten en una capacidad significativamente disminuida para apreciar las leyes o costumbres contra la incapacidad para ajustar la conducta a tales estándares. A continuación, se enumeran ejemplos de trastornos médicos, psiquiátricos y del desarrollo importantes:

- Demencia de moderada a grave.
- Discapacidad intelectual de moderada a grave (por ejemplo, CI < 60).
- Trastornos psiquiátricos, neurológicos u otros trastornos médicos graves asociados a un deterioro cognitivo suficiente para impedir la independencia en las actividades básicas de la vida diaria. El engaño puede coexistir con trastornos asociados a déficits cognitivos como la discapacidad intelectual leve, la demencia leve o el deterioro cognitivo leve. Del mismo modo, puede coexistir con trastornos psiquiátricos o neurológicos definidos por síntomas somatomorfos (por ejemplo, trastorno de síntomas somáticos, trastorno de conversión/trastorno de síntomas neurológicos funcionales, trastorno facticio, síntomas médicos inexplicables) y en presencia de otros trastornos psiquiátricos (por ejemplo, depresión).

Especificadores

A continuación, se describen los cuatro especificadores para la presentación clínica de la simulación.

Simulación de disfunción neurocognitiva

Además de cumplir los criterios A, C y D, el individuo cumple los criterios B1a, B1b o B1c de disfunción cognitiva fingida o exagerada.

Simulación de síntomas somáticos

Además de cumplir los Criterios A, C y D, el individuo cumple los Criterios B2a o B2b de síntomas somáticos fingidos o exagerados.

Simulación de síntomas psiquiátricos

Además de cumplir los criterios A, C y D, el individuo cumple los criterios B3a o B3b de síntomas psiquiátricos fingidos o exagerados.

Simulación con presentación mixta

Además de cumplir los criterios A, C y D, el individuo cumple el Criterio B4 de síntomas fingidos o exagerados en más de un ámbito (es decir, cognitivo, somático y/o psiquiátrico).