

Desarrollo y evaluación de un procedimiento empírico para detección de la simulación de enajenación mental en el contexto legal

Ramón Arce
María del Carmen Pampillón
Universidad de Santiago de Compostela
Francisca Fariña
Universidad de Vigo

En un contexto de simulación de enfermedad mental no imputable, pedimos a una muestra de sujetos, expertos y no expertos en psicopatología, que simularan una enfermedad mental no imputable en el MMPI al que previamente habían respondido con normalidad. Los resultados mostraron que la simulación era muy accesible (el 78.8% de los sujetos podían simular correctamente), y no se observan diferencias entre expertos y no expertos. Las escalas de control del MMPI, al igual que las configuraciones de las escalas de validez, evidenciaron ser efectivas aunque dejaban un margen de error, al no clasificar como simuladores al 7.5% de los sujetos que habían simulado una enajenación mental incapacitante. Además, en sí mismas pueden ser un indicador débil al dar cabida al estudio de otras hipótesis alternativas. Así, recurrimos a otro método de medida complementario, la entrevista clínica abierta. Este procedimiento resultó ser, en sí, mucho más difícil de solventar por los simuladores ya que el 58.8% no fueron capaces de simular trastorno incapacitante. Cifra que se elevaba hasta el 71.25% si considerábamos la detección de estrategias de simulación en la entrevista, pero, aun así, el margen de error sería desorbitante, el 28.75. En consecuencia, contrastamos el poder de una aproximación multimétodo, MMPI más entrevista. Esta clasificaba correctamente a todos los sujetos con, al menos, dos medios de detección. Por ello, postulamos la eficacia de una «invalidez convergente» sustentada en, cuando menos, dos métodos de medida de la inconsistencia para etiquetar un protocolo como simulación. Además, para el control de falsos positivos recurrimos de modo complementario al Modelo de Decisión Clínica para el establecimiento de la simulación (Cunnie, 1997) Por último, se esboza un pro-

toloco de actuación en el contexto legal para la detección de la simulación y falsos positivos.

Palabras clave: simulación, enajenación mental, MMPI, psicología forense, imputabilidad.

In a simulated criminal insanity task, a sample of experts in psychopathology and laypeople were asked to simulate a mental disorder in order to avoid criminal responsibility. Subjects had previously been evaluated by MMPI before the test and had provided responses within normal ranges. The results reveal a high level of malingering (78.8% of the subjects could simulate insanity), and no differences were observed between laypeople and experts. The effectiveness of the control scales of the MMPI and the configuration of the scales of validity was confirmed though they have a margin of error, since they fail to classify 7.5% of subjects who simulated criminal insanity in order to be acquitted. Moreover, the scales are not a robust index because alternative hypotheses need to be studied. A complementary instrument was therefore used, an open clinical interview with the same format as the cognitive interview. This procedure was much more difficult to simulate, since 58.8% were incapable of simulating insanity; the figure rose to 71.25% when considering the absence of internal consistency in the interview. Nevertheless, the margin of error was significant: 28.75. Consequently, a multimethod approach (MMPI and interview) was compared, and found to classify all subjects correctly. Thus, we recommend a «convergent invalidity» procedure based on at least two methods of assessing inconsistency in order to label a protocol as a malingerer. Our proposal also includes the use of the Clinical Decision Model (Cunnién, 1997) for the detection of false positives. Decision Practical guidelines are established for the detection of simulated and false positives in the legal context.

Key words: malingering, faking, mental disorder, criminal responsibility, forensic psychology, MMPI.

En los últimos años hemos asistido al resurgimiento con inusitada fuerza de la Psicología Jurídica. Este relanzamiento no sólo tuvo sus efectos en la investigación (Garrido, 1994) y docencia universitaria sino que también vino acompañado de la oferta y posterior demanda de informes para la Sala de Justicia, esto es, el reconocimiento de la Psicología Jurídica (Kury, 1997; Ibáñez y de Luis, 1994). En esta última dirección, la situación está tan consolidada como se pone de manifiesto en hechos tan evidentes como que nos encontramos con que las diversas delegaciones del Colegio Oficial de Psicólogos cuentan con listas de peritos a disposición de las demandas judiciales; o con que se han creado cuerpos de psicólogos forenses. Estos requerimientos han supuesto nuevos interrogantes a resolver por parte de los peritos psicólogos: el mejor interés del menor, el valor del daño moral, la incapacitación civil, la nulidad matrimonial, y un largo etcétera. Entre ellos, uno de los interrogantes a resolver es la evaluación de la enfermedad mental. En otras palabras, la cuestión planteada consiste en determinar, mediante una evaluación, la enfermedad mental del encausado (según la jurisprudencia, «evaluación biológica») y, en su caso, cómo afecta a las capaci-

dades cognitivas o volitivas (legalmente, «evaluación psicológica» que realiza el perito sobre la base de la evaluación biológica para el sujeto y delito concretos) que son el substrato real de la imputabilidad (sobre estas bases, la tipología de enfermedad mental y la imputabilidad legal véase la revisión de la jurisprudencia en de la Torre, 1999; Carbonell, Gómez y Mengual, 1987).

Si bien la psicología está bien dotada para la evaluación del trastorno mental con numerosas pruebas psicométricas y formatos de entrevistas, no lo está tanto para la detección de una característica singular del contexto judicial, la simulación. De hecho, Rogers (1997) advierte, tal y como ocurre en el campo clínico, que «si no investigamos la simulación no podremos encontrarla» (p. 5), no habiendo hallado referencias a casos clínicos diagnosticados como simulación. En esta línea, el DSM-IV (American Psychiatric Association, 1995a, 1995b) remite la simulación a un apartado inespecífico bajo la denominación general, «Problemas adicionales que pueden ser objeto de atención clínica», aconsejando su consideración en, entre otros, un contexto médico-legal. Así pues, el objeto de la pericia no sólo es la evaluación del estado mental, sino desechar que se trate de una simulación.

La literatura científica ha llegado a la conclusión de que es preciso adoptar una perspectiva de medida multimétodo para la identificación de la simulación (Rogers, 1997). La eficacia del MMPI, tanto en la versión original como en la revisada o en la versión para adolescentes, en la evaluación de simulación ha sido confirmada sistemáticamente (Lyons y Wheeler, 1999; Pope, Butcher y Seelen, 2000) tanto con muestras con sujetos entrenados para fingir síntomas psicopatológicos como con sujetos extraídos de poblaciones clínicas o forenses (véase, para una revisión, Aldea 1994; Rogers, 1997). En comparación con otros procedimientos de medida, se ha encontrado sistemáticamente que el MMPI es el instrumento más fiable para la detección de simulación (Shuman, 1994). Tal eficiencia se basa en las escalas de control de la validez y en las configuraciones de estas escalas. Dentro de estas últimas, dos se han mostrado efectivas en la detección de la simulación de trastornos (Duckworth y Anderson, 1995): el índice F-K, y el perfil en «V invertida». Se ha observado que el índice F-K, también conocido como «índice de Gough», es un indicador robusto de los intentos de proporcionar una mala imagen, en tanto el perfil en «V invertida» advierte de la exageración de síntomas (Nicholson *et al.*, 1997). No obstante, este instrumento no es suficiente para establecer fehacientemente simulación porque el diagnóstico de simulación es compatible con la formulación de otras hipótesis alternativas a la simulación (Graham, 1992; Roig Fusté, 1993) y porque no clasifica correctamente a todos los simuladores (Bagby, Buis y Nicholson, 1995). En otras palabras, basándonos sólo en este instrumento se pueden distinguir dos tipos de errores: falsos positivos (catalogar a enfermos reales como simuladores) y errores de omisión (no detectar como simuladores a sujetos que realmente están simulando). Todo ello ha llevado a que se proponga la adopción de una estrategia de evaluación multimétodo a fin de minimizar estas fuentes de error (Rogers, 1997).

Por ello nos planteamos llevar a cabo un estudio experimental con el fin de perfilar la accesibilidad de la simulación de enfermedad mental, y, subsiguientemente, de enfermedad que mitigue o exima la responsabilidad penal de un acu-

sado/condenado, así como de los modos de detección de la misma con el fin último de aislarla. Adoptaremos, en consonancia con la literatura precedente, una perspectiva de evaluación multimétodo. En suma, pretendemos la creación y contrastación de un procedimiento de medida de la enfermedad mental no imputable en el contexto penal español que sea sensible a los intentos de simulación por parte del encausado/acusado. Acotaremos el término enfermedad mental a lo que se conoce en la literatura jurídica (véase Carbonell, Gómez y Mengual, 1987) como «el supuesto de enajenación mental» excluyendo la oligofrenia (retraso mental) que sigue otros procedimientos de medida y que, hoy en día, se puede verificar en pruebas documentales. También se descartarán otros supuestos como el Trastorno Mental Transitorio, las alteraciones de la percepción o la minoría de edad, bien por ser de difícil medida objetiva, y de relación con los hechos y demarcación temporal, bien por caer fuera del alcance de la evaluación psicológica. Así pues, nos delimitaremos a la simulación de la esquizofrenia o paranoia.

Método

Sujetos

Un total de 80 sujetos componen la muestra. La mitad, 40, tenían conocimientos de psicología y, más en concreto, de psicopatología, asignatura que habían cursado en la carrera de Psicología en la Universidad de Santiago. A éstos nos referiremos como expertos ya que tendrían facilidades para poder conocer los modos de simulación de enfermedad mental. Los restantes 40 no poseían conocimientos específicos de psicología, por lo que se consideran no expertos. Del total, 29 eran varones, el 36.25%, en tanto 51 eran mujeres, el 63.75% de la muestra. En relación con la edad, el rango oscilaba entre los 19 y los 45 años, con una media de 25.09 ($s_x = 5.32$). Por estudios, los sujetos eran mayoritariamente de nivel universitario (un total de 64, el 80%), con 14, un 17.5%, de nivel medio y tan sólo 2, un 2.5%, de nivel bajo (p. ej. estudios primarios).

Instrumentos de medida

Como instrumento psicométrico de medida utilizamos el MMPI original con los baremos españoles (Tea, 1988), centrándonos únicamente en los resultados de los sujetos en las «escalas de validez», orientadas a la detección de anomalías en las respuestas de los sujetos, y en el «perfil» obtenido a partir de sus respuestas en las «escalas clínicas», que nos indicaría qué trastorno o trastornos padece/simula el sujeto. Las escalas de validez son cuatro: interrogantes (CS), sinceridad (L), validez (F) y el factor corrector (K). Las escalas clínicas son diez: hipocondriasis (Hs), depresión (D), histeria (Hy), desviación psicopática (Pd), masculinidad-feminidad (Mf), paranoia (Pa), psicastenia (Pt), esquizofrenia (Sc), hipomanía (Ma) e introversión social (Si). No tomamos la nueva versión, el

MMPI-2, ya que, en lo referente a las escalas clínicas y de control, es convergente con la original (Ben-Porath y Butcher, 1989). Así, optamos por la versión original por estar extendida y baremada en más países, lo que confiere a nuestros resultados poder de generalización al contexto hispanohablante.

Además, y con el fin de poder contrastar la validez convergente de las medidas, también se evaluó a los sujetos por medio de una «*entrevista de orden clínico en formato de discurso libre*». El proceder consistió en pedir a los sujetos que relataran los síntomas, conductas y pensamientos que tenían en el momento de cometer el crimen (esto es, EEAG en el eje V del DSM-IV). Si los sujetos no respondían de *motu proprio*, les era requerido por medio de preguntas abiertas, de acuerdo con el eje V del DSM-IV, que informaran sobre sus relaciones familiares (EEGAR) y relaciones sociales (EEASL). El objetivo fundamental de este proceder era, además de acercarnos a una determinación de la validez convergente, enfrentar al sujeto a una tarea distinta a la requerida en el MMPI. En éste proceden a una tarea de reconocimiento de síntomas, en tanto en la entrevista clínica se les demanda un conocimiento de síntomas. Por ello la entrevista no es en formato de interrogatorio sino de tipo «no directiva» y orientada a la reinstauración de contextos. Es decir, seguimos el procedimiento de la entrevista cognitiva (Geiselman y otros, 1984; Fisher y Geiselman, 1992) en lo relativo a la reinstauración de contextos con relevancia clínica tal y como se recoge en el DSM-IV. Este mismo procedimiento se mostró fiable, válido y productivo en la detección de simulación de un trastorno de estrés postraumático ligado a una falsa agresión sexual o intimidación (Freire, 2000; Arce, Fariña y Freire, en prensa).

Procedimiento y diseño

En un primer momento, los sujetos tuvieron que responder al cuestionario MMPI bajo las instrucciones estándar, esto es, todos ellos fueron instruidos tal como recoge el manual del instrumento. Los expertos lo cumplimentaron en grupo como una práctica en la asignatura de Psicología Jurídica, en tanto los no expertos lo cubrieron bien individualmente o en pequeños grupos. Esta evaluación la tomamos con tres objetivos: *a*) como medida de control a fin de detectar, y en su caso eliminar, posibles enfermos mentales o respuestas distorsionadas ($CS \geq 30$); *b*) como medida de contraste con la que abordar el estudio de la capacidad de simulación en el MMPI); y *c*) como medida para contrastar con la respuesta de simulación a fin de poder conocer de las habilidades de simulación de los sujetos (Bagby, Buis y Nicholson, 1995).

En un segundo momento, a los sujetos del estudio se les pidió que se colocaran en la tesitura de haber cometido fehacientemente un delito extremadamente grave (p. ej., homicidio). Por ello, iban a ser sometidos a una evaluación psicológica con la pretensión de conocer su estado real de salud mental. Tal evaluación tenía su origen en la demanda del abogado defensor que presumía que su defendido, esto es, el sujeto del estudio, padecía una enfermedad mental severa. Obviamente, éste no era el caso (controlado a través de la medida bajo instrucciones estándar del MMPI), con lo que, a fin de poder seguir la estrategia de evitación de

una condena, debía simular padecer una enfermedad mental que lo etiquetara como no imputable. En otras palabras, le demandábamos que se viera en la situación de haber cometido un delito muy grave con pruebas irrefutables que lo condenaban, siendo la alternativa más fructífera ser declarado «culpable, pero mentalmente enfermo» y, por tanto, no responsable (instrucciones de simulación). Con objeto de acercarnos más a un contexto real, se les proporcionó un tiempo para que reflexionaran, se pusieran en situación y se formaran en lo que ellos considerasen para simular una enfermedad mental incapacitante (el tiempo dado no era, en ningún caso, inferior a 7 días y tampoco se demoraba mucho más). Tras este tiempo, se les administró el MMPI bajo las «instrucciones de simulación».

En un tercer momento, en torno a dos semanas y sin previo aviso, los participantes fueron evaluados por medio de una entrevista en formato de discurso libre en la que debían responder, en línea con la simulación anterior en el MMPI (instrucciones de simulación), a cómo se sentían, cómo pensaban, acerca de su estilo de vida, hábitos, etc., en el momento de los hechos que se le imputan. Se recurrió a este tipo de entrevista, grabando su discurso en vídeo, para posteriormente analizarlo y codificarlo, con el fin de detectar qué trastorno simula el sujeto sobre la base de los criterios diagnósticos recogidos en el DSM-IV. Estos datos se contrastaron con los obtenidos a través del MMPI, para evaluar el grado de consistencia de su simulación a lo largo del tiempo y a través de diferentes medidas. En la entrevista se evaluaría, a través de un análisis de contenido, la presencia de alguna estrategia de simulación descrita en la literatura.

En suma, el diseño consistió en dos medidas con el MMPI (instrucciones estándar e instrucciones de simulación) y una medida a través de entrevista para poder analizar la consistencia de las respuestas inter-medidas. Además, se consideraron dos niveles en el factor de agrupamiento: *a) sujetos legos* ($n = 40$), formado por sujetos sin conocimientos sobre psicopatología; y *b) sujetos expertos* ($n = 40$), compuesto por estudiantes de últimos cursos de psicología y con conocimientos de psicopatología.

Análisis de los protocolos

Las entrevistas en formato de discurso libre, tras ser grabadas en vídeo, fueron sometidas a un análisis de contenido sistemático. El objeto de este análisis de contenido era detectar criterios diagnósticos de enfermedad mental. Las categorías de análisis fueron tomadas del DSM-IV. Así, creamos un sistema categorial mutuamente excluyente, fiable y válido, en lo que Weick (1985) ha denominado sistemas de categorías metódicas. Del total de criterios posibles, sólo se consideraron aquellos que se mostraron productivos (Wicker, 1975)¹.

Si bien la gran mayoría de los síntomas, incluidos los más adversos, pueden ser informados directamente por los sujetos (Lewis y Saarni, 1993), algunos sólo pueden observarse. En consecuencia, la detección de las categorías respon-

1. No se expone el listado completo por ser demasiado amplio. El lector interesado puede solicitarlo directamente a los autores.

día a dos métodos complementarios: expresión directa del sujeto e inferencias de los codificadores tras analizar los protocolos. Por ejemplo, el deterioro de memoria puede ser manifestado directamente por el sujeto o ser inferido por el codificador tras la entrevista.

Realizado el análisis de contenido de la entrevista de cada sujeto, los codificadores evaluaban si, a tenor de los criterios detectados, se cumplía uno o más trastornos, y, en su caso, si dicho trastorno era motivo de inimputabilidad. Finalmente, valoraban la existencia de las diferentes estrategias de simulación (presencia vs. ausencia) en el discurso del sujeto sobre la base de la identificación de alguna de las seis estrategias que la literatura ha descrito que siguen los simuladores y que son productivas en formato de entrevista no directiva: síntomas raros, combinación de síntomas, síntomas obvios, consistencia de síntomas, síntomas improbables, y severidad de síntomas (véase, para una definición y ejemplos de cada categoría, Rogers y Mitchell, 1991).

Entrenamiento de codificadores

En este estudio participaron cuatro codificadores expertos y con conocimientos de evaluación psicopatológica, que habían colaborado previamente en otras codificaciones, mostrándose fiables (Jólluskin, 2000; Freire, 2000). Los cuatro fueron exhaustivamente entrenados en éste y otros sistemas de codificación. El entrenamiento consistió, tras presentarles y ejemplificarles cada categoría de análisis, en la ejecución con material del estudio que no iban a codificar posteriormente, utilizando el índice de concordancia [$IC = \text{Acuerdos}/(\text{acuerdos} + \text{desacuerdos})$] como instrumento de cotejo de la ejecución, lo que permitiría, si se constataba inconsistencia con los criterios sostenidos por los experimentadores, subsanar errores de codificación a través de la discusión en grupo y la homogeneización de criterios. Las definiciones de las categorías de análisis fueron tomadas de los criterios diagnósticos del DSM-IV (1995a, 1995b); y las definiciones y ejemplos de cada categoría de análisis en la detección de las estrategias de simulación de Rogers y Mitchell (1991).

Fiabilidad

Los codificadores se dividieron en dos grupos, con dos codificadores cada uno, que valoraban conjuntamente y de forma consensuada su observación. Cada grupo codificó el 50% de las entrevistas, incluyendo cada uno la mitad de sujetos legos y la mitad de expertos. De este modo, cada grupo analizó 40 de las 80 entrevistas, 20 de sujetos *legos* y otras 20 de *expertos*. Primero codificaron los protocolos en los criterios diagnósticos del DSM-IV. En un segundo momento temporal, se repitió el mismo procedimiento en relación a la evaluación de las estrategias de simulación en los protocolos.

Para calcular la fiabilidad intrajueces, los codificadores de cada grupo repitieron, una semana después del fin de la codificación original, el 10% de las

entrevistas que ellos mismos habían codificado. Asimismo, para obtener la fiabilidad interjueces los codificadores de cada grupo codificaron el 10% de los protocolos que inicialmente había codificado el otro grupo.

Se consideran evaluaciones concordantes aquellas que superan el punto de corte .80 (Tversky, 1977), que es más restrictivo que los valores kappa. Contrastados nuestros resultados², en los 93 criterios diagnósticos del DSM-IV que se mostraron productivos y en la evaluación de las estrategias de simulación en los protocolos, podemos sostener que las evaluaciones son consistentes inter- e intra-evaluadores, en el tiempo e inter-contextos. En suma, los datos son fiables (Wicker, 1975).

Hipótesis

Hipótesis 1.- Esperamos que los sujetos, tanto legos como expertos, sean capaces de simular enfermedad mental en el MMPI (Rogers, 1997).

Hipótesis 2.- Hipotetizamos que los sujetos podrán simular en las escalas clínicas del MMPI una enfermedad mental de tipo no imputable, en concreto, enajenación mental esquizofrénica o paranoide.

Hipótesis 3.- Prevedemos, de acuerdo con la «hipótesis de la metasimulación» (Porot, 1977; Aldea, 1994), que los expertos simulen en el MMPI con mayor efectividad que los legos una enfermedad mental.

Hipótesis 4.- Inferimos que las escalas de control del MMPI serán sensibles, en general, a la simulación (Lyons y Wheeler, 1999; Pope, Butcher y Seeley, 2000).

Hipótesis 5.- Pensamos, contrariamente a los postulados del modelo criminológico, que el simulador no adoptará como estrategia la evitación de la evaluación (Rogers, 1992; Lewis y Saarni, 1993; Bagby *et al.*, 1997) porque no asociará una evaluación de enfermedad mental no imputable a la no-respuesta a la evaluación, esto es, en la escala de interrogantes (CS) presentará puntuaciones inferiores al punto de corte establecido para la desestimación de la medida (<30).

Hipótesis 6.- Aguardamos que los simuladores no presenten, en su intento por simular una enfermedad mental no imputable, una imagen de deseabilidad social (imagen positiva de sí mismo), o sea, las puntuaciones en la escala L (mentira) estarán dentro de la normalidad o serán bajas ($T < 70$). Obviamente, el simulador no debe intentar presentar una imagen positiva de sí mismo porque sería contrario a la simulación de una enfermedad mental dirigida a la obtención de un beneficio legal a través bien de la atenuación o exoneración de la responsabilidad penal. Por el contrario, la deseabilidad social está ligada a la disimulación (Baer, Wetter y Berry, 1995).

Hipótesis 7.- Confiamos en que la escala F, encaminada a evaluar respuestas inusuales, sea sensible a la simulación, esto es, que los simuladores de enfermedad mental incapacitante presenten puntuaciones elevadas en esta escala ($T > 70$).

2. No exponemos los datos directos por ser excesivamente amplios y carecer de valor en sí. Éstos pueden solicitarse a los autores.

Hipótesis 8.— Estimamos que el factor corrector K será un detector robusto de invalidez del protocolo, ya que se orienta a la detección de simulación de un deterioro.

Hipótesis 9.—Entendemos que las configuraciones de las escalas de validez del MMPI relevantes para el estudio de la simulación, el índice F-K y el perfil en «V invertida» detectarán efectivamente la simulación (Duckworth y Anderson, 1995).

Hipótesis 10.— Creemos que la manipulación de enfermedad mental incapacitante no será accesible para los sujetos en la entrevista, contrariamente a las escalas clínicas del MMPI, al implicar una tarea de conocimiento e implementación de los síntomas en la vida diaria.

Hipótesis 11.— El estudio de las estrategias de simulación en los protocolos de las entrevistas, junto con el poder discriminativo de la entrevista clínica no directiva, supondrán un indicador robusto de simulación de trastorno mental no imputable.

Resultados y conclusiones

Estudio de la simulación en las escalas clínicas del MMPI

El primer interrogante a responder era conocer si los sujetos, en general, eran capaces de simular una enfermedad mental para, posteriormente, perfilar si ésta era de tipo incapacitante, y si la simulación resultaba accesible para expertos y legos.

Los resultados muestran que, tomando el punto de corte de 70 en puntuaciones típicas, esto es, el referente a partir del cual se considera patológico un sujeto (Roig Fusté, 1993), los sujetos son capaces de simular efectivamente desviación psicopática (Pd), paranoia (Pa) y esquizofrenia (Sc) (véanse las Tablas 1 y 2). Por submuestras, los legos simulan con efectividad paranoia (Pa) y esquizofrenia (Sc), en tanto apreciamos una patología en desviación psicopática marginalmente significativa. Por su parte, los expertos sobrepasan significativamente el punto de corte patológico en desviación psicopática (Pd), paranoia (Pa) y esquizofrenia (Sc).

La desviación psicopática hace referencia básicamente a un comportamiento antisocial, ya que tiene su origen en el intento de perfilar a aquellos sujetos con problemas continuos con la ley, con carencia de control de las respuestas emotivas, incapaces de obtener provecho de la experiencia, así como de establecer relaciones sociales cálidas y profundas. Es importante resaltar que esta dimensión se perfila al margen de la privación cultural, de trastornos de ansiedad (anteriormente neuróticos) y psicóticos, y en personas de inteligencia normal. En suma, de esta evaluación no se desprende que el sujeto vea alteradas por enfermedad mental las capacidades cognitivas o volitivas. En consonancia con todo ello, la jurisprudencia del Tribunal Supremo (TS) califica la psicopatía de enfermedad mental endógena (4 de octubre de 1982, RA 5593), precisando su alcance

en la consideración como «simples atipias caracterológicas no llevan aparejadas necesariamente disminuciones del campo de la conciencia en la esfera volitiva y afectiva del sujeto» (Sentencia del TS de 8 de abril de 1980, RA 1097). En resumen, la psicopatía, en genérico, no se entiende como causa de no imputabilidad. De hecho, no existe base de evaluación alguna para derivar una limitación en las capacidades cognitivas o volitivas de los psicópatas. Una sentencia del TS de 22 de abril de 1982 resume muy acertadamente el estado de la cuestión: «la doctrina legal, de acuerdo con la ciencia psiquiátrica, tiene reiteradamente declarado que los psicópatas no son enfermos mentales, sino que pueden estimarse como personas normales a los efectos de la imputabilidad por no tener limitadas sus facultades intelectivas y volitivas, sin perjuicio de que a efectos extrapenales puedan ser psiquiátricamente calificados de anormales» (RA 2104).

La medida de la paranoia (Pa) se ha edificado sobre la base de pacientes con ideas o delirios persecutorios, de referencia o grandeza, hipersensibilidad e ingenuidad. Puntuaciones elevadas en esta escala indican perfeccionismo, rigidez, preocupación, terquedad, proyección de culpa y hostilidad (Roig-Fusté, 1993) o posible psicosis paranoide (Graham, 1990). Aisladamente no es un indicador fiable de no imputabilidad. De nuevo, la jurisprudencia camina, usualmente, en esta dirección tal y como se desprende, por ejemplo, de la Sentencia del TS de 15 de octubre de 1984 (RA 4839) que establece: *a)* el método biológico de diagnosticar paranoia es irrelevante para la imputabilidad si no perturba la inteligencia y voluntad del sujeto; *b)* debe aplicarse un método de evaluación biológico-psicológico en el que se diagnostique la enfermedad, las áreas a las que afecta y la relación entre el hecho punible y el delirio. El objeto de la evaluación psicológica debe dirigirse a la identificación de delirios ya que éstos guían, más allá del conocimiento de que el acto es ilícito, al enfermo al delito (de la Torre, 1999).

La escala de esquizofrenia (Sc), confeccionada a partir de pacientes psiquiátricos caracterizados por la presencia de comportamientos y pensamientos

TABLA 1. MEDIAS Y DESVIACIONES TÍPICAS EN LAS ESCALAS CLÍNICAS DEL MMPI.
RESPUESTAS SIMULADAS

Variable	Media	s_x	Media	s_x	Media	s_x
	Total (n = 80)	Total	Legos (n = 40)	Legos	Expertos (n = 40)	Expertos
Hipocondría	65.03	18.62	67.63	17.09	62.43	19.91
Depresión	64.80	15.07	67.25	14.41	62.35	15.49
Histeria	59.33	13.48	60.33	13.37	58.33	13.69
Desviación psicopática	75.61	12.62	73.88	13.32	77.35	11.79
Masculinidad-feminidad	57.13	12.50	55.58	11.53	58.68	13.38
Paranoia	81.22	14.38	80.00	14.16	82.45	14.68
Psicastenia	63.24	11.63	64.38	11.09	62.10	12.17
Esquizofrenia	80.78	12.49	79.55	12.87	82.00	12.14
Hipomanía	64.91	13.45	62.23	14.14	67.60	12.31
Introversión social	66.71	12.69	69.45	10.33	63.98	14.29

insólitos o extravagantes, experiencias inusuales, apatía, indiferencia y aislamiento emocional, hace referencia a sujetos con graves problemas de conducta. En relación con la imputabilidad, los esquizofrénicos presentan distorsiones en la percepción (área cognitiva) y escaso control de los impulsos (área volitiva). Obviamente, los esquizofrénicos pueden ser no imputables. A la sazón, la jurisprudencia (Sentencia del TS de 22 de diciembre de 1984, RA 6633) excluye incluso la necesaria vinculación directa entre comportamiento punible y enfermedad porque ésta está siempre presente (Sentencia del TS de 2 de noviembre de 1979, RA 3799). Esto es, sólo sería menester una evaluación de orden biológico sin necesidad de precisar las implicaciones psicológicas.

TABLA 2. T PARA UNA MUESTRA CON EL VALOR DE PRUEBA DE 70. RESPUESTAS SIMULADAS

Variable	T*	p	T**	p	T**	p
	Total	Total	Legos	Legos	Expertos	Expertos
Hipocondría	-2.390	.019	-.879	.385	-2.407	.021
Depresión	-3.087	.003	-1.207	.235	-3.124	.003
Histeria	-7.082	.000	-4.578	.000	-5.392	.000
Desviación psicopática	3.978	.000	1.840	.073	3.944	.000
Masculinidad-feminidad	-9.210	.000	-7.915	.000	-5.355	.000
Paranoia	6.981	.000	4.466	.000	5.365	.000
Psicastenia	-5.203	.000	-3.208	.003	-4.106	.000
Esquizofrenia	7.717	.000	4.694	.000	6.254	.000
Hipomanía	-3.384	.001	-3.477	.001	-1.234	.225
Introversión social	-2.317	.023	-.337	.738	-2.667	.011

* g.l. (79); ** g.l. (39).

El estudio de casos en relación con la imputabilidad pone de manifiesto que, en total, 62 sujetos simulan esquizofrenia, de los cuales 29 son legos y 33 expertos, siendo igual el número de legos y expertos simuladores, $\chi^2_1(p = 0.258)$. En relación con la paranoia, 67 sujetos muestran una puntuación T en esta escala superior a 70. De éstos, 32 son legos y 35 expertos. Nuevamente, no encontramos el previsible mayor poder de simulación de los expertos frente a los legos, $\chi^2_1(p = 0.134)$.

En resumen, los sujetos son capaces, a tenor de la jurisprudencia revisada, de simular efectivamente una enfermedad mental que les convierta en no imputables en instrumentos psicométricos fiables como el MMPI. Se confirman, pues, las hipótesis 1 y 2, resultando incluso más fiables los sujetos legos al descartar, en mayor medida, la desviación psicopática como enfermedad mental no imputable. Además, no sigue una estrategia de simulación total, esto es, de puntuar positivamente en todos los síntomas. A nuestro modo de ver esta facilidad para la simulación se sustenta en el hecho de que no es necesario un conocimiento explícito de los síntomas al estar frente a una tarea de reconocimiento y no de conocimiento.

Estudio del cambio en la evaluación clínica en el MMPI

Conocida la posibilidad de que los sujetos simulen enfermedad mental, en instrumentos contrastados, es necesario tomar una salvaguarda, ¿es significativo el cambio de la normalidad a la simulación? Por ello, ejecutamos un MANOVA con el factor instrucciones (instrucciones estandar vs. instrucciones de simulación) y como variables dependientes las escalas clínicas del MMPI. Los resultados evidencian un cambio significativo de la condición de ejecución bajo las instrucciones estándar de cumplimentación del MMPI a las instrucciones de simulación de un trastorno mental no imputable en términos de patología, $F_{\text{multivariada}}(10,149) = 59.222$; $p^{***} < .001$; T.E. = .799. Por su parte, los efectos univariados (véase la Tabla 3) exhiben diferencias en todas las variables en el sentido de que los sujetos pueden manipular sus indicadores clínicos de patología. Tomando como punto de corte una $T > 70$, observamos que los perfiles clínicos salientes son desviación psicopática (PD), paranoia (PA) y esquizofrenia (SC). Esto es, simulan una situación comórbida y efectiva en términos de la no imputabilidad, desechando una simple combinación de síntomas. Todo ello presta apoyo experimental a las hipótesis 1 y 2.

TABLA 3. CONTRASTES UNIVARIADOS, INTRASUJETOS

Variable	MC	F	p	η^2	$\bar{x}_{\text{normalidad}}$	$\bar{x}_{\text{simulado}}$
Hipocondría	6451.6	26.684	.000	.252	52.325	65.025
Depresión	5856.4	31.508	.000	.285	52.700	64.800
Histeria	1305.306	10.372	.002	.116	53.613	59.325
Desviación psicopática	20205.025	184.724	.000	.700	53.137	75.613
Masculinidad-feminidad	2472.756	19.855	.000	.201	49.263	57.125
Paranoia	36966.4	35.397	.000	.309	50.825	81.225
Psicastenia	10416.756	102.766	.000	.565	47.100	63.238
Esquizofrenia	45630.025	465.133	.000	.855	47.000	80.775
Hipomanía	15820.506	119.68	.000	.602	45.025	64.912
Introversión social	15464.556	116.486	.000	.596	47.050	66.713

g.l. (1,79).

Estudio de las variables conocimientos y simulación

Asimismo, ejecutamos un MANOVA con el factor submuestra (expertos vs. legos) y como variables dependientes las escalas clínicas del MMPI. Los resultados muestran que el factor submuestra no media diferencias en la simulación de un estado psicopatológico de inimputabilidad, $F_{\text{multivariada}}(10,69) = 1.345$; T.E. = .163. A su vez, los efectos univariados (constátense en la Tabla 4) no muestran efectos significativos en ninguna de las variables que conforman el estudio patológico o, lo que es lo mismo, expertos y legos simulan por un igual. No encontramos, por tanto, apoyo a la hipótesis generalmente asumida de que los expertos tendrían mayor capacidad de simulación efectiva (hipótesis 3). Dos son

las posibles interpretaciones a estos resultados: a) el supuesto rol de experto no capacita a los sujetos para realizar mejores simulaciones, y b) las enfermedades mentales a simular en relación con la no imputabilidad son accesibles también a legos. Una observación de las medias en la Tabla 4 indica claramente que hay una capacidad por ambas partes para simular adecuadamente. Ahora bien, este resultado no se puede generalizar a otros contextos en los que se requieran conocimientos más específicos tal y como encontramos en el caso de la simulación de trastorno de estrés postraumático donde sólo encontramos, entre una muestra de 105 sujetos, que dos expertos sujetos eran capaces de simular este trastorno (Arce, Fariña y Freire, en prensa).

TABLA 4. PRUEBAS DE LOS EFECTOS INTER-SUJETOS

Variable	MC	F	p	η^2	\bar{x}_{legos}	$\bar{x}_{experto}$
Hipocondría	540.800	1.572	.214	.020	67.625	62.425
Depresión	480.200	2.146	.147	.027	67.250	62.350
Histeria	80.000	.437	.511	.006	60.325	58.325
Desviación psicopática	241.513	1.526	.220	.019	73.825	77.350
Masculinidad-feminidad	192.200	1.233	.270	.016	55.575	68.675
Paranoia	120.050	.577	.450	.007	80.000	82.450
Psicastenia	103.513	.764	.385	.010	64.375	62.100
Esquizofrenia	120.050	.768	.384	.010	79.550	82.000
Hipomanía	577.813	3.288	.074	.040	62.225	67.600
Introversión social	599.513	3.857	.053	.047	69.450	63.975

g.l. 1,78).

Estudio de la simulación en las medidas de control del MMPI

Estudio del cambio en las escalas de validez del MMPI

Las escalas de control de la validez de las respuestas son sensibles a la manipulación del factor simulación (instrucciones estándar vs. instrucciones de simulación), $F_{\text{multivariada}}(4,155) = 175.4$; $p^{***} < .001$; T.E. = .819. Se valida, en general, la hipótesis 4.

TABLA 5. PRUEBAS DE LOS EFECTOS INTER-SUJETOS

Variable	MC	F	p	η^2	\bar{x}_{normal}	$\bar{x}_{simulado}$
Escala de interrogantes (?)	11.556	.602	.439	.004	46.238	45.7
Escala L (mentira)	93.025	1.038	.310	.007	55.438	53.913
Escala F (incoherencia)	69764.256	623.990	.000	.798	48.412	90.175
Escala K (corrector)	10160.156	93.918	.000	.373	54.288	38.35

g.l. (1,158).

Los efectos univariados (consúltense en la Tabla 5) ponen de manifiesto que se produce un aumento significativo en los valores de la escala F (incoherencia) y un descenso en la escala K. Por su parte, la escala de interrogantes (CS) así como la escala L (mentira) no son sensibles a la simulación. Las puntuaciones de la escala de interrogantes dan ambas respuestas, bajo instrucciones estándar y de simulación, como válidas. Asimismo, los valores promedio en la escala L (mentira) están dentro de un perfil válido tanto en las respuestas en instrucciones estándar ($\bar{x} = 55.4$) como de simulación ($\bar{x} = 53.9$). Estos resultados dan base a la utilidad de las escalas de validez, específicamente las escalas F y K, en este contexto forense. No obstante, el criterio de las escalas de validez del MMPI no es suficiente por sí mismo ya que los «verdaderos enfermos» pueden presentar en ellas puntuaciones invalidantes, de modo que tras la detección de «presunta invalidez» es necesario estudiar las hipótesis alternativas (véanse, por ejemplo, en Zabalegui, 1990; Graham, 1992; Roig-Fusté, 1993). Veremos a continuación, pormenorizadamente y por separado, cada uno de los índices así como sus combinaciones a fin de establecer su «verdadero» valor en simulaciones con $n = 1$ en el ámbito forense.

Estudio de las escalas de validez por separado

• *Escala de interrogantes (CS).*

La no respuesta de los sujetos puede mediar el resultado de las escalas clínicas con lo que un número elevado de no respuestas (se considerarían no válidos aquellos cuestionarios cuyas puntuaciones en la escala de interrogantes superara 30), anularía el protocolo. La simple observación de la Tabla 6 nos informa de que esta contingencia no se da en ningún caso. En consecuencia, los sujetos con intento de simulación descartan como estrategia la no respuesta probándose así la hipótesis número 5. Por el contrario, se confirman las críticas al modelo criminológico apuntadas por Rogers (1992), Lewis y Saarni, (1993) o Bagby y otros (1997) que pronosticaban que los simuladores cooperarían en la evaluación. En resumen, el simulador utiliza una estrategia racional de simulación en la que presupone que debe demostrar la enfermedad y ésta no se demostraría no apareciendo como enfermo en las medidas.

TABLA 6. VALORES DIRECTOS DE NO RESPUESTA (VALORES T). ESCALA DE INTERROGANTES

Valores directos (valores T)	Frecuencia	Porcentaje	Interpretación ¹
0 (46)	66	82.5	Válido
1 (47)	5	6.3	Válido
2 (49)	3	3.8	Probable válido
3 (51)	1	1.3	Probable válido
6 (54)	2	2.5	Probable válido
7 (54)	1	1.3	Probable válido
8 (55)	2	2.5	Probable válido

1. Estas interpretaciones cualitativas están basadas en las puntuaciones del MMPI-2 con lo que podrían existir ligeras diferencias con el MMPI original. En todo caso son orientativas.

• *Escala L (mentira).*

La escala L (mentira) se orienta a la evaluación del grado en que el sujeto intenta falsificar sus respuestas en el sentido de manipular su imagen hacia la deseabilidad social ($T < 70$). Tal y como cabría esperar ($\bar{x} = 43.91$; $s_x = 11.1$), los simuladores no tienden a presentar una imagen de deseabilidad social, $t(79) = -21.014$; $p^{***} < .001$. En el estudio de casos hallamos que únicamente dos sujetos parecen estar movidos a presentar una imagen positiva, esto es, rechazo de faltas (mírese la Tabla 7). En consecuencia, esta escala, por sí sola, no es de un valor relevante a la hora de descubrir simulación con lo que se verifica la hipótesis 6.

TABLA 7. VALORES T EN LAS RESPUESTAS DE SIMULACIÓN. ESCALA L

Valores T	Frecuencia	Porcentaje	Interpretación
30	6	7.5	Posible actitud de «mala imagen» ²
34	11	13.8	Posible actitud de «mala imagen»
37	15	18.8	Posible actitud de «mala imagen»
41	11	13.8	Posible actitud de «mala imagen»
45	7	8.8	Posible actitud de «mala imagen»
48	6	7.5	Posible actitud de «mala imagen»
52	8	10	Válido
56	7	8.8	Válido
60	6	7.5	Probablemente válido
67	1	1.3	Probablemente válido
71	2	2.5	Validez cuestionable

2. La mala imagen no se puede inferir sólo de esta escala, es un indicador de potencia cuando se une a puntuaciones K bajas. Por sí mismas, puntuaciones bajas son indicio de veracidad.

• *Escala F (incoherencia).*

Esta escala se orienta a la medición de respuestas inusuales en personas normales (contenidos insólitos, experiencias raras, temas religiosos, reacciones impulsivas o paranoides, etc.). En principio, sólo el 10% de las respuestas del grupo normativo ponderan en la dirección del test. Una puntuación T mayor de 80 es un indicio de intentar presentar una imagen negativa (Roig-Fusté, 1993). En general ($\bar{x} = 90.18$; $s_x = 11.62$), este índice es efectivo, $t(79) = 7.831$; $p^{***} < .001$, y, además, clasifica correctamente al 88.7% de los simuladores. Encontramos, de este modo, apoyo a nuestra hipótesis experimental número 7. Aun así, nuestros sujetos pasarían en un 11.3% este control de validez. Por otra parte, este índice también es orientativo en relación a sujetos que padecen realmente trastornos psicóticos cuyos valores suelen darse entre 70 y 80 (Roig Fusté, 1993). En otras palabras, la potencia, en nuestro contexto de medida, de la hipótesis alternativa de que este tipo de puntuaciones son propias de sujetos realmente psicóticos reduce significativamente su valor.

• *Escala K (corrector).*

El factor K de corrección de determinadas escalas clínicas, también tiene un poder discriminativo de la actitud del sujeto en la evaluación. Sujetos con valores altos ($T > 70$) probablemente estén fingiendo una buena imagen de sí mismos, en tanto valores bajos ($T < 50$) advierten de la posible simulación de un deterioro (véase la Tabla 8). En términos generales, el factor K detectaría correctamente la simulación. De hecho, la puntuación media ($\bar{x} = 38.66$; $s_x = 9.06$) en esta escala es significativamente inferior, [$t(79) = -10.671$; $p^{***} < .001$], a 50; en tanto el estudio de casos pone de manifiesto que el 87.5% de los sujetos son detectables por esta escala, o sea, tienen puntuaciones $T < 50$. Por el contrario, el 12.5% de la población superaría este sistema de control de la validez. De estos resultados se desprende que éste es un indicador robusto de simulación, $\chi^2_1 = 45$; $p^{***} < .001$, ratificándose la hipótesis 8. Nuevamente, este indicador señala que los sujetos no tienden a presentar una buena imagen de sí mismos ($T > 70$) al no encontrarse ningún caso bajo esta contingencia.

TABLA 8. VALORES T EN LAS RESPUESTAS DE SIMULACIÓN, ESCALA K

Valores T	Frecuencia	Porcentaje	Interpretación
23	1	1.3	Perfil invalidado
25	4	5.0	Perfil invalidado
28	3	3.8	Perfil invalidado
30	10	12.5	Perfil invalidado
32	9	11.3	Perfil invalidado
35	7	8.8	Perfil invalidado
37	8	10.0	Perfil invalidado
39	7	8.8	Perfil invalidado
41	9	11.3	Perfil invalidado
44	2	2.5	Perfil invalidado
46	6	7.5	Perfil invalidado
48	4	5.0	Perfil invalidado
51	2	2.5	Perfil válido
53	2	2.5	Perfil válido
55	1	1.3	Perfil válido
58	3	3.8	Perfil válido
60	1	1.3	Validez cuestionada
62	1	1.3	Validez cuestionada

Configuraciones de las escalas de validez

El índice F-K, también conocido como «índice de Gough», se manifiesta, en general, como un indicador potente de los intentos de distorsión de los resultados en el sentido de intentar dar una mala imagen de sí mismo. De hecho, nuestros datos evidencian, en el conjunto de la muestra, una puntuación media superior y significativa [$\bar{x} = 51.825$; $s_x = 16.5359$; $t(79) = 11.805$; $p^{***} < .001$] en comparación con el punto de corte establecido para determinar un intento de simulación ($T \geq 30$) (Rogers, 1997; Zabalegui, 1990). Además, este índice clasifi-

caría correctamente al 90% de los simuladores, ajuste que le convierte en fiable, $\chi^2_1 = 51.2$; $p^{***} < .001$.

El perfil en «V invertida», con L y K por debajo de 50 y F sobre 80, sugiere la posibilidad de exageración de síntomas. 48 sujetos (un 60%) serían detectados por este sistema como personas que exageran síntomas para aparecer enfermos, lo que no le confiere a este índice un estatus de fiable, $\chi^2_1 = 3.2$; y, por tanto, limita claramente la potencia de su uso en este contexto de evaluación. Sucintamente, estos resultados validan las expectativas de la hipótesis número 9 para el «índice F-K», en tanto que desestiman el valor del perfil en «V invertida».

Estudio de la simulación en la entrevista

Del total de sujetos, 35 (un 43.8%) no son capaces de simular trastorno mental alguno, en tanto 45, que representan el 56.2%, pueden simular al menos un trastorno mental concreto (véanse los resultados directos por trastorno específico en la Tabla 9). De estos datos directos se deriva que no existen diferencias entre simuladores de enfermedad mental efectivos (es decir, que simulan alguna enfermedad mental) y no efectivos (es decir, que no logran simular enfermedad mental alguna), $\chi^2_1 = 1.125$.

TABLA 9. CONTINGENCIA DE LOS TRASTORNOS SIMULADOS EN LA ENTREVISTA

Trastorno	Frecuencia	Porcentaje
Ningún trastorno	35	43.8
Trastorno delirante de grandiosidad***	3	3.8
Trastorno delirante celotípico***	1	1.3
Trastorno delirante persecutorio***	4	5.0
Esquizofrenia paranoide***	13	16.3
Esquizofrenia indiferenciada***	4	5.0
Depresión con síntomas psicóticos***	1	1.3
Trastorno psicótico no especificado***	4	5.0
Trastorno antisocial de la personalidad	2	2.5
Trastorno esquizoide de personalidad	4	5.0
Trastorno paranoide de personalidad	1	1.3
Trastorno esquizotípico de personalidad	1	1.3
Trastorno narcisista de personalidad	1	1.3
Trastorno de personalidad por evitación	1	1.3
Trastorno distímico	1	1.3
Trastorno explosivo intermitente	1	1.3
Esquizofrenia paranoide + trastorno antisocial de la personalidad***	1	1.3
Trastorno distímico con síntomas psicóticos***	1	1.3
Esquizofrenia paranoide + Trast. esquizoide de personalidad***	1	1.3

*** Simulación efectiva de no imputabilidad.

Por su parte, a los 35 sujetos que no habían simulado trastorno alguno en la entrevista habría que añadir otros 12 que simulan un trastorno que los convertiría en imputables, resultando en un total de 47 sujetos imputables (58.8% del

total). De los resultados directos inferimos que este indicador clasifica por un igual, o sea, en torno a la mitad de la población, correctamente. En esta línea, no constatamos diferencias entre número de simulaciones correctas y simulaciones detectadas, $\chi^2_1 = 2.667$. En suma, por sí mismas las entrevistas no son criterios fiables a la hora de establecer fidedignamente un criterio de imputabilidad, requiriéndose otros métodos complementarios. Sin embargo, se corrobora la hipótesis 10, esto es, los sujetos no tienen accesible cognitivamente (en terminología legal, «no tienen un camino conducente») la simulación de enfermedad mental a través de la entrevista en formato de discurso libre, al contrario de lo que ocurría con las escalas clínicas del MMPI.

Estudio de casos de la capacidad de simulación de no imputabilidad y consistencia interna de los instrumentos de medida, MMPI y entrevista

En las escalas clínicas del MMPI, 63 (78.8%) de los sujetos pueden simular efectivamente no imputabilidad (contémplese en la Tabla 10), esto es, la probabilidad de simulación de no imputabilidad es significativa, $\chi^2_1 = 26.45$; $p^{***} < .001$. Así, este tipo de instrumentación psicométrica no puede tomarse como determinante sin considerar el control de las escalas de validez, lo que avala, de nuevo, la hipótesis 2.

TABLA 10. CONTINGENCIAS DE IMPUTABILIDAD POR FORMATO DE EVALUACIÓN

<i>Imputabilidad</i>	<i>MMPI</i>	<i>MMPI + Escalas de validez</i>	<i>Entrevista</i>	<i>Entrevista + estudio de simulación</i>
Imputable	17 (21.25%)	74 (92.5%)	47 (58.75%)	57 (71.25%)
No imputable	63 (78.75%)	6 (7.5%)	33 (41.25%)	23 (28.75%)

Nota: El estudio de la simulación incluye la detección de las estrategias de simulación: síntomas raros, combinación de síntomas, síntomas obvios, consistencia de síntomas, síntomas improbables, y severidad de síntomas.

Uno de los modos de detección del engaño es la consistencia interna de los discursos. Este procedimiento cuenta con cobertura tanto desde el punto de vista de la evaluación empírica de la simulación (Rogers y Mitchell, 1991) como desde el ordenamiento jurídico (Hastie *et al.*, 1983; Wigmore, 1937) para conocer la fiabilidad de la información. Así, en ambos contextos, la información inconsistente internamente se estima como no creíble. En el campo de estudio de la simulación, la consistencia interna del discurso se puede operativizar a través del estudio de las seis estrategias habituales de simulación: síntomas raros, combinación de síntomas, síntomas obvios, consistencia de síntomas, síntomas improbables, y severidad de síntomas (véase Rogers y Mitchell, 1991 para una definición y ejemplos de cada categoría). Sometidas las entrevistas clínicas a este análisis encontramos que entre los simuladores de enfermedad mental de tipo no imputable en la entrevista, 10 fueron detectados como simuladores ya que recu-

rrían a una o más de las seis estrategias de simulación. Así, encontramos apoyo a la hipótesis número 11 que sostenía la efectividad de la conjunción de estos dos procedimientos de medida (entrevista no directa más análisis de las estrategias de simulación) en la detección de la simulación (véase la Tabla 10) que se eleva al 71.25% (57 sujetos), $\chi^2_1 = 14.5$; $p^{***} < .001$. No obstante, no clasifica correctamente al 28.75%.

Si bien las escalas clínicas del MMPI resultan más fáciles para manipular la imagen de imputabilidad que la entrevista clínica en formato de discurso libre, $\chi^2_1 = 23.438$; $p^{***} < .001$; $\phi = -.383$, la unión de las escalas clínicas y las de validez (y sus configuraciones), medida de la consistencia interna de esta evaluación, es tan fiable como la entrevista más el estudio de las estrategias de simulación en la misma, $\chi^2_1 = 2.206$ (repárese en la Tabla 10).

Estudio de la consistencia inter-medidas: MMPI y entrevista

Los resultados del acuerdo inter-medidas (obsérvense en la Tabla 11) indican claramente que los sujetos tienen un grado de consistencia (>25%) elevado en paranoia, esquizofrenia e introversión social. Tanto la esquizofrenia como la paranoia podrían estar asociadas a no imputabilidad. No obstante, la ausencia de concordancia se ha mostrado como un criterio robusto ya que detectaría en torno a un 70.1%, $\chi^2_1 = 10.881$; $p^{***} < .001$, de los falsos paranoicos y a un 64.3%, $\chi^2_1 = 5.714$; $p^* < .05$, 1 de los falsos esquizofrénicos. Además, considerada la inconsistencia en cualquiera de las escalas como criterio, la falta de concordancia inter-medidas se elevaría al 97.5% (sólo dos sujetos son totalmente consistentes inter-medidas). En resumen, la consistencia inter-medidas es un estimador fiable para detectar simuladores, pero evidencia ciertas limitaciones tales como el error de medida (sobre todo en el análisis de contenido de los protocolos, especialmente en estudios de casos), limitaciones del sujeto en las respuestas (más específicamente, los sujetos pueden tener dificultades o verse imposibilitados a expresar síntomas/criterios diagnósticos en formato de discurso libre); o que la

TABLA 11. CONTINGENCIA DE ACUERDO INTER-MEDIDAS

<i>Variable</i>	<i>Consistencia</i>	<i>Inconsistencia</i>
Hipocondría		35 (100%)
Depresión	2 (6.3%)	30 (93.8%)
Histeria		20 (100%)
Desviación psicopática	3 (4.9%)	58 (95.2%)
Masculinidad feminidad		12 (100%)
Paranoia	20 (29.9%)	47 (70.1%)
Psicastenia		27 (100%)
Esquizofrenia	25 (35.7%)	45 (64.3%)
Hipomanía		37 (100%)
Introversión social	32 (54.2%)	27 (45.8%)

inconsistencia sea fruto de la enfermedad mental en sí. Todo ello relativiza el alcance de este método aisladamente. En todo caso, la combinación de dos tareas, conocimiento y reconocimiento, es un indicador robusto de simulación.

Combinación de índices

Tres son los índices de control que nosotros hemos considerado en este estudio para detectar la simulación: *a)* simular enfermedad no imputable; *b)* superar las escalas e índices de control del MMPI, y un estudio de las estrategias de simulación en la entrevista; y *c)* la consistencia inter-medidas. Primero, para que un individuo sea considerado imputable se requiere que simule una enfermedad mental no imputable; en otras palabras, el sujeto que no puede simular tal enfermedad no pasa el primer punto de corte para ser declarado no imputable. Segundo, las medidas deben superar las escalas de validez del MMPI y no detectarse ninguna de las seis estrategias descritas de simulación en la entrevista. De no ser así, estamos ante un claro indicio de falta de fiabilidad de los datos. Tercero, la evaluación ha de ser consistente inter-medidas.

Pues bien, la combinación de estos índices, si partimos de las escalas de validez del MMPI, pone al descubierto que seis sujetos superaron las escalas de control del MMPI, sin embargo tres resultaron simuladores imputables y, además, dos de ellos inconsistentes inter-medidas. Los tres restantes, que simulan enfermedad mental incapacitante, no se muestran consistentes inter-medidas. Por su parte, si partiéramos de la consistencia inter-medidas nos encontraríamos con sólo dos sujetos totalmente consistentes, aunque uno de ellos imputable. El otro, un sujeto experto que simula esquizofrenia, paranoia e introversión social, y en cuyo protocolo de entrevista no se observa ninguna estrategia de simulación, es detectado por diversos índices de las escalas de control del MMPI (F, K y F-K). Aun así, para este sujeto sería imprescindible recabar otro tipo de información convergente para proceder a su invalidación. Comprobadas todas las combinaciones disponibles de detección de la simulación encontramos que, a excepción del mentado sujeto, al menos dos sistemas de control son efectivos. En resumen, es necesario (y posible a la vez que efectivo) sostenerse, cuando menos, en dos procedimientos de invalidación del protocolo concreto, esto es, es preciso proceder con lo que llamamos una «invalidación convergente».

Discusión

La primera aproximación que merece esta discusión tiene por objeto el verdadero alcance de los resultados aquí mostrados. Primero, la implicación y subsecuente ejecución de los sujetos no es exactamente igual en este contexto de laboratorio que en la realidad. No obstante, las diferencias fundamentales pueden circunscribirse al nivel de la ansiedad situacional. Ésta presenta dos versiones, una facilitadora de la ejecución (p.ej., motivadora del aprendizaje) y otra in-

hibidora de la ejecución (dificulta la consistencia inter-medidas, tiempo, etc.) con efectos inversos en la realidad. En nuestro contexto, los sujetos evidencian, en menor cuantía pero igualmente proporcional, ambos tipos de ansiedad. Segundo, los métodos de detección de la simulación son parciales por lo que no pueden tomarse como determinantes. Tercero, cabe la posibilidad de que los sujetos realmente enfermos pudieran ser detectados por estos métodos como simuladores. Hechas estas salvedades, los resultados aquí obtenidos pueden clasificarse en dos bloques: criterios positivos y negativos:

a) *Criterios positivos*: la no evitación de respuestas y la deseabilidad social. Aquellos sujetos que sean evaluados por las escalas de control del MMPI en el sentido de que se abstienen significativamente de dar respuesta (Escala CS) y que tienden a dar respuestas de deseabilidad social, no siguen las estrategias típicas de simulación, lo que debe interpretarse más como indicio de veracidad del protocolo que como un intento de simulación. Es preciso tener en mente que este criterio se ha interpretado en términos forenses como un indicador fiable de simulación (Rogers, 1992; Lewis y Saarni, 1993; Bagby *et al.*, 1997). No obstante, esta relación no se puede generalizar a otros contextos legales tales como la evaluación de daños. Así, a menor tiempo transcurrido entre el daño y la evaluación, mayor probabilidad de que se produzca un perfil inválido (debido a un exceso de omisiones) (Kane, 1999). Es decir, la evaluación requiere, en este contexto de evaluación de daño, que éste esté consolidado. Segundo, Samuel y otros (1994) encontraron que personas implicadas en litigios por daño personal dejaban más preguntas sin contestar que otras personas. Además, si se usa el MMPI-2 es preciso tener en mente que el número de preguntas que los individuos dejan sin contestar es menor, debido a que, en esta nueva versión, se eliminaron cinco de los diez ítems del MMPI que eran omitidos con más frecuencia.

b) *Criterios negativos*. Nuestros resultados ponen de manifiesto que existen una serie de criterios negativos, esto es, que anulan o mitigan la validez del protocolo. Éstos son: 1) los sistemas de medición, MMPI, entrevista u otros, no detectan, en protocolos válidos, enfermedad mental (en otras palabras, si los instrumentos de medida no detectan enfermedad incapacitante alguna no se puede sostener la misma); 2) las escalas de control de validez y sus combinaciones detectan simulación; 3) la detección de alguna estrategia de simulación en la entrevista; y 4) ausencia de concordancia inter-medidas. El primer criterio es eliminatorio, o sea, si la enfermedad mental no es medible, no se puede sostener no imputabilidad por causa de enajenación mental (téngase presente que hemos excluido la oligofrenia o retraso mental). Los restantes, en sí mismos, no son determinantes, con lo que se requerirá, cuando menos, la complementación de dos criterios y el estudio de las hipótesis alternativas para concluir en relación a la no imputabilidad. Para estos últimos indicios de no validez es para los que formulamos el concepto de «invalidez convergente» que requiere de, al menos, dos indicios de invalidez para desestimar un protocolo como válido.

Todo ello nos lleva a formular el siguiente protocolo de actuación:

a) Recurso a sistemas de medida complementarios y concordantes, no fácilmente simulables (MMPI, Millon, entrevista, entre otros posibles). Es recomendable recurrir a dos sistemas de medida que impliquen tareas distintas tal

como reconocimiento (p.e., MMPI) y de conocimiento (p.e., entrevista no directiva). De acudir a instrumentos psicométricos concordantes debe tenerse presente que la tasa de consistencia no va a ser total. No se nos debe escapar que incluso las medidas test-retest distan mucho de la perfección.

b) Análisis de la consistencia interna de las medidas (tales como escalas de control, consistencia temporal en la entrevista, o consistencia interna de la entrevista en relación con el análisis de las seis estrategias que la literatura ha descrito que siguen los simuladores: síntomas raros, combinación de síntomas, síntomas obvios, consistencia de síntomas, síntomas improbables, y severidad de síntomas).

c) Es aconsejable que dos evaluadores, por separado, lleven a cabo la evaluación, con lo que podremos comprobar la consistencia inter-evaluador. Este sistema nos sirve para controlar posibles sesgos de medida e interpretación en el evaluador.

d) Estudio de la fiabilidad: consistencia interna, inter-medidas, inter-contextos (antecedentes, pruebas documentales, etc.), e inter-evaluadores (Wicker, 1975).

e) Control de falsos positivos, esto es, enfermos reales, a través de un estudio de los antecedentes e historia general del sujeto, de las hipótesis alternativas en cada indicador de no validez (véase Roig Fusté, 1993, Graham, 1992), y del cumplimiento de los criterios del Modelo de Decisión Clínica para el establecimiento de la simulación (Cunnién, 1997).

f) Anamnesis o estudio de los antecedentes. Afortunadamente, en la realidad contamos con más información acerca del sujeto lo cual facilita, si no determina, la evaluación. Así, podemos reforzar nuestra evaluación con los antecedentes del sujeto, datos del entorno, un estudio de su comportamiento, el recabamiento de pruebas documentales, otros testimonios, etc.

g) Estudio psicológico del grado de responsabilidad. Las medidas objetivas proporcionan datos respecto a la información biológica, pero ésta necesita de una complementación psicológica que clarifique la relación de la enajenación mental con la responsabilidad penal. Recordemos que ésta se refleja en las capacidades cognitivas o volitivas del sujeto que deben estar anuladas para ser declarado no imputable, o deterioradas para reducir su grado de responsabilidad. Todo ello referido al momento de los hechos.

h) Finalmente, también se puede poner a prueba la validez discriminante. En otras palabras, realizar una medida no relacionada con el caso, tal como de valores o personalidad (16-PF, S.I.V., o S.P.V.), esperando que no presente relación alguna con la evaluación objetivo y se descarte un intento de manipulación de la imagen, tanto en positivo como en negativo.

El sistema de evaluación debería ceñirse a las siguientes categorías: «probablemente simulador»; «probablemente no simulador». Es importante dejar a un lado la intención de establecer la certeza, y utilizar estos términos probabilísticos, puesto que nuestros sistemas de evaluación, tal y como pudo verse en el presente estudio, están sujetos a error. Sin embargo, el TS exige la seguridad plena, no la alta probabilidad, pero como toda medida está sujeta a error, debemos reconocerlo, aunque absteniéndonos de establecer el grado de certeza (p.e.,

sentencia del TS de 29 de octubre de 1981, RA 3902). Finalmente, debe tenerse en mente la máxima «in dubio pro reo», también aplicable al estudio de la imputabilidad, esto es, en caso de duda razonable decantarse a favor del acusado. Sería recomendable para estos casos recurrir a una tercera categoría «indefinido», ya que posicionarse por “probablemente no simulador» supone la asunción de un riesgo más allá del margen de error científico.

REFERENCIAS

- Aldea, M. J. (1994). Revisión del estado actual de los conocimientos sobre la neurosis de renta. *Informaciones Psiquiátricas*, 138, 411-436.
- American Psychiatric Association (1995a). *DSM-IV. Manual diagnóstico y estadístico de los trastornos mentales*. Barcelona: Masson.
- American Psychiatric Association (1995b). *DSM-IV Breviario. Criterios diagnósticos*. Barcelona: Masson.
- Arce, R., Fariña, F. y Freire, M. J. (en prensa). Contrastando la generalización de los métodos empíricos de detección del engaño. *Psicología: Teoría, Investigación e Práctica*.
- Bagby, R. M., Buis, T. & Nicholson, R. A. (1995). Relative effectiveness of the standard validity scales in detecting fake-bad and fake-good responding: replication and extension. *Psychological Assessment*, 7, 84-92.
- Bagby, R. M., Rogers, R., Buis, T., Nicholson, R. A., Cameron, S. L., Rector, N. A., Schuller, D. R. & Seeman, M. V. (1997). Detecting feigned depression and schizophrenia on the MMPI-2. *Journal of Personality Assessment*, 68 (3), 650-664.
- Baer, R.A., Wetter, M.W. & Berry, D.T.R. (1995). Effects of information about validity scales on underreporting of symptoms on the MMPI-2: an analogue investigation. *Assessment*, 2 (2), 189-200.
- Ben-Porath, Y.S. & Butcher, J.N. (1989). The psychometric stability of rewritten MMPI items. *Journal of Personality Assessment*, 53, 645-663.
- Carbonell, J. C., Gómez, J. L. y Mengual, J. B. (1987). *Enfermedad mental y delito. Aspectos psiquiátricos, penales y procesales*. Madrid: Cívitas.
- Cunnien, A. J. (1997). Psychiatric and medical syndromes associated with deception. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 23-46). New York: Guilford Press.
- Duckworth, J.C. & Anderson, W.P. (1995). *MMPI and MMPI-2: interpretation manual for counselors and clinicians*. Bristol: Accelerated Development.
- Fisher, R. P. & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing*. Springfield: Charles C. Thomas.
- Freire, M. J. (2000). *Evaluación de la credibilidad de testigos identificadores y simulación de victimación*. Tesis de Licenciatura. Universidad de Santiago de Compostela.
- Garrido, E. (1994). Psychology and law in Spain. *Applied Psychology: An International Review*, 43 (2), 303-311.
- Geiselman, R. E., Fisher, R. P., Firstenberg, I., Hutton, L. A., Sullivan, S., Avetissian, I. & Prosk, A. (1984). Enhancement of eye-witness memory: an empirical evaluation of the cognitive interview. *Journal of Police Science and Administration*, 12, 74-80.
- Graham, J.R. (1990). *MMPI-2: assessing personality and psychopathology*. New York: Oxford University Press.
- Graham, J.R. (1992). *Interpretation of MMPI-2. Validity and clinical scales*. Brujas (workshop).
- Hastie, R., Penrod, S.D. & Pennington, N. (1983). *Inside the jury*. Cambridge, Mass.: Harvard University Press (trad. cast.: *La institución del jurado en Estados Unidos. Sus intimidades*. Madrid: Cívitas, 1986).
- Ibáñez, V. y Luis P. de (1994). Psicología forense aplicada en España. En J. Sobral, R. Arce y A. Prieto (Eds.), *Manual de psicología jurídica* (pp.213-225). Barcelona: Paidós.
- Jólluskin, G. S. (2000). *Influencia social mayoritaria, minoritaria y decisiones de jurados*. Tesis Doctoral. Universidad de Santiago de Compostela.
- Kane, A. W. (1999). Essentials of malingering assessment. En M.J. Ackerman (Ed.), *Essentials of forensic psychological assessment*. New York: John Wiley and Sons.
- Kury, H. (1997). Law and psychology in Europe: current status and future perspectives. En S. Redondo, V. Garrido, J. Pérez & R. Barberet (Eds.), *Advances in psychology and law. International contributions* (pp.3-25). Berlín: Walter de Gruyter.
- Lewis, M. & Saarni, C. (1993). *Lying and deception in everyday life*. New York: Guilford Press.

- Lyons, J. A. & Wheeler, C. T. (1999). MMPI, MMPI-2 and PTSD: Overview of scores, scales and profiles. *Journal of Traumatic Stress, 12* (1), 175-183.
- Nicholson, R. A., Mouton, G. J., Bagby, R. M., Buis, T., Peterson, S. A. & Buigas, R. A. (1997). Utility of MMPI-2 indicators of response distortion: receiver operating characteristic analysis. *Psychological Assessment, 9* (4), 471-479.
- Pope, K. S., Butcher, J. N. & Seelen, J. (2000). *The MMPI, MMPI-2 & MMPI-A in court: a practical guide for expert witnesses and attorneys*. Washington, DC: American Psychological Association.
- Rogers, R. (1986). *Conducting insanity evaluations*. New York: Van Nostrand Reinhold.
- Rogers, R. (1992). *Structured interview of reported symptoms*. Odessa, FL: Psychological Assessment Resources.
- Rogers, R. (1997). *Clinical assessment of malingering and deception*. New York: Guilford Press.
- Rogers, R., Kropp, P. R., Bagby, R. M. & Dickens, S. E. (1992). *Structured interview of reported symptoms (SIRS) and professional manual*. Odessa, FL: Psychological Assessment Resources.
- Roig-Fusté, J. M. (1993). *MMPI y MMPI-2 en la exploración de la personalidad*. Barcelona: Agil Offset.
- Samuel, S. E., DeGiroloano, J., Michals, T. J. & O'Brien, J. (1994). Preliminary findings on MMPI «Cannot Say» responses with personal injury litigants. *American Journal of Forensic Psychology, 12* (4), 5-18.
- Shuman, D. W. (1994). *Psychiatric and psychological evidence*. Deerfield, IL: Clark Boardman Callaghan.
- TEA (1988). *MMPI, suplemento técnico*. Madrid: Tea Ediciones.
- Torre, J. de la (1999). La responsabilidad penal en las psicopatologías: valoraciones judiciales y jurisprudenciales. *Estudios de Psicología, 63-64*, 163-173.
- Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.
- Weick, K. E. (1985). Systematic observational methods. En G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (vol. 1) (pp. 567-634). Hillsdale, NJ: LEA.
- Wicker, A. W. (1975). An application of a multiple-trait-multimethod logic to the reliability of observational records. *Personality and Social Psychology Bulletin, 4*, 575-579.
- Wigmore, J.H. (1937). *The science of judicial proof*. Boston: Little, Brown.
- Zabalegui, L. (1990). *La medida de la personalidad por cuestionarios*. Madrid: Universidad Pontificia de Comillas.