

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE PSICOLOGÍA
DEPARTAMENTO DE METODOLOGÍA DE LAS CIENCIAS DEL COMPORTAMIENTO



TESIS DOCTORAL

**Utilidad diagnóstica del inventario estructurado de simulación de síntomas
(SIMS) en población española**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Pablo Santamaría Fernández

Directores

Rosario Martínez Arias
Héctor González Ordi

Madrid, 2014

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE PSICOLOGÍA

Departamento de Metodología de las Ciencias del Comportamiento



TESIS DOCTORAL

**UTILIDAD DIAGNÓSTICA DEL INVENTARIO ESTRUCTURADO DE
SIMULACIÓN DE SÍNTOMAS (SIMS) EN POBLACIÓN ESPAÑOLA**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Pablo Santamaría Fernández

Bajo la dirección de los doctores

Rosario Martínez Arias

Héctor González Ordi

Madrid, 2014

“La ausencia de evidencia no es lo mismo que la evidencia de ausencia”

Teuber (1969, p.20)

“No puedo daros más que gracias y más gracias y siempre gracias.”

W. Shakespeare, Noche de Reyes, Acto III, Escena III

*A todos los que de una u otra manera habéis hecho posible esta tesis, con vuestro tiempo,
con vuestro ánimo o con vuestras opiniones, correcciones y sugerencias.*

Sin vosotros, hubiera sido sencillamente imposible.

Índice

Índice de tablas	VIII
Índice de figuras	XIV
Introducción.....	1
Marco teórico.....	5
1. Concepto y delimitación de la simulación	7
1.1. Definición y modelos explicativos de la simulación	7
1.1.1. El concepto de simulación en el DSM	7
1.1.2. El modelo adaptativo	12
1.2. Ámbitos de simulación	17
1.3. Prevalencia	18
1.4. Estrategias de detección de simulación	22
1.5. Criterios diagnósticos de simulación.....	29
1.5.1. Criterios diagnósticos de Slick et al. (1999).....	30
1.5.2. Criterios diagnósticos de Bianchini et al. (2005)	34
1.5.3. Limitaciones y mejoras sugeridas a los criterios diagnósticos de simulación	36
2. Metodología de investigación en simulación.....	43
2.1. Diseños de investigación habitualmente empleados en estudios de simulación	43
2.1.1. Análogos	48
2.1.2. Grupos diagnósticos conocidos.....	56
2.1.3. Diseños de medidas autosuficientes (<i>bootstrapping comparisons</i>)	57
2.1.4. Prevalencia diferencial	58
2.1.5. Grupos mixtos	59
2.2. Análisis estadísticos empleados habitualmente en estudios de simulación.....	59
2.2.1. Análisis de las diferencias entre grupos	60
2.2.2. Análisis de la precisión clasificatoria	60
3. Instrumentos para la detección de la simulación: el papel de los autoinformes y del Inventario Estructurado de Simulación de Síntomas, SIMS	65
3.1. El papel de los autoinformes en la detección de simulación	67
3.1.1. La evaluación de la exageración de síntomas: puntos de corte y contexto.....	68
3.1.2. Factores que afectan a la sensibilidad de las escalas de exageración de síntomas.....	69
3.1.3. Exageración de síntomas y simulación.....	71
3.2. Principales autoinformes empleados en la detección de exageración de síntomas	73

3.2.1. MMPI-2.....	74
3.2.2. MMPI-2-RF.....	75
3.2.3. PAI.....	79
3.3. Principales autoinformes de screening en la detección de exageración de síntomas.....	81
3.4. SIMS, Inventario Estructurado de Simulación de Síntomas	82
3.5. Estudios de simulación análoga con el SIMS.....	85
3.5.1. Estudio de simulación análoga de Smith y Burger (1997).....	85
3.5.2. Estudio de simulación análoga de Rogers y Cruise (1998).....	88
3.5.3. Estudio de simulación análoga de Edens, Otto y Dwyer (1999).....	88
3.5.4. Estudio de simulación análoga de Theron et al. (2001)	90
3.5.5. Estudio de simulación análoga de Merckelbach y Smith (2003).....	91
3.5.6. Estudio de simulación análoga de Cima, Hollnack, et al. (2003).....	93
3.5.7. Estudio de simulación análoga de Rogers et al. (2005).....	94
3.5.8. Estudio de simulación análoga de retraso mental (Graue et al., 2007).....	95
3.5.9. Estudio de simulación análoga de trastorno por estrés postraumático de Efendov (2007).....	96
3.5.10. Estudio de simulación análoga de amnesia relacionada con la realización de un crimen (Giger, Merten, Merckelbach y Oswald, 2010)	98
3.5.11. Estudio de simulación análoga de TDAH (Buijs, 2011).....	98
3.5.12. Estudio de simulación análoga de Dandachi-FitzGerald y Merckelbach (2012)	100
3.5.13. Estudio de simulación análoga con adolescentes con diagnóstico dual en tratamiento psicológico (Rogers, Hinds y Sewell, 1996)	101
3.5.14. Estudio de simulación análoga con adolescentes con grupo clínico de referencia (Dearth, 2007).....	102
3.6. Estudios de simulación análoga y resistencia al entrenamiento con el SIMS.....	104
3.6.1. Estudio de detección de simulación de psicosis en simuladores análogos con diferente nivel de entrenamiento (Jelicic, Hessels y Merckelbach, 2006).....	104
3.6.2. Estudio de detección de simulación de psicosis en simuladores análogos con diferente nivel de conocimiento de psicopatología (Jelicic, Peters, Leckie y Merckelbach, 2007)	105
3.6.3. Estudio de detección de simulación de psicosis en simuladores análogos expertos en psicopatología (Jelicic, van Gaal y Peters, 2013).....	106
3.6.4. Estudio de detección de simulación de daño cerebral en simuladores análogos con diferente nivel de entrenamiento (Jelicic, Merckelbach, Candel y Geraerts, 2007).....	106
3.6.5. Estudio de detección de simulación de daño cerebral en simuladores análogos con diferente nivel de entrenamiento de Jelicic, Ceunen, Peters y Merckelbach (2011)	107
3.6.6. Estudio de detección de simulación de lesiones por esguince cervical en simuladores análogos con diferente nivel de entrenamiento (Merten, Diederich y Stevens, 2008).....	108

3.6.7. Estudio de detección de simulación de estrés postraumático en simuladores análogos con diferente nivel de entrenamiento (Merten, Lorenz y Schlatow, 2010)	109
3.7. Estudios de grupos diagnósticos o diseños autosuficientes en contexto forense penal con el SIMS.....	110
3.7.1. Estudio de simulación de psicopatología en presos de Poythress et al., (2001) y Edens et al. (2007).....	110
3.7.2. Estudio de detección de simulación de psicopatología en presos de nuevo ingreso (Laffoon, 2009)	112
3.7.3. Estudio de simulación de incapacidad para ser juzgados de Lewis, Simcox y Berry (2002)	112
3.7.4. Estudio de simulación de incapacidad para ser juzgados de Heinze y Purisch (2001).....	114
3.7.5. Estudio de simulación de incapacidad para ser juzgados de Vitacco, Rogers, Gabel y Munizza (2006)	116
3.8. Estudios de grupos diagnósticos conocidos o diseños autosuficientes en contexto forense civil con el SIMS	117
3.8.1. Estudio de simulación de lesiones para la obtención de indemnizaciones o compensaciones económicas de Alwes (2006).....	117
3.8.2. Estudio de simulación de síntomas depresivos para obtención de incapacidad de Clegg, Fremouw y Mogge (2009)	122
3.8.3. Estudio de simulación de lesiones para la obtención de indemnizaciones o compensaciones económicas de Wisdom et al. (2010)	125
3.8.4. Estudio con pacientes sometidos a evaluación en un centro de valoración (Merten et al., 2007).....	126
3.8.5. Estudio con pacientes con sintomatología somatoforme y de baja laboral (González-Ordi, Santamaría y Fernández Marín, 2010)	127
3.8.6. Estudio con pacientes diagnosticados de trastorno mixto ansioso-depresivo y adaptativo en el contexto medico-legal (Blasco Saiz y Pallardó Durá, 2013).....	128
3.9. Estudios en muestras de pacientes clínicos	130
3.9.1. Estudio con muestras de pacientes psiquiátricos (Dandachi-FitzGerald et al., 2011)	130
3.9.2. Estudio con muestras de pacientes con esquizofrenia (Peters, Jelacic, Moritz, Hauschildt y Jelinek, 2013).....	131
3.9.3. Estudio de prevalencia de exageración de síntomas en pacientes con trastornos neurológicos de origen psicológico (van Beilen, Griffioen, Gross y Leenders, 2009)	132
3.9.4. Estudio de prevalencia de exageración de síntomas en pacientes con epilepsia de origen psicológico (Benge et al., 2012).....	133
3.9.5. Estudio de Geraerts, Jelacic y Merckelbach (2006) sobre abuso sexual y simulación de síntomas	135

3.10. Visión general de las investigaciones realizadas con el SIMS	136
3.11. Conclusiones.....	157
Investigación empírica.....	159
4. Prevalencia de simulación por trastornos en incapacidad temporal: percepción de los profesionales de la salud	165
4.1. Objetivos	165
4.2. Método.....	166
4.2.1. Participantes.....	166
4.2.2. Procedimiento y materiales	166
4.3. Resultados	167
4.3.1. Percepción sobre las patologías más simuladas	167
4.3.2. Percepción sobre el perfil sociodemográfico prevalente en simulación	171
4.3.3. Contexto y mecanismo de simulación.....	175
4.3.4. Datos que hacen sospechar simulación a los profesionales	178
4.4. Discusión	182
4.5. Conclusiones.....	187
5. Utilidad diagnóstica del SIMS en la detección de simulación de psicopatología	191
5.1. Objetivos e hipótesis de trabajo.....	191
Estudio 1	192
5.2. Método.....	192
5.2.1. Participantes.....	192
5.2.2. Material	193
5.2.3. Procedimiento	193
5.3. Resultados	194
5.3.1. Puntuaciones en las escalas por grupo	194
5.3.2. Utilidad diagnóstica.....	198
5.4. Discusión	203
Estudio 2	208
5.5. Método.....	208
5.5.1. Participantes.....	208
5.5.2. Materiales.....	208
5.5.3. Procedimiento	208
5.5.4. Diseño.....	209
5.6. Resultados	210

5.6.1. Puntuaciones en las escalas por grupo	210
5.6.2. Utilidad diagnóstica	211
5.6.3. Relación con otras pruebas	214
5.7. Discusión	216
5.8. Conclusiones.....	218
6. Utilidad diagnóstica del SIMS en la detección de simulación de lumbalgia	221
6.1. Objetivos e hipótesis de trabajo.....	221
6.2. Método.....	223
6.2.1. Participantes.....	223
6.2.2. Materiales.....	224
6.2.3. Procedimiento	224
6.3. Resultados	225
6.3.1. Puntuaciones en las escalas por grupo	225
6.3.2. Utilidad diagnóstica	227
6.3.3. Relación con otras pruebas	229
6.4. Discusión	230
6.5. Conclusiones.....	232
7. Discusión general y conclusiones.....	235
Extended summary	247
Referencias bibliográficas	253
Apéndice A. Cuestionario de percepción de los profesionales de la salud de la prevalencia de simulación por trastornos en incapacidad temporal.	306
Apéndice B. Instrucciones a simuladores análogos de psicopatología.....	309
Apéndice C. Instrucciones a simuladores análogos de lumbalgia	311

Índice de tablas

Tabla 1.1 Modelos explicativos de simulación	14
Tabla 1.2 Tasas de simulación por contexto en el estudio de Mittenberg et al. (2002).....	20
Tabla 1.3 Tasas de simulación por patologías en el estudio de Mittenberg et al. (2002).....	21
Tabla 1.4 Estrategias de detección de exageración de síntomas en trastornos mentales.....	24
Tabla 1.5 Estrategias de detección de síntomas de trastornos mentales que requieren validación adicional	26
Tabla 1.6 Estrategias de detección de exageración de síntomas en trastornos físicos o médicos.....	27
Tabla 1.7 Criterios de diagnóstico de simulación de deterioro cognitivo de Slick et al. (1999).....	33
Tabla 1.8 Resumen de los posibles sesgos en los criterios de Slick et al. (1999) de acuerdo a Rogers, Bender, et al. (2011a).....	40
Tabla 2.1 Visión general de diseños de investigación empleados en simulación: diseños análogos y de grupos conocidos (Rogers, 2008a)	46
Tabla 2.2 Visión general de diseños de investigación empleados en simulación: diseños de medidas autosuficientes (bootstrapping comparison) y de prevalencia diferencial (Rogers, 2008a).....	47
Tabla 2.3 Aspectos clave a tener en cuenta en los diseños análogos (Rogers, 2008e)	48
Tabla 2.4 Clasificación de resultados de un estudio de validez de un nuevo test	62
Tabla 3.1 Factores e incentivos relativos a la exageración de síntomas (basado en Iverson, 2007)	72
Tabla 3.2 Escalas de exageración de síntomas del MMPI-2.....	76
Tabla 3.3 Escalas de exageración de síntomas del MMPI-2-RF.....	78
Tabla 3.4 Escalas de exageración de síntomas del PAI.....	80
Tabla 3.5 Descripción de las escalas y la puntuación total del SIMS.....	83
Tabla 3.6 Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y control, en los estudios originales de Smith y Burger (1997)	86
Tabla 3.7 Puntuaciones medias en las escalas del SIMS del grupo control y del grupo global de simuladores en los estudios de Smith y Burger (1997).....	86
Tabla 3.8 Sensibilidad y especificidad del SIMS en la muestra de estudio de Smith y Burger (1997) ...	87
Tabla 3.9 Sensibilidad y especificidad del SIMS en la muestra de validación cruzada de Smith y Burger (1997)	88
Tabla 3.10 Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y de respuesta sincera en el estudio de Edens et al. (1999).....	89
Tabla 3.11 Correlaciones entre las escalas del SIMS y el nivel de sintomatología psicopatológica en el estudio independiente de Edens et al. (1999)	89
Tabla 3.12 Sensibilidad y especificidad del SIMS en participantes bajo tratamiento psicológico o con un alto nivel de sintomatología psicopatológica en el estudio de Edens et al. (1999).....	90
Tabla 3.13 Puntuaciones medias en la escala total del SIMS en los estudios de la adaptación holandesa del SIMS de Merckelbach y Smith (2003).....	92
Tabla 3.14 Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos del estudio de Cima, Hollnack, et al. (2003).....	93
Tabla 3.15 Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales de Cima, Hollnack, et al. (2003).....	94
Tabla 3.16 Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y de respuesta sincera del estudio de Rogers et al., 2005	95

Tabla 3.17 <i>Puntuaciones medias y desviaciones típicas en la puntuación total y en la escala Baja inteligencia del SIMS de los grupos de simulación análogo y de personas con retraso mental moderado en el estudio de Graue et al. (2007)</i>	96
Tabla 3.18 <i>Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y diagnosticados con TEPT en el estudio de Efendov (2007)</i>	97
Tabla 3.19 <i>Tamaños del efecto (d de Cohen) e intervalo de confianza (95%) de cada uno de los grupos de simulación análoga frente al grupo de reclamantes de TEPT en el estudio de Efendov (2007)</i>	97
Tabla 3.20 <i>Puntuaciones medias en el SIMS del grupo de respuesta sincera, simuladores de amnesia no advertidos y simuladores de amnesia advertidos en el estudio de Giger et al. (2010)</i>	98
Tabla 3.21 <i>Características demográficas de los grupos de estudio de Buijs (2011)</i>	99
Tabla 3.22 <i>Puntuaciones medias en el SIMS del grupo de respuesta sincera de universitarios y pacientes con TDAH y del grupo de simulación de universitarios y de pacientes con TDAH en el estudio de Buijs (2011)</i>	99
Tabla 3.23 <i>Puntuaciones medias y desviaciones típicas en la puntuación total de los grupos de respuesta honesta, simulación en contexto civil y simulación en contexto criminal en el estudio de Dandachi-FitzGerald y Merckelbach (2013)</i>	100
Tabla 3.24 <i>Puntuación media y desviación típica en el grupo mixto que accedían a evaluación en el estudio de Dandachi-FitzGerald y Merckelbach (2013)</i>	101
Tabla 3.25 <i>Sensibilidad de las escalas del SIMS en adolescentes en el estudio de Rogers, Hinds y Sewell (1996)</i>	102
Tabla 3.26 <i>Puntuaciones medias en el SIMS de los grupos del estudio de Dearth (2007)</i>	103
Tabla 3.27 <i>Sensibilidad, especificidad y poder predictivo positivo y negativo de la puntuación Total del SIMS (Dearth, 2007)</i>	103
Tabla 3.28 <i>Puntuaciones medias y desviaciones típicas en la puntuación Total en el SIMS en el estudio de (Jelicic, Peters, et al., 2007) de simulación análoga de psicosis</i>	105
Tabla 3.29 <i>Puntuaciones medias y desviaciones típicas en la puntuación Total en el SIMS en el estudio de Jelicic, Merckelbach, et al. (2007) de simulación análoga de daño cerebral</i>	107
Tabla 3.30 <i>Puntuaciones medias y desviaciones típicas en la puntuación Total en el SIMS en el estudio de Jelicic et al. (2011) de simulación análoga de daño cerebral tras accidente</i>	108
Tabla 3.31 <i>Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales de simuladores análogos de esguince cervical en el estudio de Merten et al. (2008)</i>	109
Tabla 3.32 <i>Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales de simuladores análogos de trastorno por estrés postraumático. Estudio independiente de Merten et al. (2010)</i>	110
Tabla 3.33 <i>Correlación de la puntuación total del SIMS con las escalas de validez del PAI y la puntuación total en el SIRS (Edens et al., 2007)</i>	111
Tabla 3.34 <i>Correlación de la puntuación total del SIMS con las escalas de validez del PAI y la puntuación total en el SIRS (Edens et al., 2007)</i>	112
Tabla 3.35 <i>Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores. Estudio independiente de Lewis, Simcox y Berry (2002)</i>	113
Tabla 3.36 <i>Precisión predictiva de la puntuación total del SIMS. Estudio independiente de Lewis, Simcox y Berry (2002)</i>	114

Tabla 3.37 <i>Puntuaciones medias en el SIMS del grupo de simuladores de incapacidad para ser sometidos a juicio y medias estimadas de la población clínica y de simuladores en el estudio de Heinze y Purisch (2001)</i>	115
Tabla 3.38 <i>Sensibilidad de las escalas del SIMS para la detección de simuladores en el estudio de Heinze y Purisch (2001)</i>	115
Tabla 3.39 <i>Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores en el estudio de Vitacco et al. (2006)</i>	116
Tabla 3.40 <i>Sensibilidad y especificidad de las escalas del SIMS para la detección de simuladores en el estudio de Vitacco et al. (2006)</i>	117
Tabla 3.41 <i>Descripción de los grupos de la investigación de Alwes (2006)</i>	118
Tabla 3.42 <i>Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores de psicopatología en el estudio de Alwes (2006)</i>	118
Tabla 3.43 <i>Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores de problemas neurocognitivos en el estudio de Alwes (2006)</i>	119
Tabla 3.44 <i>Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores de psicopatología y de problemas neurocognitivos (ambos) en el estudio de Alwes (2006)</i>	119
Tabla 3.45 <i>Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores de psicopatología o de problemas neurocognitivos (alguno de los dos) en el estudio de Alwes (2006)</i> .	119
Tabla 3.46 <i>Tamaño del efecto (d de Cohen en valor absoluto) de las escalas del SIMS en la detección de simuladores en el estudio de Alwes (2006)</i>	120
Tabla 3.47 <i>Sensibilidad y especificidad de la puntuación total en el SIMS en los distintos grupos en el estudio de Alwes (2006)</i>	121
Tabla 3.48 <i>Tamaño del efecto (d de Cohen en valor absoluto) de las escalas del MMPI-2, SIMS, SIRS, TOMM, VSVT y LMT en la detección de simuladores en el estudio de Alwes (2006)</i>	122
Tabla 3.49 <i>Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y de respuesta sincera en el estudio de Clegg et al. (2009)</i>	123
Tabla 3.50 <i>Precisión predictiva de la puntuación total del SIMS al clasificar a los participantes como del grupo de respuesta honesta o del grupo de simuladores. Muestra de evaluación de discapacidad. Estudio de Clegg (2007)</i>	124
Tabla 3.51 <i>Precisión predictiva de la puntuación en cada una de las escalas del SIMS al clasificar a los participantes como del grupo de respuesta honesta o del grupo de simuladores. Muestra clínica. Estudio independiente de Clegg (2007)</i>	125
Tabla 3.52 <i>Puntuaciones medias y desviaciones típicas en la puntuación total del SIMS de los grupos de respuesta honesta, probable simulación y clara simulación en el estudio de Wisdom et al. (2010)</i>	125
Tabla 3.53 <i>Sensibilidad y especificidad de los distintos puntos de corte alternativos para la puntuación total del SIMS en el estudio de Wisdom et al. (2010)</i>	126
Tabla 3.54 <i>Puntuaciones medias y desviaciones típicas del SIMS y prueba t de Student entre los participantes clasificados como con y sin sospecha de simulación de acuerdo a sus resultados en el WMT (Green, 2005)</i>	127
Tabla 3.55 <i>Estadísticos descriptivos del SIMS del grupo de probable pacientes clínico genuino y de probable simulador en el estudio de González Ordi et al. (2010)</i>	128
Tabla 3.56 <i>Precisión predictiva de la puntuación total del SIMS al clasificar a los participantes como del grupo de probable paciente clínico o de probable simulador en el estudio de González Ordi et al. (2010)</i>	128

Tabla 3.57 <i>Medias, desviaciones típicas y estadístico U de Mann-Whitney del grupo con sospecha de simulación y del grupo sin sospecha de simulación del estudio de Blasco Saiz y Pallardó Durá (2013)</i>	129
Tabla 3.58 <i>Descriptivos en el SIMS del grupo de pacientes psiquiátricos ambulatorios. Estudio independiente de Dandachi-FitzGerald et al.(2011)</i>	130
Tabla 3.59 <i>Descriptivos en la puntuación total del SIMS del grupo de pacientes con esquizofrenia en el estudio de Peters et al. (2012)</i>	131
Tabla 3.60 <i>Especificidad de puntos de corte alternativos para la puntuación total del SIMS en la muestra global de pacientes con esquizofrenia (n=41) en el estudio de Peters et al. (2012)</i>	132
Tabla 3.61 <i>Puntuaciones medias y desviaciones típicas en la puntuación total de los grupos de trastornos neurológicos con y sin origen psicológico en el estudio de van Beilen et al. (2009)</i>	133
Tabla 3.62 <i>Descriptivos en el SIMS del grupo de epilepsia y epilepsia de origen psicológico. Estudio independiente de Bengé et al. (2012)</i>	133
Tabla 3.63 <i>Área bajo la curva de Total, Deterioro neurológico y Trastornos afectivos del SIMS para diferencia entre epilepsia y epilepsia de origen psicológico. Estudio independiente de Bengé et al.(2012)</i>	134
Tabla 3.64 <i>Sensibilidad, especificidad y poder predictivo positivo y negativo de la puntuación Total del SIMS para diferencia entre epilepsia y epilepsia de origen psicológico (Bengé et al., 2012)</i>	134
Tabla 3.65 <i>Sensibilidad, especificidad y poder predictivo positivo y negativo de la puntuación Deterioro neurológico del SIMS para diferencia entre epilepsia y epilepsia de origen psicológico (Bengé et al., 2012)</i>	135
Tabla 3.66 <i>Sensibilidad, especificidad y poder predictivo positivo y negativo de la puntuación Trastornos afectivos del SIMS para diferencia entre epilepsia y epilepsia de origen psicológico (Bengé et al., 2012)</i>	135
Tabla 3.67 <i>Puntuaciones medias y desviaciones típicas en la puntuación total en el SIMS de los grupos de abuso sexual y de respuesta sincera en el estudio de Geraerts et al., 2006</i>	136
Tabla 3.68 <i>Características de las investigaciones realizadas sobre el SIMS en la literatura científica</i>	138
Tabla 3.69 <i>Estudios estadísticos empleados en las investigaciones realizadas sobre el SIMS</i>	141
Tabla 3.70 <i>Tamaño del efecto (d Cohen) de las investigaciones que han empleado diseños de simulación análoga sin grupo clínico de referencia: comparación de simuladores análogos frente a respuesta honesta normales.</i>	145
Tabla 3.71 <i>Tamaño del efecto (d Cohen) de las investigaciones que han empleado diseños de simulación análoga con grupo clínico de referencia: comparación de simuladores análogos frente a respuesta honesta clínicos.</i>	147
Tabla 3.72 <i>Tamaño del efecto (d Cohen) de las investigaciones que han empleado diseños de grupos conocidos o autosuficientes: comparación de simuladores frente a clínicos en contextos reales.</i>	148
Tabla 3.73 <i>Puntuaciones medias en Total SIMS en diferentes grupos en las investigaciones encontradas en la literatura científica</i>	150
Tabla 3.74 <i>Puntuaciones medias en Total SIMS según contexto civil o penal en muestras de simuladores y clínicos litigantes encontradas en la literatura científica</i>	150
Tabla 3.75 <i>Puntuaciones medias en Total SIMS según trastorno instruido en diferentes grupos de simulaciones análogas encontradas en la literatura científica</i>	151
Tabla 3.76 <i>Puntuaciones medias en las escalas del SIMS en diferentes condiciones de respuesta encontradas en la literatura científica</i>	151

Tabla 3.77 Puntuaciones medias en las escalas del SIMS según contexto civil o penal en muestras de simuladores y clínicos litigantes encontradas en la literatura científica.....	152
Tabla 3.78 Puntuaciones medias en las escalas del SIMS según trastorno instruido en diferentes grupos de simulaciones análogas encontradas en la literatura científica	152
Tabla 3.79 Sensibilidad y especificidad de la puntuación total en el SIMS en los diferentes estudios de grupos diagnósticos conocidos o diseños autosuficientes en simulación de psicopatología o trastornos cognitivos en adultos.....	153
Tabla 3.80 Sensibilidad y especificidad de la puntuación en las escalas del SIMS en los diferentes estudios de grupos diagnósticos conocidos o diseños autosuficientes en simulación de psicopatología o trastornos cognitivos en adultos	154
Tabla 3.81 Correlaciones halladas en la literatura científica entre el SIMS y otras pruebas de evaluación	155
Tabla 4.1 Ranking de prevalencia de simulación en diferentes patologías de acuerdo a la percepción de los profesionales (1= máxima prevalencia de simulación; 10 = mínima prevalencia de simulación)	168
Tabla 4.2 Otros cuadros citados por los profesionales en la percepción de 1 a 10 de prevalencia de simulación (1= máxima prevalencia de simulación; 10 = mínima prevalencia de simulación)	169
Tabla 4.3 Porcentaje de prevalencia de simulación percibido por los profesionales en diferentes patologías.....	170
Tabla 4.4 Percepción de los profesionales sobre la presencia de mujeres y varones en simulación ...	172
Tabla 4.5 Ranking de prevalencia de simulación en los grupos de edad de acuerdo a la experiencia y percepción de los profesionales (1= máxima prevalencia de simulación; 5 = mínima prevalencia de simulación)	173
Tabla 4.6 Ranking de prevalencia de simulación según los niveles de ingresos de acuerdo a la experiencia y percepción de los profesionales (1= máxima prevalencia de simulación; 6 = mínima prevalencia de simulación).....	173
Tabla 4.7 Profesiones más vinculadas a la presencia de exageración de síntomas de acuerdo a la percepción de los profesionales	175
Tabla 4.8 Percepción de los profesionales sobre el tipo de contingencia donde se da más simulación	175
Tabla 4.9 Percepción de los profesionales sobre qué porcentaje de simuladores utiliza cada una de las siguientes estrategias de simulación.....	177
Tabla 4.10 Datos clínicos que hacen sospechar a los profesionales la existencia de simulación de acuerdo a su experiencia.....	179
Tabla 5.1 Medias, desviaciones típicas, estadístico F y tamaños del efecto para los grupos control, clínico y simuladores análogos.....	194
Tabla 5.2 Medias, desviaciones típicas, estadístico F y tamaños del efecto para los grupos clínico y simulador análogo por condición	196
Tabla 5.3 Estadísticos de utilidad diagnóstica de la puntuación total en el SIMS en la discriminación entre pacientes clínicos y simuladores análogos.....	199
Tabla 5.4 Estadísticos de utilidad diagnóstica de la puntuación total en el SIMS en la discriminación entre pacientes clínicos y simuladores análogos por condición	200
Tabla 5.5 Estadísticos de utilidad diagnóstica de las escalas del SIMS en la discriminación entre pacientes clínicos y simuladores análogos en la condición de trastorno psicótico	201

Tabla 5.6 Estadísticos de utilidad diagnóstica de las escalas del SIMS en la discriminación entre pacientes clínicos y simuladores análogos en la condición de deterioro neurológico	202
Tabla 5.7 Estadísticos de utilidad diagnóstica de las escalas del SIMS en la discriminación entre pacientes clínicos y simuladores análogos en la condición de trastorno afectivo	203
Tabla 5.8 Medias, desviaciones típicas, estadístico F y tamaño del efecto para los grupos de probable paciente clínico genuino y probable simulador	210
Tabla 5.9 Mediana, rango intercuartílico (IQR), rango promedio y estadístico U de Mann-Whitney del grupo sin sospecha de simulación y con sospecha de simulación en las escalas del SIMS.....	211
Tabla 5.10 Estadísticos de utilidad diagnóstica de la puntuación total en el SIMS en la discriminación entre pacientes con y sin sospecha de simulación	212
Tabla 5.11 Estadísticos de utilidad diagnóstica de la puntuación en la escala Psicosis (Ps) en la discriminación entre pacientes con y sin sospecha de simulación	212
Tabla 5.12 Estadísticos de utilidad diagnóstica de la puntuación en la escala Deterioro neurológico (Dn) en la discriminación entre pacientes con y sin sospecha de simulación.....	213
Tabla 5.13 Estadísticos de utilidad diagnóstica de la puntuación en la escala Trastornos amnésicos (Am) en la discriminación entre pacientes con y sin sospecha de simulación.....	213
Tabla 5.14 Estadísticos de utilidad diagnóstica de la puntuación en la escala Baja inteligencia (Bi) en la discriminación entre pacientes con y sin sospecha de simulación	213
Tabla 5.15 Estadísticos de utilidad diagnóstica de la puntuación en la escala Trastornos afectivos (Af) en la discriminación entre pacientes con y sin sospecha de simulación	214
Tabla 5.16 Correlaciones entre las escalas del SIMS y las escalas de exageración y minimización de síntomas del MMPI-2 en pacientes evaluados en contexto forense civil.	214
Tabla 5.17 Correlaciones entre las escalas del SIMS y las escalas de exageración y minimización de síntomas del MMPI-2-RF en pacientes en contexto forense civil.	215
Tabla 6.1 Medias, desviaciones típicas, estadístico F y tamaños del efecto para los grupos control, clínico sin sospecha de simulación, clínico con sospecha de simulación y simuladores análogos	226
Tabla 6.2 Estadísticos de utilidad diagnóstica de la puntuación Total en el SIMS en la discriminación entre pacientes clínicos con y sin sospecha de simulación.	228
Tabla 6.3 Estadísticos de utilidad diagnóstica de la escala Deterioro neurológico (Dn) en la discriminación entre pacientes clínicos con y sin sospecha de simulación.....	229
Tabla 6.4 Estadísticos de utilidad diagnóstica de la escala Trastornos afectivos (Af) en la discriminación entre pacientes clínicos con y sin sospecha de simulación.....	229
Tabla 6.5 Correlaciones entre las escalas del SIMS y las escalas de exageración y minimización de síntomas del MMPI-2-RF	230

Índice de figuras

<i>Figura 1.1.</i> Propuesta de modelo adaptativo de la simulación como estilo de respuesta ante la enfermedad, tomado de González Ordi et al. (2012).	15
<i>Figura 3.1.</i> Puntuaciones medias en el total del SIMS por tipo de muestras en las distintas investigaciones existentes en la literatura científica	149
<i>Figura 4.1.</i> Pregunta 1 del cuestionario	167
<i>Figura 4.2.</i> Pregunta 2 del cuestionario	170
<i>Figura 4.3.</i> Pregunta 8 del cuestionario	171
<i>Figura 4.4.</i> Percepción de los profesionales sobre la presencia de mujeres y varones en los pacientes que simulan sus síntomas	171
<i>Figura 4.5.</i> Preguntas 3 y 5 del cuestionario	172
<i>Figura 4.6.</i> Pregunta 4 del cuestionario	174
<i>Figura 4.7.</i> Preguntas 6 y 7 del cuestionario	176
<i>Figura 4.8.</i> Pregunta 9 del cuestionario	178
<i>Figura 4.9.</i> Número de menciones por parte de los profesionales a cada dato de sospecha de simulación	181
<i>Figura 5.1.</i> Puntuaciones medias para el grupo control, clínico y de simulador análogos en las escalas del SIMS.....	195
<i>Figura 5.2.</i> Medias para la condición de trastorno psicótico: grupo clínico y grupo de simuladores análogos.	197
<i>Figura 5.3.</i> Medias para la condición de deterioro neurológico: grupo clínico y grupo de simuladores análogos	197
<i>Figura 5.4.</i> Medias para la condición de trastornos afectivos: grupo clínico y grupo de simuladores análogos	198
<i>Figura 5.5.</i> Medias para los grupos de probable respuesta honesta y probable simulación	211

Introducción

Las ciencias de la salud descansan implícitamente en la asunción de que el paciente es un participante voluntario del proceso que proporciona información veraz y honesta en la medida de sus posibilidades y capacidades (Rogers, 1988). Ahora bien, ¿qué ocurre cuando la información que aporta no es veraz? ¿Qué sucede cuando busca obtener algún tipo de beneficio secundario de su patología y, por lo tanto, crea, exagera o mantiene en el tiempo su patología?

Desafortunadamente, esta exageración o simulación de la patología no parece ser un hecho aislado o espurio, tal y como erróneamente se ha concebido en ocasiones. La literatura científica informa de tasas del 7% en la práctica clínica general y de entre un 14 y un 30% en casos de solicitud de discapacidad o indemnización laboral (Mittenberg, Patton, Canyock y Condit, 2002; Rogers, Salekin, Sewell, Goldstein y Leonard, 1998).

Estas prevalencias hacen que la correcta determinación de la veracidad de los síntomas informados por el paciente sea un primer paso necesario en muchos contextos. Particularmente, se recomienda evaluar la posible presencia de simulación en aquellos casos que los resultados de la evaluación tengan implicaciones relevantes para la vida del paciente (Bush et al., 2005; Rogers, 2008a). Además, este problema no solo afecta a la práctica profesional sino a la investigación y a la creación de conocimiento científico. Por ejemplo, Rosen (2006) ha alertado de la necesidad de descartar la presencia de simulación a la hora de establecer tasas de prevalencia del trastorno de estrés postraumático y la necesidad de que revisores de revistas, editores y autores incluyan este requisito para estas publicaciones. Algo similar podría afirmarse con respecto a otros cuadros con tasas relevantes de simulación como la ansiedad, la depresión, la lumbalgia, la cervicalgia o síndromes complejos de dolor (v.g. Victor, Boone y Kulick, 2010).

Ahora bien, ¿cómo proceder a detectarla? Podría decirse que hasta hace unas décadas era una labor que recaía en la impresión clínica (Sweet, 2008). Como decía en 1964 el juez Potter Stewart de la Corte Suprema de EEUU en una cita habitualmente mencionada al tratar de definir la pornografía: "I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description ["hard-core pornography"] (...) But I know it when I see it" (O'Meara y Shaffer, 1964). Así, durante muchos años, los profesionales de las ciencias de la salud han presumido de no entrar a definir la simulación ya que "lo saben cuando lo ven (I know it when I see it)". No obstante esta visión fue puesta en tela de juicio cuando durante la década de 1980 y 1990 una serie de estudios empíricos mostraron que los clínicos no identificaban a los simuladores

más allá de lo esperable por azar (Faust y Ackley, 1998; Hall y Pritchard, 1996). Estudios recientes de meta-análisis han apuntado en esta misma dirección, indicando que los psicólogos no destacan frente a la población general en su capacidad para detectar mentiras (Aamodt y Custer, 2006).

Esta toma de conciencia de la falibilidad del “juicio clínico” provocó un progresivo cambio hacia un enfoque proactivo, que desde el principio incluyera la evaluación de la simulación como algo relevante y que debiera ser sistemáticamente considerado. Tal y como afirmaba Sweet (1999):

We can no longer take comfort in the belief, now shown to be false, that we simply know when malingering is evident, without addressing the issue deliberately and prospectively. (...) Gross observations of behavior as a means of determining presence or absence of malingering are not longer acceptable as sole criteria for ruling out insufficient effort. (p. 278)

Como consecuencia lógica de la mayor conciencia de esta necesidad, se ha producido un auge de la investigación en éste área. Una búsqueda en PsycINFO introduciendo el término “malingering” nos arroja 2.681 registros de investigaciones en la última década (2000-2010), frente a 738 en la década anterior (1990-1999) y 605 en el periodo comprendido entre 1896 y 1989. Parece que, si bien las ciencias de la salud estaban relativamente bien dotadas de herramientas de evaluación de la patología, estaban menos provistas para la detección de la característica singular que supone la simulación (Arce, del Carmen Pampillón y Fariña, 2002), lo que ha provocado a nivel internacional esta auténtica explosión de desarrollo y publicación de instrumentos con este objetivo (Otto y Heilbrun, 2002).

Dada la relativa novedad de muchos de los índices creados y su enorme impacto social y económico, la cuestión sobre su precisión y correcta clasificación ha ocupado múltiples investigaciones y ha provocado incesantes debates (McGrath, Mitchell, Kim y Hough, 2010; Merten y Merckelbach, 2013; Morey, 2012; Rohling et al., 2011) llegando incluso a llamar la atención de periódicos como el *Wall Street Journal* (Armstrong, 2008). No es de extrañar, por lo tanto, que tal y como llegaron a afirmar Farkas, Rosenfeld, Robbins y van Gorp (2006) “malingering test accuracy is increasingly a major issue in psychology and law” (p. 659).

Aunque el proceso de detección de simulación mayoritariamente vigente actualmente es multimétodo y multisistema (v.g., González Ordi, Santamaría Fernández y Capilla Ramírez, 2012, McDermott y Feldman, 2007), en él han llegado a ocupar un lugar destacado los instrumentos de evaluación mediante autoinforme (Farkas et al., 2006; Inda Caro, Lemos Giráldez, López Rodrigo y

Alonso Rionda, 2005; Samuel y Mittenberg, 2005), que han acaparado una importante parte de la investigación realizada (Sweet y Guidotti Breting, 2013).

El uso de las adaptaciones españolas de estos autoinformes requiere de estudios específicos que avalen su uso y establezcan puntos de corte específicos con el fin de evitar errores diagnósticos debidos a las diferencias culturales e idiomáticas existentes (Puente y Perez-Garcia, 2000a, 2000b), más aún cuando son escasas las pruebas específicas desarrolladas en este campo en países de lengua española (Verdejo, Alcazar-Corcoles, Gomez-Jarabo y Perez-Garcia, 2004). En esta línea, Inda Caro et al. (2005) afirmaban que:

La Psicología tiene en España un importante reto por delante en lo que se refiere a la determinación, con base científica, de la validez de los testimonios y de la sintomatología aducida por personas que están en proceso de enjuiciamiento o que han sido víctimas de accidentes o agresiones. En particular, es necesario desarrollar procedimientos estructurados y estandarizados, que permitan emitir juicios fundamentados sobre la posible existencia de simulación. (p. 106)

La presente tesis pretende precisamente abordar la utilidad en muestras españolas de una de las herramientas de *screening* más empleadas y relevantes a nivel internacional en detección de simulación de sintomatología, el Inventario Estructurado de Simulación de Síntomas (SIMS; Smith y Burger, 1997; Widows y Smith, 2005). Esta prueba ha sido considerada, por la amplitud de áreas evaluadas, una de las herramientas más valiosas para la medida del sesgo de respuesta (Dandachi-FitzGerald y Merckelbach, 2013; Wisdom, Callahan y Shaw, 2010).

El objetivo de esta tesis es, por lo tanto, determinar la utilidad de sus distintas escalas en diferentes patologías y contextos en población española, así como los puntos de corte más adecuados en nuestro país, de forma que puedan establecerse criterios que ayuden a la detección de posibles simuladores en la práctica clínica de acuerdo a la evidencia empírica proporcionada.

Para la elaboración de esta tesis doctoral se ha dividido el trabajo en dos grandes apartados: el primero, centrado en el marco teórico; el segundo, dedicado a la investigación empírica.

El primer apartado, marco teórico, se divide a su vez en tres bloques. En el primer capítulo, “concepto y delimitación de la simulación”, se presenta una aproximación conceptual a la simulación, los modelos teóricos que han tratado de articularla, los estudios de prevalencia disponibles y los

criterios diagnósticos más mencionados en la actualidad. En el segundo capítulo, “metodología de investigación en simulación”, se trata de sintetizar los diseños y técnicas de análisis más frecuentes en la investigación en el campo de la simulación. En el tercer capítulo se presentan las distintas técnicas de detección de la simulación, centrándose en el uso de los autoinformes y, particularmente, en el Inventario Estructurado de Simulación de Síntomas, SIMS, así como las evidencias empíricas disponibles sobre su uso. En su conjunto, las investigaciones realizadas en las dos últimas décadas han permitido afianzar la delimitación conceptual de la simulación, los posibles modelos explicativos y los criterios diagnósticos. Este hecho, junto con el desarrollo de la metodología apropiada para su estudio ha confluído en el avance realizado en el campo y ha sentado las bases para su crecimiento y desarrollo.

El segundo apartado, dedicado a la investigación empírica, se estructura también en tres bloques que describen los tres estudios empíricos que componen esta tesis. El primer estudio empírico, Prevalencia de simulación por trastornos en incapacidades temporales: percepción de los profesionales de la salud, aborda qué patologías son percibidas por los profesionales como más susceptibles de simulación en el contexto médico-legal y qué estrategias utilizan para su detección. Los estudios segundo y tercero se centran respectivamente en el análisis de la capacidad discriminativa del SIMS en la detección de simulación de psicopatología y de dolor crónico. En concreto, el segundo estudio, Utilidad diagnóstica del SIMS en la detección de simulación de psicopatología, se compone a su vez de dos estudios: uno con un diseño de simulación análoga y otro con un diseño de medidas autosuficientes en contextos médico-legales. Finalmente, el tercer estudio, Utilidad diagnóstica del SIMS en la detección de simulación de lumbalgia, se centra específicamente en la detección de la simulación de la lumbalgia crónica como uno de los trastornos más prevalentes en dolor crónico.

A partir de los datos que se obtengan se pretende aportar procedimientos que permitan cribar las sospechas clínicas de simulación de los profesionales de la salud, además de contribuir a un mayor conocimiento sobre el patrón de exageración de síntomas en población española.

Marco teórico

1. Concepto y delimitación de la simulación

1.1. Definición y modelos explicativos de la simulación

1.1.1. El concepto de simulación en el DSM

El concepto de simulación aparece referido por primera vez en el Manual Diagnóstico y Estadístico de los Trastornos Mentales en su tercera edición (DSM-III), publicada en 1980. Su definición y criterios de sospecha no han sufrido grandes variaciones desde la versión inicial hasta el actual DSM-5 (APA, 2013), en el que se define la simulación como:

The intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs. Malingering differs from factitious disorder in that the motivation for the symptom production in malingering is an external incentive, whereas in factitious disorder external incentives are absent. Malingering is differentiated from conversion disorder and somatic symptom-related mental disorders by the intentional production of symptoms and by the obvious external incentives associated with it. (American Psychiatric Association, 2013, pp. 726–727)

Esta definición resalta cuatro componentes clave: la invención o exageración de síntomas, la magnitud requerida, la finalidad utilitaria para obtener beneficios externos y la voluntad consciente del engaño.

El **primer componente** determina que la simulación puede ser debida tanto a la invención de síntomas falsos como a su exageración. Resnick (1997) diferenciaba, a este respecto, la simulación pura (fingimiento de un trastorno no existente) de la simulación parcial (exageración consciente de síntomas presentes o de un trastorno ya superado). Asimismo, Lipman (1962) clasificaba la simulación de síntomas como inventados (la persona no tiene ningún síntoma pero representa que lo tiene de manera engañosa), exagerados (la persona refiere síntomas más graves de los reales) o perseverados (los síntomas que en algún momento estuvieron presentes han cesado, pero se continúa alegando su presencia).

Considerada de forma estricta esta definición, quedaría excluida la falsa imputación (en terminología de Resnick) o la transferencia genuina (en terminología de Lipman), consistente en síntomas reales que son atribuidos fraudulentamente a un daño particular. Por ejemplo, una persona puede alegar alteraciones cognitivas tras un accidente de coche, cuando realmente dichas alteraciones pueden estar causadas por su historial previo de abuso de alcohol. En estos casos, el evaluado puede ocultar aspectos previos de su historial o presentarse como en perfecto estado mental previamente al accidente, para crear una falsa impresión de pérdida de funcionalidad (Faust y Ackley, 1998). De hecho, estos intentos de dirigir la atención del clínico hacia otros factores causales, que no son los verdaderos, suponen prácticas habituales de simulación en contextos legales que, dadas sus implicaciones, requieren la máxima atención de los profesionales e investigadores.

Podríamos considerar, por lo tanto, que los simuladores pueden: (a) simular *ex novo* sus síntomas y sus correlatos clínicos; (b) exagerar de forma flagrante la discapacidad o alteración causada por una patología real, motivo de cierto grado de compensación o discapacidad, o (c) atribuir fraudulentamente determinados síntomas reales a una causa o condición que pueda ser indemnizable o compensable (Rogers, Payne, Berry y Granacher, 2009).

Es importante remarcar, por tanto, que se considera simulación, tanto su presentación en su forma más pura (no existe ningún síntoma y se inventan o generan), como en su forma parcial (existiendo algún síntoma se exagera, persevera o atribuye erróneamente). Por tanto, la presencia acreditada de síntomas genuinos en la persona (mediante resonancia magnética o cualquier otra técnica diagnóstica) no excluye automáticamente la posible existencia de simulación. En estos casos, puede estar dándose una exageración de los mismos o de la discapacidad que causan, una atribución fraudulenta a una causa indemnizable o una perseveración de sus consecuencias. Tal y como afirma Iverson (2010), resulta absurdo descartar automáticamente la presencia de simulación en estos casos ya que eso sería equivalente a afirmar que estas personas, por su trastorno, son incapaces de implicarse en conductas orientadas a objetivos (esto es, la exageración, la perseveración o la atribución fraudulenta de los síntomas para obtener compensaciones, indemnizaciones o cualquier tipo de beneficio externo).

El **segundo componente (la magnitud requerida)** remarca que lo que se pretende detectar es la asignación o reconocimiento de múltiples y muy marcados síntomas psicológicos. En otras palabras, no se considera simulación una leve exageración o una relación de síntomas aislados. Tal y como afirma Rogers (2008): *“An important consideration is the magnitude of the dissimulation; it must be the fabricant or gross exaggeration of multiple symptoms. The presence of minor exaggerations or isolated symptoms does not qualify as malingering”* (p. 5).

Los **otros dos componentes** incluidos en la definición del DSM, presencia de incentivos externos y requisito de intencionalidad, establecen criterios diferenciales con respecto a los trastornos facticios y somatomorfos, respectivamente. A diferencia de los trastornos facticios, que se caracterizan por la meta de asumir el papel de enfermo, en la simulación, el objetivo perseguido por la persona debe ser un beneficio externo. Ahora bien, la mera existencia de incentivos externos no excluye automáticamente el diagnóstico de trastorno facticio, sino que ha de justificarse de un modo razonado, fundamentando las principales metas externas que se infiere que está siguiendo la persona evaluada. No parece ayudar en este complejo diagnóstico diferencial la difusa delimitación de los trastornos facticios en el DSM y la falta de especificidad en los criterios de inclusión (Berry y Nelson, 2010; Hamilton, Feldman y Cunnien, 2008).

En cuanto a los trastornos somatomorfos, la diferencia con la simulación viene marcada por el último componente requerido, la intencionalidad de la conducta. La simulación debe ser realizada intencional y conscientemente. No obstante, este asunto ha resultado particularmente polémico y diversos autores han cuestionado la necesidad y viabilidad de establecer la intencionalidad de una conducta en relación con el diagnóstico de simulación (v.g. Boone, 2007b, Drob, Meehan y Waxman, 2009).

Al igual que lo hacía el DSM-IV-TR, el DSM-5 añade a la definición anterior **cuatro criterios que deben hacer sospechar de simulación:**

- (1) presentación en un contexto médico-legal;
- (2) discrepancia acusada entre el estrés o la alteración explicados por la persona y los datos objetivos de la exploración médica;
- (3) falta de cooperación durante la valoración diagnóstica e incumplimiento del régimen de tratamiento prescrito; y
- (4) presencia de trastorno antisocial de la personalidad.

No obstante, estos criterios adicionales parecen haber resultado poco específicos para ser fiables en la detección de simulación (Gerson, 2002). La principal crítica que han recibido es que, si bien son características comunes a los simuladores, no son distintivas de estos. Esto es, son criterios que con frecuencia se dan en simuladores, pero que también se dan en otras personas no simuladoras, sin que resulten suficientemente discriminativos. De hecho, su uso de forma única puede llegar a incrementar dramáticamente el número de falsos positivos (Rogers y Vitacco, 2002; Rogers, 1990b). En esta línea, el estudio de Rogers (1990b) mostró que el empleo aislado de estos criterios provocaba que cuatro de cada cinco individuos clasificados como sospechosos de simulación

fueran falsos positivos. No en vano, tal y como señalan Resnick y Zuchowski (2007), en un contexto penal es muy probable que cualquier persona cumpliera casi por defecto al menos tres de las cuatro características mencionadas. A continuación se revisará cada uno de estos criterios.

(1) Presentación en un contexto médico-legal. Si bien es cierto que el contexto de evaluación influye en las tasas base de simulación (Mittenberg et al., 2002), la referencia al entorno médico-legal que se hace en el primer criterio parece resultar imprecisa ya que engloba múltiples contextos donde no necesariamente se da simulación y puede dejar fuera otros escenarios adversos para la persona donde puede aparecer la simulación. En este sentido, parece resultar un criterio poco útil. Por ejemplo, si pensamos en las evaluaciones legales para la custodia de los hijos, no será probable que encontremos simulación sino más bien lo contrario ya las que las personas suelen minimizar sus síntomas, aún siendo un contexto médico-legal. Igualmente, en las evaluaciones médico-legales a presos para decidir sobre su futura adaptación a la sociedad o posibilidad de reincidencia, es más probable que esté presente la disimulación que la simulación. Es por estos motivos que este criterio adolece de una limitada especificidad (Rogers, 1990b) y es preferible definir las características que debe requerir un contexto (p. ej., que sea adversos a los intereses de la persona, que estén disponibles ganancias difícilmente conseguibles por otros medios diferentes a la simulación...), en vez de ceñirse a esa etiqueta general de contexto médico-legal.

(2) Discrepancia con la información objetiva. En relación con el segundo criterio, resulta de difícil uso práctico debido al inherente carácter subjetivo, característico de la mayoría de las evaluaciones de trastornos mentales y determinados trastornos físicos vinculados a dolor crónico y a trastornos cognitivos. Incluso cuando se trata de cruzar los datos de autoinformes y heteroinformes para comprobar la fiabilidad y objetividad del informador, ha de hacerse con extrema precaución, dadas las distorsiones que es frecuente encontrar en ambas partes. Las investigaciones con pacientes normales o clínicos suelen mostrar niveles moderados o bajos de concordancia entre distintos informantes. Sirva de ejemplo el estudio de Meyer (2002), que evidenció una correlación promedio de 0,26 entre informantes en relación con distintos aspectos de funcionamiento en adultos, o el estudio de Klonsky, Oltmanns y Turkheimer (2002), que determinó una correlación promedio de 0,39 entre diversos informantes en relación con la presencia de trastornos de la personalidad. En síntesis, tal y como concluyen Achenbach, Krukowski, Dumenci y Ivanova (2005), en un meta-análisis sobre la correlación entre autoinformes y heteroinformes: "Do people see themselves as others see them? The mean cross-informant correlations suggest that they may not. The findings indicate that people may often provide different pictures of their problems than would be obtained from informants who know them (...) The limited agreement indicated by the cross-informant correlations presents challenges for diagnostic assessment ..." (p.370).

(3) Falta de cooperación. Puede considerarse que el tercer criterio peca de similares deficiencias que los criterios anteriores debido a su escasa especificidad ya que, por ejemplo, es habitual observar que personas con esquizofrenia, con trastornos de la conducta alimentaria o con problemas de abuso de sustancias sean, a menudo, poco cooperativas y desobedientes con sus citas de evaluación y tratamiento (Rogers, 1997) o que, incluso, el incumplimiento de las pautas terapéuticas indicadas por el profesional sea más bien resultado de la negación del problema más que de la existencia de simulación (Faust y Ackley, 1998).

(4) Trastorno antisocial. Finalmente, no parece que **el trastorno antisocial** sea un buen indicador, siendo escasamente específico y presentando un grave riesgo de conducir a un alto número de falsos positivos (Rogers, 1990b). Tal y como apuntaban Kucharski, Falkenbach, Egan y Duncan (2006): “the predictive utility of APD [antisocial personality disorder] in discriminating malingerers from honest responders was poor, with many with APD not attempting to mangle and those without APD presenting potentially malingered profiles” (p. 195). Además, el hecho de centrarse en la conducta antisocial puede hacer que el clínico o forense esté menos atento a su detección en personas que no sean antisociales o psicopáticas. Cabe añadir, también, que su inclusión como criterio puede dar lugar fácilmente a confusión, pudiendo llegar a inferirse erróneamente que la presencia de un trastorno antisocial de la personalidad descarta la presencia de otros trastornos genuinos comórbidos, esto es, la presencia de sociopatía no ofrece una inmunidad natural a los trastornos del eje I (Poitress, Edens y Watkins, 2001).

En síntesis, estos criterios parecen caracterizarse por una limitada especificidad y escasa utilidad ya que establecen características comunes en los simuladores, pero no distintivas. En palabras de Berry y Nelson (2010): “The evolution of symptom validity and malingering literature in recent decades has culminated in a sophisticated conception of malingering that essentially renders DSM criteria obsolete” (p. 296).

Otros autores, como Rogers (1990a), han alertado también de que los criterios del DSM, aparte de su limitada utilidad diagnóstica, reflejan un modelo puritano de la simulación, que la conceptualiza como manifestación de maldad: “the only unifying theme of this discordant amalgam is that of “badness”: namely, a bad person (sociopath) in a bad situation (forensic assessment), who is a bad participant (lack of cooperation)” (Rogers, 1990b, p. 183). Basta con encontrarse en una situación desacertada para sospechar de cualquier persona como potencial simuladora, lo que ha hecho que este modelo sea denominado modelo criminológico o puritano (Rogers, 2008a).

Rogers también señala críticamente el escaso sustrato teórico de estos criterios. Considera que la concepción teórica de la simulación debe cumplir con su papel de generadora de hipótesis, a través de un modelo que oriente y englobe la investigación. En este sentido, el modelo puritano o criminológico subyacente al DSM-5 parece responder más a la pregunta sobre quién simula que a la cuestión de por qué simula (Rogers, 1990b).

Con el objetivo de desarrollar un adecuado marco teórico y metodológico, Rogers ha desarrollado una línea de investigación articulando un modelo adaptativo de la simulación, que será descrito en los próximos apartados (Rogers, Sewell y Goldstein, 1994; Rogers, 1990a, 1990b).

1.1.2. El modelo adaptativo

Todas las personas en algún momento de su vida es probable que hayan simulado una dolencia para conseguir algo a cambio ya fuera un día de descanso del colegio, el retraso de una tarea o, incluso, algo más de atención o cuidado de familiares o amigos. La mentira parece ser un proceso de interacción social cotidiano, siendo un acto habitual de la vida diaria, más que un hecho extraordinario (DePaulo, Kashy, Kirkendol, Wyer y Epstein, 1996). Por lo general, se aprende a engañar a edades muy tempranas y es una conducta que aparece en todos los estatus socioeconómicos. En definitiva, forma parte de la conducta adaptativa que el ser humano adquiere en su desarrollo.

La simulación, concebida dentro del modelo adaptativo, estaría en el extremo de ese continuo que es el engaño. Distintos estudios apuntan precisamente en esta dirección, indicando que la exageración o fabricación de síntomas es un constructo dimensional, en vez de categórico (Walters et al., 2008; Walters, Berry, Rogers, Payne y Granacher, 2009).

El desarrollo del modelo adaptativo pretende proporcionar un modelo explicativo de la simulación que deje de lado las consideraciones patologizantes o criminalizantes de otros modelos conceptuales, para enfocarlo desde la teoría de la decisión, según la cual las decisiones y opciones se hacen bajo condiciones de incertidumbre y basadas en la probabilidad y utilidad esperada (Kahneman y Tversky, 1984). De esta manera, esta aproximación defiende que un potencial simulador hace un análisis de coste-beneficio de sus opciones en contextos adversos o indiferentes y puede optar por la fabricación o exageración de síntomas, como vía para lograr sus objetivos o los incentivos perseguidos ya sea la evitación de un problema o castigo (evitación del dolor) o la obtención de una compensación o beneficio (búsqueda del placer), que constituyen los dos grandes ejes de la motivación (Resnick y Knoll, 2005). Esto no quiere decir que los simuladores sean precisos

en sus análisis, ni que dicha conducta sea “adaptativa” o correcta desde un punto de vista moral, legal o social. Tan solo enmarca que la simulación acontece dentro de un juego de costes-beneficios percibidos, que la persona realiza en situaciones inciertas con el fin de lograr sus propios objetivos.

De acuerdo a este modelo, la conducta de simulación será, lógicamente, más probable en aquellos casos en que los incentivos sean mayores, el contexto sea más adverso y no haya otras vías alternativas disponibles o accesibles. Diversos estudios han encontrado estas relaciones entre los síntomas observados y las ganancias posibles (v.g. Binder y Rohling, 1996; Cassidy et al., 2000; Øverland et al., 2008; Robinson y Loeser, 2012). No es de extrañar, por tanto, que, tal y como afirman Echeburúa, Amor y Corral (2003) o Iverson (2010), las causas más frecuentes de simulación sean alegar irresponsabilidad penal por un delito cometido, rehuir obligaciones, conseguir indemnizaciones (p. ej., tras un accidente de coche) u obtener algún tipo de invalidez o discapacidad. Un análisis de casos de simuladores a partir de 220 expertos forenses, encontró esta evaluación de coste-beneficios en contextos adversos como prototípico de simuladores (Rogers et al., 1998).

En síntesis, podría afirmarse que el modelo adaptativo predice la presencia de simulación en la medida en que aparezcan tres condiciones:

- a) la persona percibe el contexto de evaluación como adverso;
- b) la persona considera que tiene algo que ganar si simula;
- c) la percibe no percibe otro medio más efectivo para conseguir la meta perseguida.

En la siguiente tabla pueden compararse las asunciones implícitas subyacentes a otros modelos teóricos existentes de simulación, como son el modelo patogénico o el modelo criminológico del DSM.

Tabla 1.1
Modelos explicativos de simulación

Modelo	¿Por qué simula una persona?	¿Cuáles son los factores que facilitan la presencia de simulación?	¿La simulación es un rasgo de la persona o un estado?
Patogénico	Como un medio para controlar una psicopatología grave y creciente. Pretende controlar dichos síntomas produciendo intencionalmente y bajo su control síntomas simulados.	La presencia de sintomatología grave en el eje I y II.	Mixto (rasgo y estado)
Criminológico	La simulación se enmarca en el contexto más general de la conducta antisocial y es una expresión específica de esta para conseguir ganancias no merecidas.	La presencia de antecedentes antisociales o problemas con la justicia.	Considera la simulación como un atributo estable o un rasgo de personas antisociales.
Adaptativo	Como una respuesta situacional que busca obtener el mejor resultado ante unas circunstancias adversas a sus intereses.	La magnitud de las consecuencias y la carencia de alternativas.	Considera la simulación como una respuesta situacional a un contexto específico (un estado).

Un aspecto clave del modelo adaptativo es que la persona considera que la simulación es el mejor medio para conseguir el objetivo perseguido, que, además, se considera correcto o avalado moralmente. Por ejemplo, un lesionado en un accidente de tráfico puede estar enfadado, tanto por los perjuicios que le ha causado, como por cómo ha sido tratado por los seguros y por el sistema, por lo que considerará adecuada la exageración de síntomas como justa respuesta a su ira y a su sensación de maltrato, con el fin de conseguir una generosa compensación. También puede que considere necesario “quejarse”, exagerar sus síntomas, para obtener la atención que requiere, dependiendo del sistema de creencias y valores predominantes en la persona. Sin duda, esto también tendrá relación con el nivel de desarrollo moral alcanzado por la persona (Walker y Frimer, 2011).

Recientemente, González Ordi et al. (2012) han propuesto una versión del modelo adaptativo de simulación que trata de integrar variables clave en su definición y servir de base conceptual para la generación de hipótesis y de modelo de trabajo para futuras investigaciones en el campo, que precisen el papel que juega cada componente. A continuación se incluye una representación de dicho modelo (figura 1.1.).

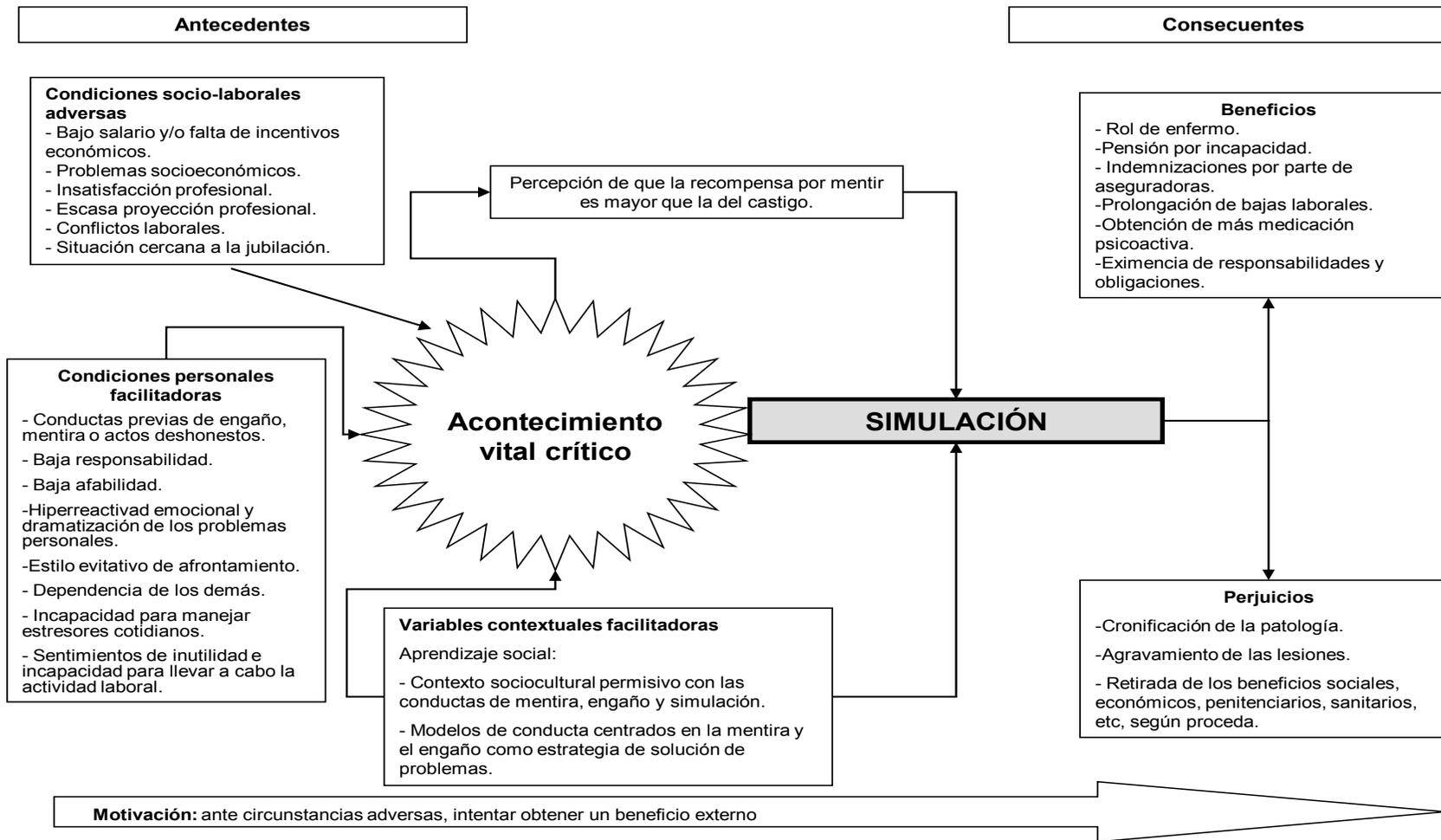


Figura 1.1. Propuesta de modelo adaptativo de la simulación como estilo de respuesta ante la enfermedad, tomado de González Ordi et al. (2012).

Tal y como se puede advertir en el modelo de la figura 1.1, la simulación, como toda conducta humana, tiene unos antecedentes que facilitan su aparición y unos consecuentes que refuerzan o extinguen su probabilidad de ocurrencia. Ante la aparición de un acontecimiento vital crítico (p.ej., un accidente, una lesión o trastorno o una situación de conflicto laboral o personal) en el que exista la posibilidad de obtener un beneficio material externo, la opción de simular cierta sintomatología discapacitante dependerá en buena medida de que el individuo perciba claramente que la recompensa por mentir es mayor que el castigo por hacerlo. Esta percepción puede estar mediatizada por diversas variables antecedentes, como son:

(1) Condiciones socio-laborales adversas: diversos estudios ponen de manifiesto que los presuntos simuladores suelen realizar trabajos percibidos como motivacionalmente insatisfactorios, por los que cobran un salario bajo y con escasos incentivos económicos y limitada o nula proyección profesional; asimismo, es probable que presenten problemas socioeconómicos, cargas familiares, conflictos laborales en su empresa o con otros empleados, o se encuentren en una situación cercana a la jubilación (Samuel y Mittenberg, 2005);

(2) Condiciones personales (intrapésicas) facilitadoras: aunque en este aspecto no existe clara investigación empírica disponible, algunos estudios forenses vinculan ciertos estilos de personalidad específicos con una mayor tendencia a la simulación (Mohino Justes et al., 2004; Wise, 2002). En esta línea, Meyers (2002) y Aguerrevere, Greve, Bianchini y Meyers (2008) observaron que los simuladores, además de puntuaciones elevadas en las escalas propias de exageración de síntomas, presentaban puntuaciones extremadamente bajas en la escala suplementaria Fuerza del yo (Es). Tal y como sugieren dichos autores, esto podría deberse a que ciertas características intrapésicas -como la baja responsabilidad, la baja afabilidad, la hiperreactividad, la dramatización de los problemas personales, el estilo evitativo de afrontamiento, la dependencia de los demás, la incapacidad para manejar estresores cotidianos y los sentimientos de inutilidad e incapacidad para llevar a cabo la actividad laboral- junto con la actitud favorable a la realización de conductas de engaño, mentira y actos deshonestos podrían resultar variables moduladoras significativas en la conducta de simulación.

Otras hipótesis han relacionado la simulación con una menor inteligencia y un peor rendimiento de las funciones ejecutivas. De esta manera ya O'Connell (1960) sugirió en los albores de la investigación en este campo que la baja inteligencia se asociaba a las alegaciones de amnesia tras la comisión de un crimen, argumentando que este tipo de estrategia podría ser una maniobra defensiva atractiva para gente con menos recursos cognitivos, dada su simplicidad y aparente escapatoria fácil a la condena. Más actualmente, el estudio de Cima, Merckelbach, Hollnack y Knauer

(2003), con presos que alegaban amnesia tras la comisión de un crimen, mostró una leve correlación entre las puntuaciones totales en el SIMS, como prueba de simulación de síntomas y una menor inteligencia (correlación con WAIS-R de -0,33), un menor rendimiento de la función ejecutiva (correlación con BADS de -0,37) y una mayor tendencia a la impulsividad (correlación con BIS de $r=0,25$).

(3) Variables contextuales facilitadoras: en virtud de las leyes del aprendizaje social, resultará más probable que un individuo simule si se ha educado en el contexto de modelos culturales permisivos con las conductas de engaño, mentira, ocultación y simulación y, más aún, si en su repertorio de aprendizaje han existido modelos de conducta centrados en la mentira y el engaño, como estrategia de solución de problemas personales. Es interesante citar aquí el estudio del Insurance Research Council (1993), que encontró que un 20% de los estadounidenses consideraban aceptable la exageración y falsificación de síntomas para obtener una determinada compensación o indemnización. Este porcentaje se elevaba hasta el 46% si se analizaban específicamente los estados situados más al noreste, donde los modelos de fraude o litigancia tienen más influencia.

Finalmente, es más probable que la simulación se produzca si la conducta no tiene coste inmediato de respuesta; esto es, si el simulador queda impune por su simulación. Como ocurre con toda conducta humana, el reforzamiento y el mantenimiento de la simulación en el tiempo tiene que ver con la percepción del individuo de que los beneficios superan a los perjuicios o, simplemente, con la percepción selectiva de los beneficios y no de los perjuicios potenciales que el mantenimiento de dicha conducta pueda acarrearle.

1.2. Ámbitos de simulación

Se puede considerar que existen tres grandes ámbitos o áreas de simulación: la simulación de trastornos mentales, la simulación de deterioro cognitivo y la simulación de síntomas físico-somáticos (Greene, 2008; Rogers, Sewell y Gillard, 2010). Cada uno de ellos tiene sus particularidades, tanto en los métodos de detección como en las propias características de la simulación (Cercy, Schretlen y Brandt, 1997).

El simulador de trastornos mentales debe centrar su atención en crear un conjunto de síntomas creíbles y coherentes, incluyendo su comienzo y desarrollo, su grado de consciencia sobre los mismos, así como el nivel de impacto en su vida diaria. Por el contrario, el simulador de trastornos cognitivos no se centra tanto en esta organización de síntomas, sino en mostrarse incapaz de resolver tareas cognitivas, teniendo que determinar en qué medida falla (con qué frecuencia y en

qué grado) y cómo le afecta a su vida diaria. Ante las tareas cognitivas, deben aparentar poner todo su esfuerzo en resolverlas y, aún así, fallarlas. Finalmente, la simulación de quejas médicas es más compleja que los otros dominios por la diversidad de formas en que puede manifestarse. Puede centrarse en un único síntoma incapacitante (p. ej., dolor), puede retratar una constelación de síntomas relativamente comunes que causan malestar (p. ej., dolor de cabeza, fatigas y problemas gastrointestinales) o especializarse en síndromes complejos (p. ej., fibromialgia).

Aunque frecuentemente los simuladores entremezclan las áreas que simulan (p. ej., puede simular dolor junto con trastornos afectivos y alteraciones cognitivas), es importante destacar que, dadas las diferencias entre los dominios, las escalas de detección son relativamente específicas a cada uno de ellos (Nelson, Sweet, Berry, Bryant y Granacher, 2007). Esto es, los resultados negativos en un test de simulación psicopatológica no descartarán la presencia de simulación en las áreas cognitiva o médica o, a la inversa, su presencia no es reflejo de que el individuo esté simulando en todas ellas (Berry y Granacher, 2009).

Los avances en el área de la detección de simulación de trastornos mentales y trastornos cognitivos han sido notables en los últimos años. No obstante, los progresos en el ámbito de la detección de la simulación de problemas médicos han sido sustancialmente menores, siendo un área que generalmente ha sido obviada en la investigación (Berry y Nelson, 2010; Rogers et al., 1998) y considerada el próximo reto en este campo (Rogers y Correa, 2008).

1.3. Prevalencia

La estimación de tasas de prevalencia en simulación es un asunto particularmente complejo y de difícil solución. En primer lugar, porque los simuladores buscan activamente no ser detectados, lo que dificulta o imposibilita de facto llegar a conocer su prevalencia real; en segundo lugar, porque solo recientemente se han establecido criterios diagnósticos claros que definan operativamente qué es la simulación (Bush et al., 2005).

A estas dos serias dificultades, hay que añadir la notable influencia del contexto de evaluación en la prevalencia de la simulación, haciendo prácticamente inviable hablar de tasas de prevalencia de simulación generales, estables o consistentes en distintos escenarios. Por el propio modelo adaptativo que pretende explicar la simulación, su prevalencia varía según el motivo de evaluación, el trastorno o el contexto (civil o penal), como se ha hallado en distintos estudios (v.g., Rogers et al., 1998)

Estas tasas también se ven afectadas por el tipo de evaluación, siendo sensiblemente diferentes en evaluaciones rutinarias y sistemáticas de los casos que en evaluaciones de casos remitidos por sospecha de simulación. Los estudios en que la evaluación ha sido aplicada sistemáticamente, obtienen unas tasas de prevalencia que pueden oscilar entre un 10 y un 30%, si bien cuando la evaluación se hace a sujetos que han sido remitidos por sospecha de simulación la tasa de prevalencia probablemente exceda el 50% (Greene, 2008).

A pesar de estas dificultades y fuentes de variación, es fundamental poder disponer de ciertas estimaciones de prevalencia ya que resultarán necesarias no solo para conocer su importancia y relevancia social sino para informar sobre las decisiones clínicas y calcular los índices de utilidad diagnóstica en las investigaciones sobre indicadores de simulación que se realicen (Meehl y Rosen, 1955).

Los estudios llevados a cabo hasta la fecha para estimar la prevalencia de simulación han seguido dos estrategias: a) la consulta mediante encuestas estructuradas a jueces expertos, es decir, profesionales cualificados de la salud que habitualmente se encuentran con esta problemática en su desempeño profesional y b) la revisión de historias clínicas archivadas de individuos a los que se ha evaluado sistemáticamente la presencia de simulación.

La literatura científica disponible, abrumadoramente norteamericana, sugiere tasas de alrededor del 7% en la práctica clínica general (Dandachi-FitzGerald, Ponds y Merten, 2013; Dandachi-FitzGerald, Ponds, Peters y Merckelbach, 2011; Mittenberg et al., 2002; Rogers et al., 1998, 1994; Sharland y Gfeller, 2007) y de entre un 14 y un 30% en casos de solicitud de discapacidad o indemnización laboral (Dandachi-FitzGerald et al., 2013; Gervais, Rohling, Green y Ford, 2004; Miller, 2000; Mittenberg et al., 2002; Rogers et al., 1998, 1994; Sharland y Gfeller, 2007; Van Hout, Schmand, Wekking y Deelman, 2006). Algunos autores han elevado estas tasas hasta el rango de entre el 30% y el 50% en contextos forenses (Ardolf, Denney y Houston, 2007; Chafetz, 2008; Larrabee, 2003a; Norris y May, 1998; Walters, White y Greene, 1988), siendo particularmente más altas en contextos forenses criminales en juicios o en prisión (McDermott, Dualan y Scott, 2013). En el campo de simulación del dolor, los estudios sugieren tasas entre el 20% y el 50%, dependiendo del método de diagnóstico elegido (Brooks, Johnson-Greene, Lattie y FERENCE, 2012; Gervais et al., 2001; Greiffenstein, Gervais, Baker, Artiola y Smith, 2013; Greve, Ord, Bianchini y Curtis, 2009; Kay y Morris-Jones, 1998; Leavitt y Sweet, 1986), que son consistentemente superiores a las estimaciones iniciales de Fishbain, Cutler, Rosomoff y Rosomoff (1999).

Los escasos estudios realizados fuera de Norteamérica todavía son incipientes para determinar su coincidencia con los estudios americanos, si bien suelen apuntar tasas de simulación iguales o superiores. Así, por ejemplo, la investigación realizada en Alemania por Stevens, Friedel, Mehren y Merten (2008), con solicitantes de indemnizaciones por lesiones, reflejó tasas de simulación del 44,6% y, en cuanto al tipo de simulación, un estudio previo de dichos autores estimó un 44% de simulación de problemas cognitivos y un 51% de simulación de trastorno por estrés postraumático en pacientes tras un accidente (Merten, Friedel y Stevens, 2006). El estudio de Plohmann (2013) en Suiza, realizado a partir de una muestra de solicitantes de discapacidad o compensaciones por lesiones, presentó resultados de sospecha de simulación entre 23% y el 54% de los casos, dependiendo del protocolo de evaluación llevado a cabo. En Inglaterra, Moss, Jones, Fokias y Quinn (2003) hallaron una prevalencia del 33% de simulación en personas evaluadas por traumatismos craneoencefálicos. En Holanda, Schmand et al. (1998) obtuvo tasas de sospecha de simulación del 61% de los litigantes tras un esguince cervical y de un 29% en no litigantes que acudían a consulta clínica externa, mientras que Van Hout et al. (2006) estimó la presencia de simulación de problemas cognitivos en personas que alegaban encefalopatía tóxica crónica en un 26%. Finalmente, el estudio realizado en Brasil por Jorge y Jorge (2011) determinó unas tasas de simulación de lumbalgia del 78% en solicitantes de incapacidad.

No obstante, de todos los estudios de prevalencia realizados en el campo médico-legal hasta la fecha, el más completo es el llevado a cabo en Estados Unidos por Mittenberg et al. (2002), que obtuvo tasas estimadas de probable simulación y exageración de síntomas a partir de un total de 22.121 casos en contexto médico, 6.371 casos en evaluaciones para juicios por lesiones personales, 3.688 casos en evaluaciones para determinar el grado de discapacidad asignable y 1.341 casos evaluados en contexto criminal. A continuación, se pueden observar las tasas estimadas según el contexto y patología.

Tabla 1.2
Tasas de simulación por contexto en el estudio de Mittenberg et al. (2002)

Por contexto	%
Consulta de medicina general o psiquiatría	8
Casos criminales	19
Lesiones personales	29
Discapacidad	30

Cuando se analizan los resultados de prevalencia de simulación por patología, tal y como se expone en la tabla 1.3, puede observarse que los traumatismos craneoencefálicos leves (TCE leve)

encabezan la lista, seguidos de los trastornos asociados al dolor crónico (fibromialgia, fatiga crónica, dolor o trastornos somatoformes) Estas tasas de simulación del dolor, con valores que oscilan entre el 30% y el 35%, coinciden con las obtenidas en otros estudios como el de Gervais et al. (2001) con pacientes con fibromialgia. Respecto a la prevalencia de simulación de trastornos de carácter psicopatológico (depresión, ansiedad y trastornos disociativos), las tasas estimadas oscilan entre el 10% y el 15%. Estos tres grupos de trastornos, trastornos cognitivos (TCE), trastornos físicos o médicos (dolor, neurotóxicos, electricidad) y trastornos psicopatológicos (ansiedad, depresión, trastorno disociativo), constituyen, según este estudio, los trastornos simulados más habitualmente.

Tabla 1.3
Tasas de simulación por patologías en el estudio de Mittenberg et al. (2002)

Por patologías	%
TCE leve	38,5
Fibromialgia o fatiga crónica	34,7
Dolor o trastornos somatoformes	31,4
Trastornos neurotóxicos	26,5
Lesión por accidente con electricidad	21,9
Trastornos por depresión	14,9
Trastornos de ansiedad	13,5
Trastornos disociativos	10,5
Trastorno con base epiléptica	9,3
TCE moderado o severo	8,8
Demencia vascular	2,4

Nota. TCE : Traumatismo Craneoencefálico

El conjunto de los estudios realizados en el campo médico-legal muestra, por lo tanto, la relevancia e importancia que la simulación tiene en esta área. No obstante, cabe señalar que el impacto de la exageración y simulación de sintomatología afecta también a la práctica clínica cotidiana de los profesionales, incluso en contextos sin aparente presencia de simulación. A este respecto cabe mencionar el estudio holandés de Van Egmond, Kummeling y Balkom (2005) con una muestra de pacientes en tratamiento y sin incentivos aparentes. Los resultados evidenciaron que el 42% de los pacientes de consultas clínicas ordinarias albergaban expectativas de beneficios específicos diferentes de la propia mejoría (p. ej., conseguir ayuda para obtener la condición de discapacidad, una baja laboral o para tener acceso a otro alojamiento). Lo más reseñable de este

estudio es que la mayoría de los clínicos eran ajenos a dichas expectativas, siendo el grupo de pacientes que albergaba expectativas de beneficios el que mostró un peor pronóstico y una menor mejoría, en comparación con el grupo de pacientes que no albergaba dichas expectativas. Este estudio resalta como la presencia de exageración o simulación de sintomatología no solo afecta a los aspectos económicos, legales o sociales de la evaluación y el diagnóstico sino también a la propia eficacia de la intervención y de la práctica profesional.

1.4. Estrategias de detección de simulación

En un escenario en el que la detección de la simulación se ha realizado tradicionalmente en función de criterios clínicos subjetivos y poco específicos (respaldados por la falsa creencia en el “ojo clínico” y en la experiencia), se hace necesario establecer una lista sistemática y empíricamente verificable de qué características ha de cumplir una estrategia para poder ser considerada como de detección de la simulación.

En este marco de trabajo, una estrategia de detección puede ser definida como un método estandarizado, fundamentado conceptualmente y validado empíricamente, que permite la diferenciación sistemática de un estilo específico de respuesta (p. ej., exageración de síntomas) frente a otros (p. ej., respuesta honesta) (Rogers et al., 2010).

De esta conceptualización se desprenden los siguientes requisitos que debe cumplir una estrategia de detección de la simulación para considerarse válida:

1. *Método estandarizado*: ¿Está suficientemente operativizada la estrategia como para ser aplicada y corregida de forma fiable y sistemática? La aplicación o valoración subjetiva de reglas no detalladas carece del requerido rigor científico.
2. *Fundamentada conceptualmente*: ¿Está respaldada por un marco teórico o conceptual en el que se incluye? La justificación de la estrategia dentro de un marco teórico o conceptual es requisito indispensable para la construcción y desarrollo de un adecuado corpus científico.
3. *Validada empíricamente*: ¿Ha sido validada empíricamente usando la metodología establecida? Tal y como se expondrá en el capítulo referido a metodología, los métodos establecidos incluyen distintos tipos de diseños.
4. *Diferenciación sistemática*: Cuando se emplea la estrategia, ¿cuál es la magnitud de la diferencia? La existencia de diferencias estadísticamente significativas no es suficiente por sí misma, es requisito que sean relevantes clínicamente para lo cual su tamaño del

efecto debe ser moderado o grande (de acuerdo a los criterios que se expondrán en el capítulo referido a metodología).

5. *Estilo específico de respuesta*: ¿Ha sido validada la estrategia para un estilo de respuesta específico usando grupos clínicos relevantes de comparación? Es absolutamente indispensable que se haya validado empleando grupos clínicos de comparación que permitan determinar las diferencias entre los clínicos genuinos y los simuladores de trastornos mentales. Incluir comparaciones de simuladores frente a población normal no permite establecer ninguna diferenciación relevante ya que no posibilita investigar si la técnica diferencia entre los grupos realmente importantes, es decir, los pacientes reales frente a los pacientes simuladores.

En líneas generales, las estrategias que se han ido generando para la detección de simulación de trastornos mentales y de trastornos médicos pueden agruparse en dos grandes categorías: aquellas basadas en la inverosimilitud de los síntomas y aquellas basadas en la amplificación de los síntomas informados (Rogers et al., 2010).

La primera categoría, **inverosimilitud**, agrupa aquellas estrategias que presentan síntomas improbables o falsos que los simuladores erróneamente se imputan a sí mismos considerándolos propios del trastorno simulado. El punto clave en esta categoría es que los pacientes clínicos genuinos apenas afirmarán caracterizarse por ninguno de estos ítems. Esta diferencia será la que permitirá detectar el patrón de exageración de síntomas con estas estrategias.

La segunda categoría, **amplificación de síntomas**, agrupa aquellas estrategias que presentan síntomas que, si bien pueden darse en la población clínica, no aparecen en la clínica real con la severidad, frecuencia o intensidad con que se describen. En este caso, son ítems plausibles y veraces pero no lo son ni en su intensidad, ni en su frecuencia, ni en sus características asociadas. El punto clave en esta categoría es que haya diferencias relevantes entre el grado afirmado por los pacientes clínicos genuinos y los simuladores. Esta discrepancia será la que permitirá discriminar entre ambos casos.

A continuación se desglosan las estrategias específicas que conforman ambas categorías, tanto para la simulación de trastornos mentales como de trastornos físicos o médicos. Como podrá observarse, ambas guardan un gran paralelismo.

Tabla 1.4
Estrategias de detección de exageración de síntomas en trastornos mentales

Estrategia	Descripción	Puntos fuertes	Limitaciones
Síntomas infrecuentes	Presenta síntomas y otros indicadores de psicopatología que son muy infrecuentes (<5%) en poblaciones clínicas genuinas. Debe diferenciarse de síntomas “casi infrecuentes” en que, en este caso, deben ser síntomas que también sean muy infrecuentes en poblaciones clínicas.	Esta estrategia ha sido ampliamente aplicada en distintas pruebas. Suele proporcionar tamaños del efecto grandes o muy grandes.	
Síntomas casi-infrecuentes	Presenta síntomas y otros indicadores de psicopatología que son muy infrecuentes (<5%) en la población general. No obstante, son síntomas que pueden ser relativamente frecuentes en ciertas poblaciones clínicas. Por este motivo, es considerada una “cuasi” estrategia ya que los ítems infrecuentes pueden reflejar tanto trastornos genuinos como simulados.	Suele proporcionar tamaños del efecto grandes o muy grandes.	Dado que ciertos síntomas infrecuentes en población normal pueden ser habituales en poblaciones clínicas, los resultados obtenidos por esta estrategia pueden ser confusos (p. ej., los sujetos con esquizofrenia o trastorno por estrés postraumático suelen obtener habitualmente notables elevaciones en la escala F del MMPI-2 (Rogers, Sewell, Martin y Vitacco, 2003).
Combinación de síntomas	Emplea pares de síntomas y otros indicadores de psicopatología que pueden darse de modo aislado, pero muy raramente se presentan simultáneamente.	Esta estrategia es sofisticada y debería ser resistente al entrenamiento para evitar ser detectados como simuladores. Es fácilmente adaptable tanto a entrevistas estructuradas como a inventarios multiescalares. Suele producir tamaños de efecto grandes.	

Tabla 1.4 (continuación)

Estrategia	Descripción	Puntos fuertes	Limitaciones
Síntomas improbables	Es una variante extrema de la estrategia de “síntomas infrecuentes” ya que describe síntomas fantásticos que por sus características son prácticamente imposibles que sean auténticos.	Dada su naturaleza fantástica resulta prácticamente imposible que estos ítems sean ciertos en una persona. Por lo tanto, a diferencias de otras estrategias, en este caso, existen pocas explicaciones alternativas cuando se han marcado múltiples ítems de esta estrategia.	Su carácter fantástico limita su utilidad con simuladores más sofisticados que pueden identificar la inverosimilitud de que sean síntomas clínicos genuinos.
Patrones espurios de psicopatología	Va un paso más allá de la combinación de síntomas ya que esta estrategia supone un elaborado enfoque que identifica aquellas combinaciones de escalas y subescalas que son muy poco frecuentes en poblaciones clínicas genuinas.	Su complejidad minimiza la posibilidad de que los simuladores puedan ser entrenados para evitar su detección.	Dada su complejidad, requiere una amplia y extensa validación cruzada que garantice su validez.
Estereotipos erróneos	Presenta estereotipos que erróneamente se consideran características típicas de los trastornos mentales y presentes en los sujetos clínicos.	Parece resistente al entrenamiento ya que hasta los propios profesionales de la salud suelen tener estos estereotipos.	
Asignación indiscriminada de síntomas	Considera la proporción global de síntomas alegados y es útil para detectar a aquellos simuladores que siguen la consigna de “cuanto más mejor”. El reconocimiento masivo de múltiples y diversos síntomas no es frecuente en la población clínica genuina.	La proporción de síntomas afirmados es fácil de calcular y puede obtenerse en casi todas las medidas psicológicas.	Solo puede emplearse en la medida en que cubra un amplio rango de síntomas, de lo contrario podría llevar fácilmente a falsos positivos.

Tabla 1.4 (continuación)

Estrategia	Descripción	Puntos fuertes	Limitaciones
Severidad de los síntomas	Tiene en cuenta el grado de severidad de los síntomas para diferenciar entre exageración de síntomas y pacientes clínicos genuinos.	Es fácilmente adaptable a entrevistas estructuradas y escalas clínicas. Suele producir tamaños del efecto grandes.	
Síntomas obvios	Hace uso de síntomas que son fácilmente reconocidos como característicos de los principales trastornos mentales. Aquellos que exageran su deterioro suelen superar notablemente el número de estos síntomas que afirman en comparación con poblaciones clínicas diversas.	Suele producir tamaños del efecto grandes o muy grandes.	No está claro si es mejor utilizar esta estrategia por separado o combinarla con la detección de problemas más sutiles.
Síntomas observados frente a los informados	Analiza la inconsistencia entre el grado de deterioro informado y el observado clínicamente.	Con observaciones estandarizadas, esta estrategia proporciona una verificación independiente de los síntomas afirmados.	La estandarización de la observación es esencial para la precisa discriminación.

Tabla 1.5

Estrategias de detección de síntomas de trastornos mentales que requieren validación adicional

Estrategia	Descripción	Puntos fuertes	Limitaciones
Sobreespecificación de los síntomas	Descansa en la asunción de que los simuladores están más dispuestos a afirmar que padecen un síntoma que esté descrito con un nivel inverosímil de precisión y detalle.	Produce tamaños del efecto moderados.	A diferencia del resto de estrategias, carece de una base conceptual.
Aproximación a síntomas genuinos	Presenta síntomas falsos pero que son similares en apariencia a los síntomas genuinos, a excepción de algún detalle importante.		Los pacientes genuinos pueden fijarse en lo esencial del ítem y ser erróneamente clasificados como simuladores.

Tabla 1.6
Estrategias de detección de exageración de síntomas en trastornos físicos o médicos

Estrategia	Descripción	Puntos fuertes	Limitaciones
Quejas médicas o síntomas infrecuentes	Presenta síntomas y quejas médicas que son muy infrecuentes (<5%) en poblaciones clínicas genuinas.	Esta estrategia es actualmente la que más atención ha recibido en investigación y la que tiene más medidas que se han basado en ella.	
Quejas médicas o síntomas improbables	Es una variante extrema de la estrategia de “quejas médicas o síntomas infrecuentes” ya que describe síntomas fantásticos y que por sus características es prácticamente imposible que sean auténticos.	Dada su naturaleza fantástica resulta prácticamente imposible que estos síntomas o quejas sean ciertos en una persona. Por lo tanto, a diferencia de otras estrategias, en este caso, existen pocas explicaciones alternativas cuando se han marcado múltiples síntomas de esta estrategia.	Su carácter fantástico limita su utilidad con simuladores más sofisticados que pueden identificar la inverosimilitud de que sean síntomas clínicos genuinos.
Combinación de síntomas	Emplea pares de síntomas que pueden aparecer con frecuencia en poblaciones clínicas genuinas pero que muy raramente se presentan simultáneamente. Los simuladores suelen tener información sobre la prevalencia de los síntomas individualmente, pero desconocen que es poco probable que co-ocurrán simultáneamente.	Esta estrategia es sofisticada y debería ser resistente al entrenamiento para evitar ser detectados como simuladores. Es fácilmente adaptable tanto a entrevistas estructuradas como a inventarios multiescalares.	
Asignación indiscriminada de síntomas	Considera la proporción global de síntomas alegados y es útil para detectar a aquellos simuladores que siguen la consigna de “cuanto más mejor”. El reconocimiento al por mayor de múltiples y diversos síntomas no es frecuente en la población clínica genuina.	Si se descarta la presencia de enfermedades estructurales (p.ej., lupus), la asignación indiscriminada de síntomas puede proporcionar una excelente discriminación entre pacientes genuinos y simuladores.	

Tabla 1.6 (continuación)

Estrategia	Descripción	Puntos fuertes	Limitaciones
Intensidad de los síntomas	Esta estrategia se basa en la observación de que los simuladores de trastornos físicos suelen exagerar la frecuencia, duración o gravedad de sus síntomas o problemas.	Combina varios parámetros (p. ej., frecuencia e intensidad) para crear índices de detección compuestos. Algunos estudios han mostrado tamaños del efecto grandes con esta estrategia.	
Síntomas observados frente a los informados	Analiza la inconsistencia entre las quejas y síntomas físicos informados por el paciente y los observados. En el caso de los simuladores suele darse una divergencia sistemática entre ambas fuentes de información.	Con observaciones estandarizadas, esta estrategia proporciona una verificación independiente de los síntomas afirmados.	La estandarización de la observación es esencial para una discriminación precisa.
Dependencia de las quejas médicas	Esta estrategia está basada en la idea, aparentemente contra intuitiva, de que los simuladores estarán más abiertos a reconocer los aspectos positivos asociados a su lesión o trastorno.		Escasa investigación que la respalde.
Virtuosismo	Esta estrategia descansa en la tendencia por parte de los simuladores a evitar ser descubiertos presentándose como extremadamente virtuosos, negando la presencia de pequeños defectos personales frecuentes y asignándose cualidades ideales a nivel personal y social.	Combinada con otras estrategias puede aumentar la capacidad discriminativa entre pacientes genuinos	

Los escasos estudios realizados con simuladores análogos sobre las estrategias que han creído percibir durante la evaluación para detectarles como simuladores, han indicado que suelen considerar la inconsistencia como principal vía percibida de detección. Esto es, consideran que la manera en que se va a evaluar si son sinceros o están simulando es preguntándoles de formas diversas lo mismo para ver si son consistentes. Así, el simulador análogo se centra en resultar consistente en sus respuestas y percibe que la diferencia marcada en sus respuestas será una de las estrategias para detectar su engaño (Gillard, 2010; Rogers et al., 2010), cuando realmente esta no es una de las estrategias relevantes empleadas.

Es interesante señalar cómo las estrategias de detección de simulación mencionadas en este apartado suelen centrarse en valorar la verosimilitud o credibilidad de los síntomas presentes, de la información que nos proporciona directamente el paciente y que debemos contrastar. No obstante, cabría plantearse como vía complementaria de investigación el análisis de los síntomas ausentes, aquellos que no han sido mencionados por el paciente y que, sin embargo, pueden ser prototípicos y característicos del cuadro patológico. Esta estrategia complementaria de síntomas negativos o ausentes apenas ha sido investigada hasta la actualidad, debido a dos dificultades principales: 1) cómo determinar qué “ausencias” resultan indicativas de simulación y cuáles son compatibles con la presentación genuina del trastorno ya que sabemos que los trastornos son constelaciones de síntomas que no suelen cumplirse en su integridad; 2) cómo indagar en la “ausencia” de estos síntomas sin provocar la tendencia del simulador a asignarse dichos síntomas ya que sabemos que es necesario confirmar dicha ausencia (y evitar que sea debida en realidad a una omisión en el relato del paciente) pero es difícil hacerlo sin causar la aquiescencia del evaluado. Estas limitaciones hacen que todavía sea una línea de investigación extremadamente incipiente que requerirá más investigación en las próximas décadas antes de formar parte del arsenal de estrategias establecidas para la detección de simulación.

1.5. Criterios diagnósticos de simulación

Como respuesta a las limitaciones de los criterios unidimensionales, valorativos y subjetivistas del DSM en la evaluación de la simulación, se han propuesto modelos alternativos que enfatizan específicamente el examen detallado de los datos de las pruebas y la concurrencia de la información antecedente disponible.

Tal y como se señalaba en el punto anterior, la compleja naturaleza del diagnóstico diferencial de simulación ha hecho que sistemáticamente durante los últimos quince años los expertos hayan recomendado, una y otra vez, el necesario empleo convergente de múltiples métodos y estrategias para su correcta evaluación y diagnóstico (Bender y Rogers, 2004, 2004; Capilla Ramírez y González-Ordi, 2009; González-Ordi y Gancedo Rojí, 1999; Larrabee, Greiffenstein, Greve y Bianchini, 2007; Matalobos Veiga y González-Ordi, 2007; Nies y Sweet, 1994; Slick, Sherman y Iverson, 1999; Sweet, 2008; Youngjohn, Spector y Mapou, 1998).

Así, el diagnóstico de simulación se fundamenta en la idea central de la convergencia de los datos que provienen de distintos métodos, fuentes y estrategias: entrevista, historia clínica, pruebas médicas, pruebas psicológicas, evaluación conductual, etc. La mejor práctica clínica implica el uso de criterios diagnósticos múltiples en oposición a la utilización de un único test, así como la utilización de enfoques evaluativos que incluyan la combinación de tests especializados e indicadores de validez (González-Ordi et al., 2012; Kane, 2010).

De esta manera, se busca incrementar la robustez del diagnóstico de simulación incorporando múltiples fuentes de información y conceptualizando la simulación como algo dimensional, más que categórico, que abarca desde la ausencia de sospecha de simulación, a la clara y definitiva existencia de simulación, pasando por los puntos intermedios de posible simulación y probable simulación (Wygant et al., 2009).

En este marco de trabajo, los dos sistemas diagnósticos más empleados para el diagnóstico de simulación son los desarrollados por Slick et al. (1999) para la simulación de déficit neuropsicológicos (ámbito cognitivo de simulación) y los adaptados por Bianchini, Greve y Glynn (2005) para la simulación de dolor. Ambos coinciden en abordar múltiples criterios procedentes de diversas fuentes de información para el diagnóstico de simulación, así como , en requerir la existencia de incentivos externos significativos, en la exclusión de factores psicopatológicos, neurológicos o evolutivos que expliquen los resultados y, por último, en clasificar en gradientes del grado de sospecha de simulación (posible, probable, definitiva).

1.5.1. Criterios diagnósticos de Slick et al. (1999)

Slick et al. (1999) adaptaron el enfoque empleado como referencia para el diagnóstico de la enfermedad de Alzheimer y lo aplicaron a la simulación. Al igual que en el diagnóstico de Alzheimer, consideraron que en la simulación no hay evidencias definitivas y concluyentes que establezcan un

diagnóstico, sino que debe hacerse uso de múltiples fuentes de información y determinar distintos niveles de seguridad en el diagnóstico establecido.

Estos autores establecieron tres niveles de seguridad en el diagnóstico (posible, probable y definitivo o firme), a partir de cuatro criterios diagnósticos que consideran múltiples fuentes de información y datos:

- (A) presencia de incentivos externos sustanciales,
- (B) sesgo en sus respuestas a las pruebas cognitivas,
- (C) exageración de síntomas informados o aportación de información no veraz y engañosa y
- (D) las conductas observadas en los criterios B y C no son explicadas completamente por otros factores psiquiátricos, neurológicos o de desarrollo.

A continuación se desglosa cada uno de los criterios:

- **Criterio A. Presencia de un importante incentivo externo**, el cual debe estar presente en el momento de la valoración (indemnizaciones por lesiones personales, obtención de pensión de invalidez, evasión de un posible enjuiciamiento penal...).

- **Criterios B. Resultados en las pruebas neuropsicológicas**. Debe existir una exageración o fabricación de disfunción cognitiva en el rendimiento en los tests neuropsicológicos, demostrado al menos por uno de los siguientes aspectos:

1. *Claro sesgo de respuesta negativa*. Una tasa de acierto o rendimiento por debajo del azar en una o más pruebas de función cognitiva con ítems de elección múltiple.

2. *Probable sesgo de respuesta*. Resultados consistentes con simulación en uno o más tests psicométricos validados, que evalúen la exageración o fabricación de déficits cognitivos.

3. *Discrepancia entre los resultados en los tests y los patrones conocidos de alteraciones cognitivas*. Un patrón de rendimiento en los tests neuropsicológicos que es altamente discrepante con el modelo aceptado del funcionamiento del sistema nervioso central normal y anormal. Por ejemplo, muestra una severa disfunción en las pruebas de atención pero un rendimiento normal en las pruebas de memoria.

4. *Discrepancia entre los resultados del test y el comportamiento observado*. El rendimiento en dos o más tests neuropsicológicos dentro de un dominio es discrepante con el nivel observado de funcionamiento del evaluado, de forma que sugiere una exageración o fabricación de síntomas (p. ej., resultados muy bajos en las pruebas de memoria en personas que no muestran ningún tipo de dificultad para contar los detalles de cómo sufrieron sus

lesiones, su historia médica y otros aspectos posteriores al accidente relacionados con su afectación).

5. *Discrepancia entre el rendimiento de las pruebas y los informes colaterales obtenidos.* El rendimiento en dos o más tests neuropsicológicos dentro de un dominio es discrepante con el nivel de funcionamiento cotidiano del evaluado descrito por, al menos, una fuente de información colateral fiable, de forma que sugiere una exageración o fabricación de síntomas (p. ej., un paciente que se ocupa de las finanzas de la familia, pero es incapaz de realizar cálculos matemáticos simples en las pruebas).

6. *Discrepancia entre el rendimiento de las pruebas y la historia documentada.* Inexplicablemente se presenta un bajo rendimiento en dos o más tests estandarizados de funcionamiento cognitivo, que es inconsistente con la historia neurológica o psiquiátrica documentadas (p. ej., un paciente sin historia de pérdida de consciencia o amnesia postraumática y sin antecedentes de otras alteraciones del sistema nervioso central o enfermedades, obtiene resultados similares a los que obtienen las personas que han estado más de un mes en coma).

• **Criterio C. Resultados de los autoinformes.** Los siguientes comportamientos son indicadores de una posible simulación de deterioro cognitivo, pero su presencia no es suficiente para el diagnóstico. Sin embargo, la presencia de uno o más de estos criterios aporta evidencia adicional para sustentar el diagnóstico de simulación. Estos criterios incluyen inconsistencias significativas o discrepancias en la información de síntomas que proporciona el paciente y que sugieren un intento deliberado por exagerar o fabricar déficits cognitivos.

1. *Discrepancia entre la historia informada por el sujeto y la historia documentada.* La historia reportada por el mismo paciente es claramente discrepante con la historia médica o psicosocial documentada y sugiere intentos por exagerar la severidad en la lesión o negar disfunciones neuropsicológicas premórbidas (p. ej., exagera la severidad de daño físico, la duración de pérdida de consciencia o el periodo de amnesia postraumática o niega una historia psiquiátrica u otra lesión cerebral previa).

2. *Discrepancia entre los síntomas informados por el sujeto y el patrón conocido de funcionamiento del cerebro* (p. ej., el paciente refiere tener una extensa amnesia sobre su propia biografía posterior a un trastorno craneoencefálico leve sin pérdida de consciencia).

3. *Discrepancia entre los síntomas informados por el sujeto y el comportamiento observado* (p. ej., afirma tener seriamente deteriorada su memoria, pero no presenta ninguna dificultad en narrar su historia clínica o los acontecimientos relacionados con el accidente, previos y posteriores).

4. *Discrepancia entre los síntomas informados por el sujeto y los informes colaterales obtenidos* (p. ej., afirma tener serios problemas de memoria, pero las entrevistas con otros informantes fiables no indican que hayan disfunciones de memoria en su casa).

5. *Hay evidencia de exageración o fabricación de alteraciones psicológicas*. Las alteraciones informadas por el propio paciente discrepan sustancialmente con los comportamientos observados, con la información colateral fiable o con las escalas de validez de autoinformes (p.ej., MMPI-2), que claramente sugieren la exageración o fabricación de síntomas en el autoinforme.

- **Criterio D.** Los comportamientos necesarios de los criterios B y C no son mejor explicados completamente por factores psiquiátricos, neurológicos o del desarrollo, que pueden producir una disminución significativa de la capacidad para considerar las normas (tal como la necesidad psicológica para desempeñar el papel de enfermo o en respuesta a una alucinación). Los comportamientos necesarios de los criterios B y C son productos de un esfuerzo racional y motivado con el objetivo de, al menos, adquirir u obtener un incentivo externo, según lo define el criterio A.

Tabla 1.7

Criterios de diagnóstico de simulación de deterioro cognitivo de Slick et al. (1999)

Criterio	Indicadores
A	A. Claro y sustancial incentivo externo.
B. Pruebas cognitivas	B1. Claro sesgo de respuesta negativa (por debajo del nivel de azar). B2. Probable sesgo de respuesta. B3. Discrepancia entre los resultados en los tests y los patrones conocidos de alteraciones cognitivas. B4. Discrepancia entre los resultados del test y el comportamiento observado. B5. Discrepancia entre el rendimiento de las pruebas y los informes colaterales obtenidos. B6. Discrepancia entre el rendimiento de las pruebas y la historia documentada.
C. Autoinforme	C1. Discrepancia entre la historia informada por el sujeto y la historia documentada. C2. Discrepancia entre los síntomas informados por el sujeto y el patrón conocido de funcionamiento del cerebro. C3. Discrepancia entre los síntomas informados por el sujeto y el comportamiento observado. C4. Discrepancia entre los síntomas informados por el sujeto y los informes colaterales obtenidos. C5. Hay evidencia de exageración o fabricación de alteraciones psicológicas.
D. Exclusión	Los comportamientos necesarios de los criterios B y C no son mejor explicados completamente por factores psiquiátricos, neurológicos o del desarrollo.

De acuerdo a estas premisas, el **diagnóstico de simulación** se considerará:

- *Definitivo o firme*: Si se cumplen los criterios A + B1 + D.
- *Probable*:
 - Si se cumplen los criterios A + dos o más criterios entre B2 y B6 + D ó
 - Si se cumplen los criterios A + uno o más criterios entre B2 y B6 + uno o más criterios de C1 a C5 + D.
- *Posible*:
 - Si se cumplen los criterios A + uno o más criterios de C1 a C5 + D.
 - Si se dan las condiciones para el diagnóstico definitivo o probable pero no se cumple el criterio D (no pueden descartarse que sean otras condiciones, psiquiátricas, neurológicas o del desarrollo, las que expliquen dichas conductas).

1.5.2. Criterios diagnósticos de Bianchini et al. (2005)

Bianchini et al. (2005) propusieron una serie de estándares diagnósticos referidos a la simulación de dolor inspirándose en el modelo previo de Slick et al., (1999). A continuación se desglosan los criterios determinados por estos autores:

- **Criterio A. Evidencia de incentivos externos significativos contingentes con un patrón de exageración o fabricación de síntomas** (estatus de litigancia, pensión por discapacidad, indemnizaciones por parte de aseguradoras, prolongación de bajas laborales, obtención de más medicación psicoactiva, etc.).
- **Criterio B. Evidencia procedente de la evaluación física.** Evidencia de que las habilidades, capacidades o limitaciones demostradas a lo largo de una evaluación física formal son inconsistentes con la exageración de las quejas y síntomas somáticos y la disminución exagerada de la capacidad física. Este aspecto puede operativizarse mediante la detección de:
 1. *Sesgos en el rendimiento físico*: realización de pruebas de rendimiento que evidencian la existencia de una disminución de la capacidad física que no resulta concordante con la patología diagnosticada;
 2. *Discrepancia entre la valoración subjetiva de dolor (p. ej., EVA) y la reactividad psicofisiológica concomitante*;
 3. *Los síntomas o las quejas informadas por el paciente son sustancialmente divergentes con la evidencia patofisiológica existente*;

4. *Discrepancia entre la actividad física mostrada por el paciente durante la evaluación formal con la actividad física demostrada cuando el individuo no sabe que está siendo evaluado.*

• **Criterio C. Evidencia procedente de la evaluación neuropsicológica.** Evidencia de que los déficits neurocognitivos (capacidad de concentración, percepción, memoria, etc.) demostrados a lo largo de una evaluación formal son consistentes con la exageración o el falseamiento de la discapacidad cognitiva. Este criterio se operativiza en:

1. *Rendimiento por debajo del azar en pruebas de elección forzada que evalúan funciones cognitivas o perceptivas;*

2. *Rendimiento consistente con la exageración de la discapacidad en pruebas que evalúan funcionamiento cognitivo o perceptivo;*

3. *Discrepancias entre los resultados de las pruebas neuropsicológicas y el conocimiento existente del funcionamiento cerebral;*

4. *Discrepancias entre los resultados de las pruebas neuropsicológicas y la conducta observable y objetivable del paciente.*

• **Criterio D. Evidencia procedente de la evaluación mediante autoinforme.** Evidencia que los síntomas, quejas y limitaciones autoinformadas por el paciente son consistentes con un patrón de exageración de síntomas y discapacidad a nivel físico, cognitivo o emocional. Este criterio puede operativizarse a través de:

1. *La discrepancia de la conducta del paciente cuando es formalmente evaluado frente a cuando no sabe que está siendo evaluado,*

2. *Las quejas autoinformadas no son consistentes con la evolución de la historia clínica del paciente;*

3. *Los síntomas autoinformados por el paciente son discrepantes con los patrones conocidos de funcionamiento fisiológico y neurológico del dolor y la discapacidad referida a la patología en cuestión;*

4. *Los síntomas autoinformados por el paciente son discrepantes con la conducta observable y objetivable;*

5. *Presentación de un patrón de exageración de síntomas en pruebas de personalidad de amplio espectro y pruebas de screening de simulación de síntomas.*

• **Criterio E. Los criterios mencionados no se explican mejor por factores psicopatológicos, neurológicos o evolutivos.** La simulación se entiende como un patrón de conductas de carácter intencional y dirigido a la obtención de un beneficio externo, generalmente de tipo económico.

El **diagnóstico de simulación** se considerará:

- *Definitivo o firme*: Si se cumplen la mayoría de los criterios y subcriterios A, B, C, D y E.
- *Probable*: Si se cumplen los criterios A + dos o más subcriterios de B, C o D + E.
- *Posible*: Si se cumple el criterio A + un subcriterio de B, C o D. No es necesario que se cumpla E.

1.5.3. Limitaciones y mejoras sugeridas a los criterios diagnósticos de simulación

Si bien la inclusión de estos criterios ha supuesto un avance definitivo en el campo, no han faltado las críticas que han recibido, muchas de ellas sólidamente fundamentadas y que apuntan a una próxima tercera generación de criterios diagnósticos de simulación. Entre estas críticas caben destacar las aportaciones de Larrabee et al. (2007), Boone (2007b) y Rogers, Bender y Johnson (2011a).

El equipo de Larrabee et al. (2007) realizó una revisión de los criterios de simulación de Slick et al., (1999) a partir de toda la investigación y práctica clínica que se había ido acumulando desde su publicación, alertando de las siguientes cinco limitaciones:

1. Los criterios considerados tienden a no prestar atención a la evidencia disponible sobre simulación en otras áreas, tales como la emocional o la física.
2. Ciertas combinaciones de variables pueden predecir simulación con un alto grado de especificidad. Sin embargo, en el algoritmo de decisión de diagnóstico de simulación no se consideran suficiente para el diagnóstico de probable simulación ya que se encuadran dentro de un mismo subcriterio B (habitualmente B2 o B6).
3. Los criterios presentan un exceso de confianza en los resultados de los tests cognitivos (particularmente en los resultados de B1, B2 y B6), mientras que otras fuentes de evidencias (p.ej., las incluidas dentro de C5), aunque potencialmente son indicadores robustos de simulación, son infravalorados en su contribución diagnóstica.
4. El criterio C, tal y como aparece definido, depende en gran medida del juicio clínico. La fiabilidad y precisión del juicio clínico sobre la incongruencia de un autoinforme de acuerdo a otras evidencias (comportamentales, perfil de desarrollo...) estará en función de la profundidad y amplitud de conocimiento del neuropsicólogo.

5. Los criterios podrían ser demasiado conservadores (p. ej., se requieren dos evidencias en el criterio B6 para ser considerado como evidencia de sospecha de simulación).

Para paliar estas limitaciones, el equipo de Larrabee et al. (2007) establece cuatro recomendaciones:

1. Permitir que se pueda proceder al diagnóstico definitivo o probable a partir de múltiples resultados en pruebas psicométricas; en concreto, que no tengan que considerarse únicamente los porcentajes de aciertos debajo del azar, sino que pueda establecerse a partir de la existencia de puntuaciones por debajo de los puntos de corte en diversos tests de validez de síntomas.
2. Requerir que en el criterio B6 solo tenga que haber discrepancia con una prueba y no con dos.
3. Requerir múltiples evidencias en los criterios cualitativos o subjetivos o disponer de definiciones más precisas de las inconsistencias requeridas para los criterios B4, B5, C1, C2, C3 y C4.
4. Dar al criterio C un peso equivalente al del criterio B.

Boone (2007b) también realiza una revisión de los criterios diagnósticos de Slick et al., (1999) centrándose en tres aspectos:

1. El concepto de intencionalidad. Considera que es extremadamente difícil, si no imposible e incluso innecesario, llegar a determinar si una determinada conducta responde a un intento consciente o inconsciente por parte del evaluado. Según esta autora, resulta más inteligente huir de esos falsos dilemas conceptuales para centrarse en la “credibilidad” o no de los síntomas referidos por el paciente, cuestión clave por la que se solicita la evaluación al profesional. De esta manera, recomienda alejarse de la etiqueta de “simulación” para reemplazarla progresivamente por “funcionamiento neurocognitivo no creíble”, en el caso de simulación de deterioro cognitivo, o de “funcionamiento no creíble” en un contexto más general. Esta crítica y sugerencia de cambio de enfoque ha sido posteriormente refrendada por diversos autores (Berry y Nelson, 2010; Iverson, 2010).
2. El criterio para la categoría diagnóstica de definitivo o firme. Boone considera que el requisito de respuesta por debajo del azar para esta categoría reduce dramáticamente la sensibilidad del criterio y excluye a aquellos simuladores más sofisticados que son

conscientes de que responder por debajo del azar es poco creíble. Propone, como criterio para esta alternativa, de forma convergente con Larrabee (2008), que el evaluado falle al menos tres tests de validez de síntomas escasamente correlacionados y con especificidades superiores al 90%, junto con evidencias comportamentales de sospecha de simulación.

3. La advertencia a los evaluados sobre la presencia de escalas para detectar simulación. Si bien el grupo de Slick sugiere que se informe a los evaluados previamente de la existencia de escalas para detectar simulación, Boone discrepa de esta postura, argumentando que no hay evidencias empíricas que respalden la bondad de esta práctica. Como alternativa, Boone propone incluir en el consentimiento informado el siguiente texto:

“Se me ha explicado la importancia de poner el máximo esfuerzo en la realización de las pruebas que voy a realizar. También se me ha indicado que, si bien hay pacientes que tienden a exagerar sus problemas en los tests para asegurarse de que sus dificultades quedan bien reflejadas, en mi caso, más que ayudarme, podría hacer más difícil de interpretar mis resultados en las pruebas”.

Finalmente, Rogers, Bender, et al. (2011a) también han criticado los criterios de Slick et al. (1999). A diferencia de Larrabee et al. (2007), que valoraban los criterios como excesivamente conservadores, estos autores afirman que tienen un marcado riesgo de sobrediagnosticar simulación debido a su escasa especificidad. Además, advierten que no existen evidencias empíricas o estudios que avalen su uso en la práctica profesional, desaconsejando de forma rotunda su empleo.

En esta línea, consideran que los criterios de Slick et al. (1999) resultan poco exigentes para determinar simulación (criterios A, B y D) a la par que son muy exigentes para considerar hipótesis alternativas a la simulación (criterio D). A continuación, se detallan estas críticas criterio por criterio (en la tabla 1.8 puede encontrarse un resumen de las mismas).

- Criterio A. (Presencia de incentivos externos sustanciales). Rogers, Bender, et al. (2011a) señalan que existe una marcada diferencia entre la existencia de un incentivo externo y que efectivamente sea una motivación consciente para el evaluado (requisito para el diagnóstico de simulación). Consideran que es un criterio muy laxo que fácilmente puede cumplirse ya que el mero hecho de encontrarse en un contexto en que se puede obtener algún beneficio supone su cumplimiento y, por lo tanto no se requiere ninguna evidencia adicional de que el evaluado muestra motivación clara y marcada hacia la obtención de dicho incentivo externo.

- Criterios B y C (sesgo en sus respuestas a las pruebas cognitivas y exageración de síntomas informados o información no veraz y engañosa). El equipo de Rogers y colaboradores señala que con tan solo el cumplimiento de dos criterios (un criterio del de B2 a B6 y otro de C1 a C5) ya se cumple el diagnóstico de probable simulación, incluso aunque únicamente sean dos de los diez criterios que aparecen (es decir, un 20% de los criterios). Esto alarma más aún a estos autores cuando consideran que algunos de los criterios, tanto de B como de C, (concretamente los criterios B4, B5 y C1) incluyen la palabra “sugieren” dentro de su descripción, lo cual da pie a dudar sobre la solidez u objetividad como evidencias para el diagnóstico de simulación. Adicionalmente, consideran que otros criterios de C (C1, C4 y C5) tampoco resultan suficientemente sólidos o con evidencia empírica o teórica precisa que los respalde. Así, la ausencia de un adecuado *insight* o recuerdo de la propia historia (algo que, según afirman estos autores, no es infrecuente en poblaciones clínicas) es prueba de simulación; como también lo es la presencia de exageración de síntomas a partir de pruebas de autoinforme, sea cual fuera la esfera donde se exageran los síntomas, incluso aunque no tengan que ver específicamente con deterioro cognitivo. El equipo de Rogers señala que el criterio B2 se considera cumplido con un test de esfuerzo por debajo del punto de corte, independientemente de cuál sea el test seleccionado. Consideran este hecho inadecuado ya que existen marcadas divergencias en las propiedades psicométricas de los diferentes tests disponibles (distintos valores de sensibilidad y especificidad, entre otros) y no se pueden considerar intercambiables de forma arbitraria.
- Criterio D (las conductas observadas en los criterios B y C no son explicadas completamente por otros factores psiquiátricos, neurológicos o de desarrollo). En contraposición al resto de criterios, al equipo de Rogers et al, les sorprende que en el criterio D se exija que los factores psiquiátricos, neurológicos o de desarrollo deban explicar *completamente* esos resultados para descartar la simulación. Por ejemplo, el criterio habitual en el DSM para diagnosticar cualquier trastorno es que “no sea mejor explicado por” en vez de que “no sea completamente explicado” por otros factores psiquiátricos, neurológicos o de desarrollo.

La crítica que realizan estos autores incluye además una revisión bibliográfica de los estudios empíricos que se han realizado sobre los criterios de los Slick et al. (1999). Rogers y colaboradores identificaron un total de 19 estudios, de los que se pueden extraer las siguientes conclusiones:

- El criterio A no fue investigado en ninguno de los estudios.
- Tres de los cinco criterios B no fueron investigados en la mayoría de los estudios (B3 y B5 solo en un 21,1%; B4 solo en un 15,8%).
- Cuatro de los cinco criterios C no fueron investigados en la mayoría de los estudios (C1, C2 y C3 en un 10,5%; C4 en un 5,3%).
- El criterio D de exclusión no fue tratado en la mayoría de los estudios (tan solo en un 15,8%), sin que en ninguno de ellos se descartara de forma sistemática la presencia de trastornos del desarrollo. De acuerdo a estos autores, en 15 de los 19 estudios no se tenían en cuenta la mayoría de los criterios del modelo de Slick, lo cual consideran una seria limitación al respaldo empírico existente del modelo.

Tabla 1.8

Resumen de los posibles sesgos en los criterios de Slick et al. (1999) de acuerdo a Rogers, Bender, et al. (2011a)

	Criterio	Problema según Rogers et al.	Descripción del problema según Rogers et al.
Poco exigentes y específicos para diagnosticar simulación	A	Considera suficiente la presencia de incentivos externos.	La presencia de un incentivo externo es considerada equivalente a que el sujeto tiene el propósito o intención.
	B2 a B6, C1 a C5	Con pocos criterios se alcanza el diagnóstico de probable simulación.	Solo se requieren dos o más criterios (un 20% de los 10 criterios establecidos).
	B4, B5 y C1	Mínimo nivel de certeza requerido para que se cumplan algunos criterios.	Se usa la expresión “sugiere” como definición del cumplimiento del criterio.
	C1 y C4	La negación de un historial previo es considerado “prueba” de simulación.	La negación de un historial previo psiquiátrico o de problemas cognitivos es criterio de presencia de simulación.
	C5	La exageración en cualquier dominio “prueba” la presencia de simulación.	La presencia de exageración en cualquier área es criterio de presencia de simulación.
Muy exigentes para descartar simulación y considerar hipótesis alternativas	D	Máximo nivel de certeza requerido para considerar hipótesis alternativas a la simulación.	Requiere que la hipótesis alternativa a la simulación explique <i>completamente</i> (al 100%) la evidencia encontrada.
	D	Alta exigencia para descartar que subyazca una motivación para simular.	Requiere una disminución significativa de la capacidad para considerar las normas.
	C2 y C3	No se considera el efecto de un escaso <i>insight</i> como motivo de la discrepancia.	Basta con la discrepancia entre síntomas informados y patrón conocido de funcionamiento del cerebro o comportamiento observado.

Boone (2011) ha respondido a las críticas de estos autores considerando que magnifican las posibles limitaciones de los criterios de Slick y colaboradores. La autora concuerda en que tras 13 años desde su publicación hay algunos puntos que deben mejorarse en los criterios, como ya señaló en Boone (2007b) y Larrabee et al. (2007). Concretamente está de acuerdo con Rogers y colaboradores en la necesidad de:

- Aumentar la especificidad del criterio B2 requiriendo el fallo en dos o más tests de validez de síntomas (en vez de en uno como aparece ahora).
- Circunscribir el criterio C5 a las escalas de autoinforme que específicamente hayan mostrado una capacidad satisfactoria de detección de exageración de síntomas cognitivos no creíbles.
- Investigar el posible solapamiento entre algunos criterios B y C para evitar que de una misma conducta pueda deducirse el cumplimiento de criterios en B y C y llegar al diagnóstico de probable de simulación.
- Operacionalizar más claramente el criterio D para ver qué factores pueden ser excluyentes del diagnóstico de simulación.

Sin embargo, Boone discrepa en el rechazo global a los criterios y considera desproporcionada la enmienda total de Rogers y colaboradores. Específicamente, señala los siguientes puntos en relación a alguna de las críticas de Rogers y colaboradores:

- Criterio A. (Presencia de incentivos externos sustanciales). Boone (2011) considera que es un criterio necesario y acertado ya que supone que está presente el contexto donde puede darse la simulación. A efectos ilustrativos pone la metáfora de requerir para el diagnóstico de abuso de alcohol que la persona consuma alcohol; ciertamente luego habrá que ver si se cumplen los criterios de abuso, pero lo primero que tiene que darse es un contexto que pueda permitir que exista ese abuso. En relación con el debate sobre la intencionalidad o existencia de motivación consciente, considera que, como ya señaló en Boone (2007b), resulta una vía estéril de análisis por la extrema subjetividad de este hecho y valora como preferible la detección de síntomas no creíbles en un contexto donde el paciente dispone de incentivos externos que conseguir con dicha simulación.
- Criterios B y C (sesgo en sus respuestas a las pruebas cognitivas y exageración de síntomas informados o información no veraz y engañosa). Esta autora señala que Rogers y colaboradores omiten en su análisis que los criterios demandan una *marcada* discrepancia entre los resultados en las pruebas y otras informaciones, lo que limita el

hecho de que posibles discrepancias menores puedan hacer saltar los criterios y provocar su menor especificidad. De hecho, indica que, en el criterio B2, dado que los tests de esfuerzo o validez de síntomas están creados para disponer de una alta especificidad (habitualmente cercana al 0,90) es más probable la existencia de falsos negativos que de falsos positivos.

La contrarréplica de Rogers, Bender y Johnson (2011b) afirma que la mayoría de los puntos clave planteados originalmente por ellos no han sido rebatidos por Boone (p. ej., el hecho de que solo se requieran el 20% de los criterios cuando habitualmente en el DSM se requieren para el diagnóstico entre el 43% y el 63%, el empleo del verbo “sugerir” dentro de los criterios, la falta de atención a la carencia de *insight* del paciente sobre sus problemas o sobre su historial o el requisito de “ser completamente explicado” para los criterios de exclusión) y que aquellos que han sido rebatidos no lo han sido con datos empíricos en los que apoyarse.

En su contrarréplica, Rogers y colaboradores señalan que Boone (2011) se ha centrado en la defensa de la especificidad de los tests de esfuerzo (criterio B2), considerando que, incluso aunque estuviera en lo cierto, solo indicaría que uno de diez criterios (B2-B6, C1-C5) es suficientemente específico como para evitar falsos positivos, pero ¿qué pasa con el resto de criterios? Además, afirman que cuando se repasa la lista de puntos en los que está de acuerdo en su modificación (aquellos que mencionó en Boone, 2007b y los de Larrabee et al., 2007) puede verse un alto grado de concordancia en la necesidad de reformar a fondo los criterios de Slick y colaboradores para el diagnóstico de simulación.

Probablemente en los próximos años se prosiga en la depuración de unos criterios consensuados de detección de simulación en las distintas áreas. No obstante, los avances realizados en estas dos últimas décadas en la conceptualización de la simulación, en los posibles modelos explicativos y en los criterios diagnósticos han sido notables. Este hecho, junto con el desarrollo de pautas metodológicas en el campo, como se expondrá en el siguiente capítulo, ha sentado unas ciertas bases para el crecimiento y desarrollo de la investigación en esta área.

2. Metodología de investigación en simulación

El avance en el campo de la simulación frecuentemente se ha visto lastrado por la ausencia de un marco teórico y metodológico que guiara su desarrollo. Con el objetivo de mitigar estas carencias se han ido construyendo modelos conceptuales que tratan de estructurar y articular las distintas vías de investigación, tal y como se ha comentado en el capítulo anterior. Por otra parte, en el campo metodológico ha ido surgiendo una serie de estándares en cuanto a diseños a emplear y metodología de análisis de datos, que ha ido proporcionándole cierto rigor.

Este avance ha determinado que, en la actualidad, se considere como requisito necesario para considerar una estrategia de detección de simulación el que exista tanto una fundamentación conceptual como una validación empírica (Rogers et al., 2010). En el capítulo previo se revisaron aquellos aspectos relevantes a la fundamentación conceptual y en éste se revisará qué se entiende por validación empírica específicamente en esta área.

En esencia, la validación empírica requiere que se hayan utilizado diseños de investigación y técnicas de análisis que permitan responder adecuadamente a las preguntas que se hacen en el marco de la simulación y determinar si hay diferencias sistemáticas y relevantes entre los sujetos simuladores y los pacientes clínicos genuinos.

Se presentarán, en primer lugar, los diseños habitualmente utilizados en simulación, incluyendo aquellos puntos clave a tener en cuenta para su correcta implementación, sus puntos débiles y el tipo de conclusiones que nos permiten establecer.

A continuación, se describirán las técnicas de análisis habitualmente empleadas específicamente en este campo para valorar la magnitud de las diferencias y los estándares considerados en la investigación en simulación para establecer su relevancia.

2.1. Diseños de investigación habitualmente empleados en estudios de simulación

El objetivo de cualquier diseño de investigación es establecer una estrategia que permita dar respuesta a la pregunta o problema planteado por el investigador (Breakwell, Smith y Wright, 2012). Por lo tanto, todo diseño debe constituir un adecuado esquema de acción, cuya importancia es tal, que la mayor parte de los errores que invalidan un trabajo de investigación se refieren precisamente a la incorrecta elección del mismo.

El diseño requiere todo un laborioso proceso de toma de decisiones, persiguiendo dos objetivos fundamentales: proporcionar respuestas a las preguntas por las que se ha iniciado la investigación y controlar las posibles fuentes de variables contaminadoras. La finalidad del diseño no es otra que permitir al investigador obtener respuestas inequívocas a sus preguntas.

No existe un principio único que establezca qué diseño es el más adecuado, sino que la elección debe realizarse de acuerdo al contexto concreto y en virtud de una serie de consideraciones tales como simplicidad, precisión, validez (interna y externa) y ausencia de errores sistemáticos.

Dada la dificultad de investigar la simulación en condiciones reales (los simuladores no suelen identificarse voluntariamente), una gran cantidad de investigación es realizada bajo condiciones “análogas”, esto es, condiciones que tratan de asemejarse a las reales. En estos diseños, se solicita a un grupo de sujetos que simulen una determinada patología o trastorno, tratando de crear para ello unas condiciones equiparables a los contextos reales.

Las ventajas generales de la investigación análoga son su relativa facilidad de realización, su bajo coste y el control experimental que proporcionan. Estas ventajas permiten la investigación original de nuevos procedimientos e ideas que, más tarde, pueden ser llevadas a situaciones clínicas o forenses para ser investigadas. Así, la investigación en situaciones clínicas podría quedar reservada para planteamientos que han sido probados previamente y que tienen, por ello, una cierta probabilidad de aportar resultados de interés.

La investigación análoga permite al investigador controlar muchas de las variables que usualmente fluctúan en la situación clínica. En este sentido, los estudios análogos maximizan la validez interna si bien presentan limitaciones en cuanto a su validez externa. En el polo opuesto, se situarían los estudios en contextos reales que maximizan la validez externa pero que presentan restricciones notables en su validez interna. Ambos son extremos de un continuo de diseños empleados.

Por lo tanto, la investigación análoga es especialmente útil para analizar diversas facetas o aspectos de la simulación, descubrir nuevas estrategias o parámetros que influyen en la eficacia de la detección y cuestiones similares que requieren un cuidadoso control experimental. Posteriormente, podrán llevarse dichas cuestiones a otros diseños con más validez externa, como el diseño de grupos conocidos o el diseño de medidas autosuficientes, con el objetivo de replicar los resultados y obtener evidencias de generalización. Tal y como afirma Rogers (1990a), “It is only through the convergent results with both research methods that researchers can have confidence in the validity of their findings” (p. 186).

Es importante recalcar la necesidad de combinar ambos diseños en la investigación de nuevas medidas o técnicas de detección de simulación ya que solo podrán establecerse criterios y puntos de corte adecuados a partir de los resultados convergentes de ambos. Un ejemplo paradigmático es la escala FDR del PAI (Rogers, Sewell, Morey y Ustad, 1996), creada originalmente a partir de estudios análogos en los que había mostrado una excelente capacidad de discriminación y en los que se habían fijado puntos de corte con nuevas muestras independientes. No obstante, cuando se ha probado su uso en contextos reales ha mostrado una muy limitada capacidad de detección, lo que ha motivado que se desaconseje su uso actualmente, a la espera de nuevas investigaciones que puedan confirmar o descartar su uso en contextos reales (Hawes y Boccaccini, 2009; Sellbom y Bagby, 2008).

En la tabla 2.1 se incluye una breve descripción de los principales diseños de investigación empleados en simulación y sus características. En todos ellos es posible aumentar o garantizar su validez siguiendo una serie de pautas clave que se describirán posteriormente.

Antes de pasar a la descripción de estos diseños, cabe mencionar someramente el estudio de caso. Este enfoque parte del estudio de casos individuales como vía de generación de hipótesis, que pasarán a ser estudiadas con otros diseños. Por lo tanto, su objetivo es la generación de ideas y posibles nuevas hipótesis. Además, aunque la información extraíble de un caso único está considerablemente limitada en su generalizabilidad, puede ser el único diseño viable en la investigación de síntomas extremadamente raros como el síndrome de Munchausen.

Ha de advertirse que uno de los riesgos habituales en el estudio de caso es la identificación de características de los simuladores que, si bien son salientes, llamativas o destacadas, pueden no ser distintivas o discriminativas con relación a los casos clínicos genuinos. Por este motivo es requisito fundamental que las hipótesis derivadas de estos estudios sean verificadas en diseños más completos como los que se describen a continuación.

Tabla 2.1

Visión general de diseños de investigación empleados en simulación: diseños análogos y de grupos conocidos (Rogers, 2008a)

Análogos
<p><i>Descripción.</i> Se asignan aleatoriamente los participantes a las diferentes condiciones experimentales de simulación. Habitualmente los resultados de estos sujetos son comparados con los de grupos clínicos relevantes que responden honestamente.</p> <p><i>Validez interna.</i> Alta: Se incluyen procedimientos estandarizados de instrucciones, condiciones e incentivos.</p> <p><i>Validez externa.</i> Baja: Aunque los estudios análogos tratan de replicar las condiciones reales, inevitablemente existirán diferencias con las situaciones reales ya que, por ejemplo, los sujetos no se enfrentan a las graves circunstancias y consecuencias de los contextos reales ni han pasado por el periplo legal o médico que les puede permitir un mayor aprendizaje de la simulación.</p> <p><i>Clasificación.</i> Puede evaluarse la precisión de la medida en la clasificación de sujetos honestos y simuladores mediante validaciones cruzadas.</p> <p><i>Mal uso de este diseño.</i> En los casos en los que no se incluyen grupos clínicos se limita drásticamente su utilidad ya que no permiten determinar si existirán las necesarias diferencias entre sujetos simuladores de una patología y sujetos honestos con dicha patología (pacientes clínicos genuinos). Contar con grupos clínicos de comparación es requisito para la correcta implementación de este diseño.</p>
Grupos diagnósticos conocidos
<p><i>Descripción.</i> Los participantes se encuentran en contextos reales en los que es probable la presencia de simulación. Expertos independientes realizan una clasificación de los participantes de acuerdo al grado de simulación presente, estableciendo al menos dos grupos: probable simulación y probable respuesta honesta.</p> <p><i>Validez interna.</i> Baja: Los investigadores no tienen control sobre la asignación experimental a los grupos u otros aspectos del procedimiento.</p> <p><i>Validez externa.</i> Alta: los participantes, los contextos, las situaciones y los incentivos se ajustan a las condiciones reales en que ocurren.</p> <p><i>Clasificación.</i> Puede evaluarse la precisión de la clasificación de sujetos honestos y simuladores mediante validaciones cruzadas.</p> <p><i>Mal uso de este diseño.</i> El empleo de expertos poco fiables o dependientes de la medida bajo estudio puede invalidar todo el diseño. Este aspecto ha de ser cuidado en extremo en estos casos. Por otra parte, para garantizar la calidad de la clasificación de los expertos es deseable que se consideren al menos tres grupos: probable simulación, indeterminado y probable respuesta honesta. La existencia del grupo “indeterminado” evita la asignación de casos dudosos a uno de los dos grupos (probable simulación, probable respuesta honesta) considerados en los análisis.</p>

Tabla 2.2

Visión general de diseños de investigación empleados en simulación: diseños de medidas autosuficientes (bootstrapping comparison) y de prevalencia diferencial (Rogers, 2008a)

Diseños de medidas autosuficientes (*bootstrapping comparisons*)

Descripción. Los participantes se encuentran en contextos reales en los que es probable la simulación y en los que se aplica un protocolo para su detección. A partir del uso de otras medidas de detección (excluyendo aquella bajo estudio) se establecen puntos de corte estrictos (máxima especificidad), que puedan ser aplicados para la creación de los grupos de probable simulación y probable respuesta honesta.

Validez interna. Baja. Los investigadores no tienen control sobre la asignación experimental a los grupos o sobre otros aspectos del procedimiento.

Validez externa. Alta: los participantes, los contextos, las situaciones y los incentivos se ajustan a las condiciones reales en que ocurren.

Clasificación. Puede evaluarse la precisión de la medida en la clasificación de sujetos honestos y simuladores, aunque con limitaciones. La mayor incertidumbre es en relación con los falsos positivos, dado que se desconoce el porcentaje de sujetos del grupo asignado (esto es, simuladores) que realmente no pertenecen a dicho grupo. La cuidadosa aplicación de indicadores externos con valores establecidos de sensibilidad y especificidad pueden hacer de este diseño una interesante vía de estudio.

Mal uso de este diseño. Al igual que se establecía en los diseños de grupos conocidos, el empleo de criterios diagnósticos poco fiables o dependientes de la medida bajo estudio puede invalidarlo. Este aspecto ha de ser cuidado en extremo. Por otra parte, para garantizar la calidad de la clasificación es deseable que se consideren al menos tres grupos: probable simulación, indeterminado y probable respuesta honesta.

Prevalencia diferencial

Descripción. Teniendo en cuenta los incentivos presentes, se asume que un amplio número de sujetos de un grupo definido (p. ej., litigantes) presentará un patrón de exageración de síntomas en comparación con otro segundo grupo (p.ej., no litigantes) en el que no están presentes dichos incentivos.

Validez interna. Baja. Los investigadores no tienen control sobre la asignación experimental a los grupos o sobre otros aspectos del procedimiento.

Validez externa. Moderada. Los participantes habitualmente están inmersos en situaciones reales en las que tiene importantes consecuencias el éxito o fracaso en la simulación. No obstante, la falta de que un criterio claro que determine qué participantes presentan efectivamente un estilo de exageración de síntomas limita drásticamente la validez del estudio. El pertenecer a uno de los grupos no garantiza que se esté empleando dicho estilo de respuesta, solo puede afirmarse con cierto grado de seguridad que la prevalencia de la exageración de síntomas será mayor en uno de los grupos frente al otro.

Clasificación. Incontrastable. Al poder establecerse únicamente la prevalencia diferencial de ambos grupos, pero no la identificación de cada caso individual como simulador o paciente clínico genuino, no se puede verificar la precisión de la medida en la clasificación de los casos individuales.

Mal uso de este diseño. Este diseño presenta importantes limitaciones derivadas de la incertidumbre sobre qué participantes se consideran simuladores y cuáles presentan una respuesta honesta. La simple prevalencia diferencial en ambos grupos evita que pueda establecer la precisión de la medida, así como puntos de corte clasificatorios. Igualmente, limita el alcance y la interpretación de los resultados que se obtengan.

2.1.1. Análogos

La mayor parte de las investigaciones llevadas a cabo en el ámbito de la simulación ha empleado este diseño ya que es el que permite un mayor grado de control experimental y resulta más accesible y configurable. Por este motivo, será el diseño que más detalladamente se describirá en este capítulo.

En estos diseños suele disponerse de participantes no clínicos que aleatoriamente son asignados a las condiciones de respuesta honesta (grupo control) o respuesta de simulación (grupo o grupos experimentales). A estos últimos se les da instrucciones específicas de cómo responder a la prueba simulando ciertos trastornos. Habitualmente, los resultados en estos grupos control y experimental son comparados entre sí y con los obtenidos por un grupo clínico para determinar las diferencias entre los sujetos instruidos para simular y aquellos casos clínicos genuinos.

A continuación se revisarán los distintos aspectos críticos a tener en cuenta en estos diseños. En la tabla 2.3 se incluye un pequeño resumen de los mismos (Rogers, 2008e).

Tabla 2.3
Aspectos clave a tener en cuenta en los diseños análogos (Rogers, 2008e)

<i>Relevancia de las comparaciones</i>
- ¿Se incluye un grupo clínico con el que se comparan resultados de análogos y controles?
<i>Instrucciones</i>
- ¿Resultan claras y comprensibles?
- ¿Son suficientemente específicas como para reflejar situaciones reales?
- ¿Se proporciona un contexto que resulte familiar a los participantes y del que puedan tener cierta experiencia?
- ¿Se destaca la relevancia del estudio como vía de motivación e implicación en la investigación?
- ¿Se incluyen incentivos?
- ¿Se advierte que la presentación debe de ser creíble, convincente y realista?
<i>Preparación</i>
- ¿Se les da instrucciones específicas sobre los síntomas principales a ser simulados en la patología? ¿Se proporciona a los sujetos tiempo para preparar la simulación?

2.1.1.1. Relevancia de las comparaciones

Cuando el estudio análogo se realiza con población no clínica, como suele ser habitual, la mera comparación experimental entre el grupo control (que responde honesta y sinceramente) y el grupo análogo aporta poco valor sustantivo a la investigación ya que no responde a la pregunta primordial, esto es, la diferencia que se establece entre simuladores y pacientes clínicos genuinos.

Así, es posible que la discriminación de la prueba entre controles y simuladores no esté presente cuando se incluyan casos clínicos y que la prueba no sea capaz de discriminar realmente entre pacientes genuinos y los que simulan sus trastornos. Aquellas investigaciones que solo dispongan del grupo control de personas normales respondiendo honestamente y el grupo de simuladores análogos, están sesgadas intrínsecamente y no deberían ser publicadas. Esta limitación es más evidente aún cuando las muestras empleadas han sido de estudiantes, dado que los resultados que se obtengan tendrán un valor sustantivo muy limitado.

Tal y como afirmaban Faust y Ackley (1998):

La mayoría de las investigaciones sobre simulación estudian justo la discriminación que un forense o un clínico nunca tendrán que hacer, la distinción entre individuos normales respondiendo honestamente e individuos normales a los que se les ha pedido que respondan fingiendo encontrarse mal. Más bien, a lo que frecuentemente se enfrenta el profesional es a distinguir entre pacientes clínicos con algún trastorno e individuos que lo están simulando. La tarea habitual del clínico es bastante más difícil que discriminar entre normales y normales simulando (p.9).

Conscientes de esta limitación y de la necesidad de grupos clínicos de referencia, en algunas investigaciones se seleccionan grupos "clínicos" a partir del propio grupo control de estudiantes. Partiendo de sus respuestas en determinadas pruebas o de los síntomas alegados se les asigna a un grupo clínico. Esta solución, más justificada por su conveniencia que por su rigor experimental, introduce importantes limitaciones al estudio, dado que, aunque puedan manifestar sintomatología y cierto grado de malestar, es discutible el grado en que es análogo al experimentado por casos clínicos reales, en cuanto a su nivel de severidad y su impacto en el funcionamiento y las actividades cotidianas.

Incluso, aunque mediante entrevista clínica se delimitaran y justificaran los diagnósticos, el investigador tendría que ser consciente de la limitada extrapolación de sus resultados a grupos clínicos más afectados. Los resultados serían generalizables a grupos clínicos de sujetos con un alto nivel de funcionamiento y que están activamente implicados en la consecución de metas a largo plazo (la obtención de una titulación universitaria). En muchos casos, este perfil es notablemente diferente del que se encuentran en la clínica donde el grado de afectación suele ser más severo y generalizado.

En estos diseños es por lo tanto un requisito ineludible disponer de un grupo clínico de sujetos que padezcan los trastornos estudiados y que respondan honestamente a la misma. El análisis clave del estudio será la comparación entre simuladores análogos y pacientes clínicos.

Otra práctica que puede darse en estos diseños es utilizar un mismo grupo de participantes a los que se les pide responder como simuladores y posteriormente como controles o respuesta honesta. En este caso, es interesante tener en cuenta los resultados de Merckelbach, Jelicic y Pieters, (2010) a la hora de diseñar o interpretar los resultados de este tipo de estudios. Estos autores aluden a un efecto residual de la simulación que hace que, cuando posteriormente se le pide que respondan honestamente, sigan mostrando un ligero patrón de exageración de síntomas en comparación con los que no han pasado previamente por la condición de simulación.

2.1.1.2. Instrucciones proporcionadas

Gran parte del éxito que pueda tener un estudio análogo dependerá del grado en que replique las condiciones reales. En esta tarea juegan un papel clave las instrucciones que se proporcionan. Éstas tienen que ser claras y comprensibles, a la par que motivadoras, para que la persona actúe como si la situación fuera real en todos los sentidos posibles. Ha de remarcarse que ambos aspectos, el comunicativo y el motivador, son igualmente importantes, siendo el último frecuentemente olvidado o relegado.

Típicamente estas instrucciones se concretan en dos o tres párrafos descriptivos que se presentan por escrito a los sujetos. Resulta una cuestión interesante, pero desconocida en la actualidad, si otros formatos de presentación (multimedia, video...) pudieran mejorar su eficacia.

Rogers y Cruise (1998) definieron distintos requisitos para las instrucciones, concretamente indicaron que deberían ser claras, específicas, contextualizadas, motivadoras y que remarcaran la necesidad de presentar perfiles creíbles y realistas.

Claras

La cuestión fundamental es que el participante obtenga una idea clara de lo que se espera de él, que pueda captar la esencia de la tarea y el rol que debe adoptar en ella. Para ello, las frases empleadas deben ser precisas y sencillas. Es recomendable que tanto el vocabulario empleado como las estructuras sean asequibles para sujetos de cualquier nivel cultural.

Específicas

Un aspecto íntimamente relacionado con el anterior es que las instrucciones sean específicas y directas, de forma que el sujeto sepa concretamente cuál es la condición que debe simular. Esto es también extremadamente relevante de cara a que el investigador pueda extraer conclusiones apropiadas a la condición simulada por el sujeto.

Dependiendo de los objetivos concretos de la investigación, las instrucciones pueden incitar a la simulación de un trastorno específico (enajenación mental para eludir una condena, discapacidad para obtener una invalidez, estrés postraumático para conseguir una indemnización...) o de una condición más genérica.

Cabe advertir que deben evitarse las instrucciones excesivamente complejas o detalladas ya que pueden dificultar la comprensión por parte del participante y su habilidad para ponerlas en práctica de forma efectiva y cabal.

Contextualizadas

Toda simulación ocurre en un determinado contexto que le da sentido; por ejemplo, la simulación de síntomas depresivos suele darse en un contexto de solicitud de baja o de demanda de daños y perjuicios por algún hecho acaecido. Proporcionar este contexto a los participantes es fundamental para que puedan reflejar, de la forma más verosímil posible, la situación real de simulación. El contexto debe ser lo más cercano al sujeto de forma que pueda imaginarse que esa situación le está ocurriendo a él. Por ejemplo, en la investigación de Gillard (2010) con estudiantes universitarios, se encontró un mayor efecto en la simulación análoga cuando se les planteaba un contexto y objetivo vinculado a su situación actual (suspenso en una asignatura importante), que cuando el contexto y objetivo era evitar la pena capital en un juicio por asesinato. No obstante, ha de señalarse que este ajuste del contexto a la realidad de la persona puede afectar a la generalizabilidad de los resultados y a su condición de “analogía” ya que, lógicamente, no resulta equivalente el incentivo de evitar el suspenso en una asignatura que eludir la pena de muerte.

Especificar claramente en las instrucciones cuál es el contexto de simulación también es relevante para saber en qué uso concreto estamos validando la herramienta. Por ejemplo, Merckelbach, Smeets y Jelicic (2009b) observaron que los resultados de sensibilidad y especificidad eran diferentes si la simulación de un trastorno era para evitar una condena (contexto forense-criminal) o para solicitar una indemnización (contexto civil). Los síntomas que simuladores análogos informan aparentar son parcialmente diferentes en ambos contextos (Dandachi-FitzGerald y Merckelbach, 2013). Peace y Masliuk (2011) también evidenciaron diferencias en la simulación del

TEPT según cuál hubiera sido el evento traumático (accidente, desastre o asalto sexual) y la motivación por la que simulaban (compensación, venganza o ninguna motivación en particular). Si bien el número de investigaciones que han indagado en este aspecto es reducido, sistemáticamente todos han apuntado a una sensibilidad del contexto y de los objetivos de la simulación por parte de los simuladores. Esto es coherente con el modelo adaptativo de simulación en el que la persona la emplea como una respuesta estratégica y flexible hacia situaciones con consecuencias negativas para ella (González-Ordi et al., 2012; Rogers, 2008a).

Motivadoras e incentivadoras

Quizás uno de los aspectos que más puede diferenciar la situación en que responde un análogo y un caso real es el grado de implicación en la tarea. Mientras los casos reales habitualmente tienen incentivos externos muy relevantes para sus vidas, los análogos suelen ver la tarea como un mero trámite para conseguir determinados créditos o calificaciones, en el caso de alumnos, o una limitada cantidad económica, en el caso de investigaciones remuneradas.

Con el objetivo de conseguir implicar al participante en la tarea y mimetizar en la medida de lo posible las condiciones reales, puede ser conveniente emplear múltiples estrategias complementarias, entre las que se incluyen:

- plantear la tarea como un desafío (p.ej., ¿serás capaz de pasar sin ser detectado?, ¿podrás engañar al test?).
- destacar la relevancia de la investigación y de su participación (p.ej., “una mejor detección de la simulación permitiría reducir los tiempos de espera de los servicios de tratamiento y rehabilitación para las personas realmente necesitadas” o, en contextos universitarios, “ese dinero se podría emplear para unas mejores instalaciones y más recursos y prestaciones”).
- incluir algún tipo de recompensa a los sujetos que simulen la patología sin ser detectados y de penalización a aquellos que no consigan superar la prueba y sean detectados.

Si bien resulta claro que es deseable incluir alguna forma de recompensa que acentúe la implicación en la tarea, no existe consenso sobre cómo debe ser este procedimiento o qué efecto tiene cada una de las opciones posibles (Elhai et al., 2007).

Responder a la pregunta sobre qué supone un incentivo relevante para el sujeto dependerá del grupo concreto que se esté evaluando. No es lo mismo un incentivo económico de 30 euros para un estudiante universitario que para un trabajador o para un preso. Su significado puede variar notoriamente (Schretlen, Wilkins, Van Gorp y Bobholz, 1992). Además, puede resultar que para unos

sea más relevante un incentivo económico y para otros un incentivo de tipo social (v.g., reconocimiento social del éxito en la simulación). Incluso incentivos que a priori el investigador puede considerar relevantes, pueden que no lo hayan sido al ver su valoración en un cuestionario posterior al experimento (Viglione et al., 2001).

En las investigaciones se pueden encontrar incentivos económicos que oscilan entre los 2 (Martin, Bolter, Todd, Gouvier y Niccolls, 1993) y los 100 dólares (Blanchard, McGrath, Pogge y Khadivi, 2003), si bien la amplia mayoría se comprende entre los 20 y los 50 dólares.

No obstante, no está claro el efecto que puede tener un mayor incentivo, barajándose dos hipótesis: que vuelva a los sujetos más cuidadosos y, por lo tanto, menos detectables o, por el contrario, que aumente sus deseos de mostrarse afectados y, por lo tanto, hacerlos más fácilmente detectables. Aunque la evidencia disponible hasta ahora es contradictoria, parece que no hay un claro efecto de la magnitud económica en las respuestas dadas, al menos en los valores relativamente modestos que habitualmente se manejan (Bernard, 1990; Elhai et al., 2007; Martin, Bolter, Todd, Gouvier y Niccolls, 1993).

Tampoco se conoce qué efecto pueden tener otro tipo de incentivos como un viaje de fin de semana, un par de entradas a un evento artístico o deportivo... Se desconoce igualmente si es más acertado un premio pequeño para todos los que superen la prueba sin ser detectados o un premio más jugoso pero sólo para los 3 ó 4 mejores. Todas estas vías están pendientes de investigación.

Es interesante señalar también el posible papel de las penalizaciones. Los simuladores en contextos reales no solo tienen en su horizonte las posibles recompensas, sino también los posibles castigos si son detectados. Una persona que simula una depresión para obtener una baja laboral se arriesga a la reprobación pública en caso de ser detectada, así como a que esta información sea incluida en su historial, de cara a futuras reclamaciones y dolencias.

Por lo tanto, en las condiciones análogas pueden utilizarse tanto incentivos positivos como negativos, a ser posible en forma de desaprobación pública. Por ejemplo, Patrick y Iacono (1989) sugirieron publicar los nombres de las personas que fracasasen en un tablón expuesto al público (aunque posteriormente por razones éticas no se publicaran más que nombres ficticios y no los reales) o de pérdida de un beneficio (p.ej., no obtener las horas de créditos de práctica con que se premia habitualmente a estudiantes).

Los estudios realizados parecen indicar que la presencia de penalizaciones hacen que las personas se centren más en la patología a simular y hagan una menor asignación genérica de

síntomas, tal y como afirman Rogers y Cruise (1998): “The most striking finding was that the type of incentive appeared to affect simulators' presentations. In the face of a very mild negative incentive, simulators appeared to become more focused in their presentation” (p.282).

Que remarquen la credibilidad

Este punto está íntimamente relacionado con lo mencionado en el apartado anterior. Los simuladores reales tienen mucho en juego y buscan mostrarse afectados pero de una forma que resulte creíble y no dramática, midiendo sus respuestas. Puede ocurrir que los simuladores análogos, menos implicados en la tarea y menos conscientes de la necesidad de resultar creíbles, se embarquen en un patrón de respuesta exagerado ya que no se sienten intimidados o motivados a ser creíbles y mesurados.

Quizás esto puede ser una de las causas que hace los simuladores análogos sean los que suelen obtener puntuaciones más altas en los índices de exageración de síntomas, por encima de los simuladores en contextos reales. De hecho, Viglione et al. (2001) afirman que gran parte de las diferencias entre simuladores análogos y reales puede ser explicada por el nivel de precaución o cautela de ambos; los simuladores reales, frecuentemente, tienen más oportunidades para probar y refinar sus papeles de simuladores de modo que llegan a ser más sutiles con el tiempo y a reflejar con más éxito la patología alegada.

En esta investigación realizada por Viglione et al. (2001) se observó que la introducción de la advertencia de credibilidad producía puntuaciones sustancialmente más bajas en los índices de exageración de síntomas, concretamente, reducía en promedio una desviación típica los valores. Resultados similares se han encontrado en otras investigaciones (v.g., Elhai et al., 2007), sugiriendo que su introducción produce un patrón más conservador y probablemente más realista, de asignación de síntomas. Las investigaciones realizadas hasta la fecha parecen indicar que esta instrucción de credibilidad es uno de los aspectos que más drásticamente afecta a los resultados que se obtengan, por lo que será un factor clave a tener en cuenta.

Con el objetivo de introducir esta advertencia en las instrucciones, habitualmente se han incluido textos como los siguientes: “si tus respuestas son demasiado dramáticas, es más probable que seas detectado y no podrás conseguir tus objetivos. Por lo tanto, sé precavido y representa un perfil creíble, convincente y realista de acuerdo al problema que se te ha indicado” o “contesta a la prueba sin perder de vista el problema que tienes que simular, intentando que tus respuestas sean lo más convincentes posibles, pero sin que puedan detectar que estás mintiendo”.

2.1.1.3. Preparación de la simulación

Ossipov (1944) ya afirmaba que los simuladores son actores que representan sus síntomas de la mejor manera que saben, esto es, de acuerdo a su conocimiento del papel que interpretan. Esta afirmación llama la atención sobre la importancia que pueda tener la preparación de la simulación en los contextos reales y, por lo tanto, en qué medida debería ser replicada en los estudios análogos.

Habitualmente no se le ha prestado atención a este aspecto en las investigaciones con análogos y a los participantes no se les suele conceder tiempo para preparar su estrategia de simulación o para recabar información relevante al trastorno que van a simular o al proceso de evaluación. Sería deseable que, en la medida de lo posible, los participantes análogos dispusieran de un cierto tipo de preparación para reproducir las condiciones reales.

Ahora bien, qué tipo de información se les debe proporcionar a los sujetos análogos para reflejar las condiciones reales, ¿se les debe incluir algún tipo de información sobre los trastornos a simular o sobre las propias estrategias de detección empleadas? ¿Qué información previa tienen los simuladores reales? ¿Qué es exactamente lo que saben? ¿Sólo disponen de información sobre el trastorno y cómo deben actuar, o su formación va más allá y saben también cómo funcionan las escalas de validez y cómo evitarlas?

Aunque son escasas las investigaciones que han indagado en el grado de preparación que tienen los simuladores o qué estrategias utilizan en dicha preparación (Dandachi-FitzGerald y Merckelbach, 2013; Iverson, 1995; Tan, Slick, Strauss y Hultsch, 2002), los datos disponibles parecen sugerir que los simuladores se dedican activamente a su preparación y a la búsqueda de información para conseguir su objetivo. También parece evidente que, al menos en EEUU, los abogados suelen instruir a sus clientes sobre las escalas de validez (Essig, Mittenberg, Petersen, Strauman y Cooper, 2001; Wetter y Corrigan, 1995; Youngjohn, 1995).

La relevancia de los incentivos hace predecible que dediquen cierto tiempo y esfuerzo a prepararse. Además, el amplio acceso a la información que permite Internet también facilita la preparación y obtención de información sobre simulación (Bauer y McCaffrey, 2006; Ruiz, Drake, Glass, Marcotte y van Gorp, 2002).

Rogers (2008c) parte de la asunción de que los simuladores reales tienen ciertos conocimientos de las estrategias de detección y, por lo tanto, aboga por instruir directamente a los análogos en dichas estrategias para igualar las condiciones con los contextos reales. Considera,

además, que informar sobre los trastornos a simular introduce principalmente desorientación y confusión en los participantes ya que:

- Los inventarios multiescalares de personalidad son bastante imprecisos en cuanto al diagnóstico, esto es, no es necesario obtener un perfil determinado en el MMPI-2 para que sea coherente con un diagnóstico de depresión.
- Esta información puede hacer que los participantes se focalicen en recrear el trastorno mostrado (escalas clínicas) y no en la tarea principal que es evitar ser detectados como simuladores (escalas de validez); esta preparación, por lo tanto, no mejorará su capacidad de eludir la detección y no se le estará preparando realmente para simular.
- no refleja la comorbilidad real existente habitualmente en estos casos ya que se informa de trastornos específicos;

Rogers considera muy probable que no se pueda guardar el secreto profesional sobre las estrategias que se emplean y que sólo las más complicadas y sofisticadas podrán sobrevivir (p. ej., consideración simultánea de varios síntomas para considerar su coherencia), aquellas que, aunque se conozcan en abstracto, sean después prácticamente imposibles de tener en cuenta al responder al test.

No obstante, tal y como afirma Viglione et al. (2001) los simuladores “attempt to match a specific disorder and a specific set of events with a reasonable life predicament. They must identify which symptoms are indicative of the disorder they are attempting to feign and which symptoms would be considered excessive or atypical” (p. 238). Por lo tanto, se considera que la principal preparación que debe darse a los sujetos análogos es sobre los trastornos a simular ya que, independientemente de la eficacia de esta preparación para eludir la detección, es muy probable que esto recree las condiciones reales de los simuladores que buscan información sobre los trastornos para poder reproducirlos. Preferentemente esta información debe ser clara, breve y escrita para que los participantes puedan centrarse en un conjunto específico de sintomatología.

2.1.2. Grupos diagnósticos conocidos

Es muy deseable que la investigación iniciada con análogos sea replicada en grupos diagnósticos conocidos. Tal y como se exponía en apartados anteriores, la investigación con análogos puede servir como banco inicial de pruebas en el que se experimenten nuevas ideas y métodos. No obstante, antes de pasar a la práctica profesional es extremadamente recomendable que se lleven a cabo estudios en contextos reales.

Los diseños con grupos diagnósticos conocidos se caracterizan por:

- a) ser evaluaciones realizadas en contextos reales en las que los resultados de la evaluación tienen efectos en los sujetos;
- b) disponer de un criterio externo y fiable de clasificación de los sujetos como simuladores o pacientes genuinos;

El primer requisito implica que los resultados de la evaluación afectarán al sujeto y tendrán consecuencias. No se incluyen dentro de este grupo aquellos participantes evaluados con fines académicos (p. ej., entrenamiento o investigación) ya que estos contextos no se pueden considerar estrictamente reales o ser equiparables a aquellos donde se usará la prueba en su modo aplicado.

El segundo criterio, disponer de un criterio externo para la clasificación, requiere prestar extrema atención a la selección de los expertos en simulación que van a realizar la clasificación o del conjunto de medidas altamente precisas que serán empleadas para la clasificación. Si en los grupos análogos el criterio clave para la relevancia del estudio era la replicación de las condiciones reales, en este caso, es la bondad del criterio elegido para realizar la clasificación. La tarea de establecer quiénes son simuladores y quiénes pacientes clínicos genuinos debe estar reservada a expertos entrenados en esta labor debe estar basada en exhaustivas evaluaciones.

Estos diseños tienen dos fases claramente definidas. Una primera, relacionada con el establecimiento de los grupos (clínicos genuinos y simuladores) y una segunda, de análisis de las diferencias y similitudes en ambos grupos en las variables investigadas. Lógicamente, el principal problema es realizar la clasificación de ambos grupos de una forma fiable y precisa, para lo cual se tendrá que utilizar el criterio de referencia seleccionado.

2.1.3. Diseños de medidas autosuficientes (*bootstrapping comparisons*)

En los grupos diagnósticos conocidos, las clasificaciones independientes son realizadas por expertos o criterios externos establecidos. A diferencia de estos diseños, en los diseños de medidas autosuficientes usan puntos de corte muy exigentes de medidas de simulación de forma que sea máxima la probabilidad de acierto en la asignación al grupo de simuladores o de respuesta honesta; además, suelen excluir los casos inciertos que no pueden ser clasificados con certeza y que son asignados al grupo de “indeterminado”. Puede aplicarse a este diseño una aproximación al de grupos diagnósticos conocidos. Este diseño es especialmente útil en la evaluación de la validez externa y el análisis del tamaño del efecto de las medidas o criterios bajo estudio.

Los criterios que definen este tipo de diseños son los siguientes (Rogers et al., 2010):

- a) indicadores independientes de simulación;
- b) uso de puntos de corte conservadores que producen relativamente pocos errores en ambos grupos clasificados;
- c) la exclusión de los casos indeterminados.

Un ejemplo de estudio que emplea este diseño puede consultarse en Rogers et al. (2009).

2.1.4. Prevalencia diferencial

Dentro de este diseño se encuadran aquellas investigaciones que comparan grupos conforme a su diferente prevalencia de simulación. Por ejemplo, los resultados en un grupo de personas implicadas en procesos legales, que persiguen obtener una determinada indemnización o beneficio (habitualmente denominados como “litigantes”), se comparan con los de otro grupo de personas con patología, pero que no están inmersos en dichos procesos legales, ni en la búsqueda de esos beneficios externos (habitualmente denominados como “no litigantes”). En estos casos, se parte de la suposición razonable de que el porcentaje de simuladores en el grupo de litigantes (en el que está presente un incentivo externo relevante) será mayor que en el de los no litigantes (donde no consta evidencia de dichos incentivos externos) y este supuesto es el que sostiene la comparación entre ambos grupos en las medidas de simulación.

Estos diseños presentan como ventaja que tienen lugar en situaciones reales. No obstante, la falta de un criterio claro que determine qué participantes en concreto están simulando limita la validez del estudio. El pertenecer a uno de los grupos no garantiza que se esté empleando dicho estilo de respuesta; tan solo puede afirmarse con cierto grado de seguridad que la prevalencia de la exageración de síntomas será mayor en un grupo frente al otro.

Esta incertidumbre es precisamente la principal limitación de estos estudios. La simple prevalencia diferencial en ambos grupos no permite que pueda establecerse la precisión de la medida, ni puntos de corte clasificatorios. Igualmente, limita el alcance y la interpretación de los resultados que se obtengan.

Si bien es un diseño en contextos reales que puede utilizarse de forma complementaria a otros diseños y para obtener unos resultados preliminares (sobre todo, en cuanto a diferencias de medias entre ambos grupos y tamaño del efecto), sus notables limitaciones hacen que sea necesario acudir a diseños diagnósticos de grupos conocidos o diseños de medidas autosuficientes a la hora de buscar resultados más sólidos en la investigación.

De hecho, el diseño de prevalencia diferencial ha recibido duras críticas como alternativa metodológica al estudio de la simulación. Rogers (2008a) ha considerado que investigar simulación con este diseño resulta similar a estudiar aspectos de la esquizofrenia o la depresión a partir de las diferencias entre varones y mujeres, dadas las prevalencias diferenciales en ambos sexos, afirmando que “this design attempts to address external validity by assuming that differences in types of referrals can be used as an expedient proxy for known groups. Such simplistic thinking should not be tolerated in clinical research, although it may play a marginal role in advancing theory” (Rogers, 2008a, p. 11).

2.1.5. Grupos mixtos

Los grupos mixtos fueron planteados inicialmente por Frederick (2000) como una alternativa al modelo rígido de los grupos diagnósticos conocidos, que exige que todos los sujetos de cada uno de los grupos esté íntegramente formado por sujetos de las características indicadas (solo simuladores, solo clínicos...). Según este autor, resulta más adecuado/ajustado considerar grupos mixtos, en los que la característica a analizar esté presente en una determinada proporción .

Con este objetivo, Frederick (2000) propone unas fórmulas alternativas para calcular los valores de verdaderos positivos y negativos, de acuerdo a las prevalencias estimadas en cada uno de los grupos. Un ejemplo de este tipo de diseños puede consultarse en la investigación de Crawford, Greene, Dupart, Bongar y Childs (2006). Si bien, esta modalidad apenas ha recibido atención en la investigación en simulación ni ha sido utilizada habitualmente.

2.2. Análisis estadísticos empleados habitualmente en estudios de simulación

Los tipos de análisis habitualmente empleados en las investigaciones sobre simulación son tres: 1) análisis de las diferencias entre grupos, 2) eficacia o precisión clasificatoria y 3) correlaciones con otros indicadores. Estos tres análisis resultan complementarios y suelen emplearse conjuntamente en el estudio de cualquier nueva medida, índice o criterio de simulación. A continuación, se señalarán muy sucintamente algunos aspectos específicos al campo de la simulación en el uso de los dos primeros tipos de análisis. El tercer tipo de análisis (correlaciones con otros indicadores) no presenta particularidades especiales, manteniendo las pautas habituales en otros diseños de investigación y validación de instrumentos.

2.2.1. Análisis de las diferencias entre grupos

Una primera y frecuente vía de análisis es estudiar si efectivamente los pacientes clínicos genuinos y los simuladores obtienen diferentes puntuaciones en la medida objeto de estudio. Para analizar esta diferencia se suelen emplear las técnicas habituales apropiadas al caso, tales como ANOVAs o MANOVAs, con el objetivo de establecer la significación estadística de las diferencias.

Adicionalmente, suelen considerarse distintos indicadores de tamaño del efecto, aunque el más asentado en el campo de la simulación y con criterios específicos de interpretación es la *d de Cohen* (Cohen, 1992). Así, Rogers et al. (2003) y, más recientemente Rogers (2008d), han propuesto rangos interpretativos específicos para la *d de Cohen* en el ámbito de la simulación. Estos rangos son notablemente más elevados que los sugeridos originalmente por Cohen (1992) ya que, según Rogers, este hecho deriva de la necesidad de ser más exigente para establecer un tamaño del efecto “grande” dadas las consecuencias que tiene para los sujetos evaluados el diagnóstico de simulación. Así Rogers (2008d) establece los siguientes rangos:

- Tamaño del efecto **moderado** si $d \geq 0,75$
- Tamaño del efecto **grande** si $d \geq 1,25$
- Tamaño del efecto **muy grande** si $d \geq 1,50$

Existe una íntima relación entre el tipo de diseño empleado y los tamaños del efecto que se esperan. De esta manera, es frecuente encontrar valores más altos en los diseños análogos que en los diseños en contextos reales (diseños de grupos conocidos, diseños de medidas autosuficientes o diseños de prevalencia diferencial). Esto es debido a que los simuladores análogos suelen ser más extremos en sus respuestas que los simuladores en contextos reales. Por lo tanto, en la interpretación de cualquier resultado estadístico se deberá tener en cuenta el diseño empleado como modulador de los valores que cabe encontrar y su relevancia.

2.2.2. Análisis de la precisión clasificatoria

A diferencia de los análisis anteriores, que ofrecen estadísticos globales para el conjunto de la muestra, en este caso se busca analizar el funcionamiento en el nivel individual, analizando la proporción de sujetos que son correctamente clasificados de acuerdo a los puntos de corte que se establezcan. Esto hace que los resultados obtenidos en estos análisis sean más claramente interpretables y vinculados a la labor práctica del profesional (Hunsley y Meyer, 2003).

Habitualmente, se emplean curvas COR (ROC en sus siglas en inglés) en estos estudios. Un primer paso en su análisis consiste en examinar el valor del área bajo la curva (ABC; AUC en sus siglas en inglés *Area Under the Curve*). El área bajo la curva (ABC) refleja la capacidad global de la prueba para discriminar en el conjunto de todos los posibles puntos de corte, siendo aquellos valores más altos indicativos de una mejor capacidad discriminativa.

Aplicada al tema que nos ocupa, el área bajo la curva puede ser descrita como la probabilidad de que un caso extraído aleatoriamente de la muestra de simuladores obtuviera una puntuación más alta que un caso extraído aleatoriamente de la muestra de participantes que respondieron sinceramente. De esta forma, valores del área bajo la curva superiores a 0,50 indicarían que el test permite clasificar los casos mejor que el azar y conforme la curva COR se aleje más de la diagonal del gráfico (línea de referencia de acierto por azar) proporcionará mayor precisión. Hosmer y Lemeshow (2000) proponen algunas pautas para la valoración del área bajo la curva:

- ABC = 0,50 Ausencia de discriminación
- ABC = 0,70 – 0,79 Nivel de discriminación aceptable
- ABC = 0,80 – 0,89 Nivel de discriminación excelente
- ABC \geq 0,90 Nivel de discriminación sobresaliente

Otros autores han establecido niveles interpretativos diferentes (Franco Nicolás y Vivo Molina, 2007):

- ABC = 0,70 – 0,91 Nivel de discriminación bueno
- ABC = 0,92 – 0,96 Nivel de discriminación muy bueno
- ABC \geq 0,97 Nivel de discriminación excelente

Adicionalmente a este valor global, se suelen calcular cinco índices para la estimación de la eficiencia diagnóstica de los puntos de corte posibles: sensibilidad (SN), especificidad (ES), tasas total de aciertos, poder predictivo positivo (PPP) y poder predictivo negativo (PPN). En todos los casos, estos índices se calculan a partir de una tabla de 2x2 en la que se compara la clasificación del nuevo test que se está estudiando, con el criterio de referencia que es considerado veraz (*gold standard*).

En la tabla 2.4 se incluye una descripción de las categorías, resultado del cruce de ambas variables.

Tabla 2.4
Clasificación de resultados de un estudio de validez de un nuevo test

		Criterio o <i>gold standard</i>		
		Presente	Ausente	
Resultado del test	Presente	A (verdaderos positivos)	B (falsos positivos)	A+B
	Ausente	C (falsos negativos)	D (verdaderos negativos)	C+D
		A+C	B+D	

- **Sensibilidad (SN):** Es la proporción de personas que tienen el atributo y que son detectadas por el test. Refleja lo sensible que es el test para detectar la presencia del atributo.

$$SN = A / A+C$$

Valores altos indicarán que la prueba es muy sensible a la presencia de simulación y que detectará a la mayoría de los simuladores, cometiendo un reducido número de falsos negativos. Así, es recomendable que los puntos de corte de *screening* tengan valores de sensibilidad cercanos a 0,90.

- **Especificidad (ES):** Es la proporción de personas que no tienen el atributo y son detectadas por el test como ausentes del atributo. Refleja lo específico que es el test y en qué medida no detecta a aquellos que realmente no están simulando sus síntomas.

$$ES = D / (B+D)$$

Valores altos indicarán que la prueba es muy específica y que excluirá correctamente a la mayoría de los pacientes clínicos genuinos, cometiendo un reducido número de falsos positivos. Así, es recomendable que los puntos de corte más diagnósticos tengan valores de especificidad cercanos a 0,90.

Como criterio general, West y Finch (1997) sugieren considerar como satisfactorios aquellos valores de sensibilidad o especificidad superiores a 0,70; aceptables los comprendidos entre 0,50 y 0,69 y bajos aquellos menores a 0,50.

- **Tasa total de aciertos:** Es la proporción de personas que son correctamente clasificadas por el test.

$$\text{Tasa total} = (A+D) / (A+B+C+D)$$

Ahora bien, incluso por mero azar se clasificarían correctamente al 50% de los casos, por lo que se suele calcular un índice de acuerdo corrigiendo la influencia del azar, siendo el más conocido el índice kappa de Cohen (κ ; (Cohen, 1960). Otros índices empleados en este sentido son el índice kappa ponderado (Kraemer, Periyakoil y Noda, 2002) o el coeficiente phi.

Los índices de sensibilidad y especificidad nos indican cuántos casos positivos son detectados por el test y cuántos casos negativos son correctamente excluidos, ofreciéndonos una lectura de los resultados desde el criterio o *gold Standard*.

Los índices de **poder predictivo positivo (PPP) y poder predictivo negativo (PPN)**, nos muestran el acierto clasificatorio desde el punto de vista del test bajo estudio, no desde el criterio.

- **Poder predictivo positivo (PPP):** Es la proporción de personas detectadas por el test como poseedoras del atributo y que realmente lo tienen.

$$\text{PPP} = A / (A+B)$$

- **Poder predictivo negativo (PPN):** Es la proporción de personas detectadas por el test como ausentes del atributo y que realmente no lo tienen.

$$\text{PPN} = D / (C+D)$$

Dada su estrecha relación con la calidad de la decisión del profesional, tanto el poder predictivo positivo, como negativo, (PPP y PPN) son considerados elementos relevantes para el profesional y deben ser informados, junto con la sensibilidad, la especificidad y el área bajo la curva, a la hora de establecer la precisión clasificatoria de un determinado punto de corte (Nicholson et al., 1997).

A diferencia de los valores de sensibilidad y especificidad, el poder predictivo positivo y negativo dependen de la prevalencia del atributo evaluado, esto es, del porcentaje de simuladores existentes en la población de referencia y no están determinados por el porcentaje de simuladores del estudio en concreto. De ahí la importancia de disponer de adecuadas estimaciones de la prevalencia de la simulación, tal y como se señalaba en el capítulo previo.

Suele ser habitual que, en las investigaciones, particularmente en las de análogos, se disponga de un 50% de simuladores ya que las pruebas estadísticas son más eficientes en estas circunstancias. Sin embargo, este valor no refleja la prevalencia de la simulación en muchos de los contextos reales. Por este motivo, para el cálculo del poder predictivo positivo y negativo se han de emplear una serie de fórmulas que se describen a continuación.

Para calcular el PPP y el PPN teniendo en cuenta una determinada tasa de prevalencia se pueden utilizar las siguientes fórmulas en las que p es la tasa de prevalencia, s la sensibilidad y e la especificidad (Baldessarini, Finklestein y Arana, 1983).

$$PPP = [(p) * (s)] / [(p) * (s) + (1-p) * (1-e)]$$

$$PPN = [(1-p) * (e)] / [(1-p) * (e) + (p)*(1-s)]$$

Con relación a qué tasas de prevalencia considerar en el cálculo, Sellbom, Toomey, Wygant, Kucharski y Duncan (2010) recomiendan obtener los valores de PPP y PPN para las tasas de prevalencia de 0,15, 0,30 y 0,50, esto es, para una prevalencia del 15%, del 30% y del 50% respectivamente. Posteriormente, el profesional que consulte los resultados podrá elegir qué tasa de prevalencia es más adecuada a su caso y qué valores se obtienen con dicha tasa de prevalencia. En caso de considerar solo una tasa de prevalencia Greve y Bianchini, (2004) recomiendan que se emplee la del 30% (0,30).

Los valores de PPP y PPN son directamente proporcionales a la prevalencia del atributo. Conforme mayor es la prevalencia, mayor será el valor de PPP y menor el de PPN; a la inversa, conforme menor sea la prevalencia, menor será PPP y mayor PPN. Cabe indicar como, incluso para una excelente especificidad y sensibilidad (0,90), el poder predictivo positivo es bajo en prevalencias inferiores al 30% (valores de 0,50 y 0,32 respectivamente). Este aspecto deberá ser tenido en cuenta a la hora de interpretar su magnitud.

3. Instrumentos para la detección de la simulación: el papel de los autoinformes y del Inventario Estructurado de Simulación de Síntomas, SIMS

Con la finalidad de determinar la capacidad de los psicólogos y otros profesionales teóricamente especializados, tales como policías, detectives o jueces, para detectar mentiras, Aamodt y Custer (2006) realizaron un meta-análisis a partir de 193 estudios en los que se sometió a estos profesionales a una situación artificial creada específicamente para este objetivo. Los resultados mostraron que los psicólogos apenas tenían una destreza ligeramente superior a los estudiantes que participaron en la investigación (62% acierto frente a 54%). Además, la confianza con la que una persona consideraba que podía detectar el engaño no guardaba relación con su capacidad real (correlaciones de 0,05 y 0,04, respectivamente en dos meta-análisis diferentes, Aamodt y Custer, 2006 y DePaulo, Charlton, Cooper, Lindsay y Muhlenbruck, 1997).

De hecho, diversas revisiones parecen apuntar a que son limitados los indicadores verbales, no verbales o fisiológicos relacionados de modo indudable con el engaño (Ben-Shakhar y Elaad, 2003; DePaulo et al., 2003; Kleiner, 2002; MacLaren, 2001; Masip, Garrido y Herrero, 2004; Masip, Sporer, Garrido y Herrero, 2005; Vrij, 2000), existiendo múltiples factores que inducen a la confusión. Por ejemplo, se suele considerar como claro marcador de mentira el no mantener la mirada o no realizar contacto visual, cuando la investigación parece avalar que éste no es un buen indicador (Vrij, 2004). De hecho, las diferencias culturales pueden confundirse fácilmente con indicadores de mentira o engaño. Así, investigaciones realizadas en Holanda han mostrado que la evitación de la mirada es mayor en holandeses procedentes de Turquía y Marruecos frente a holandeses nacidos en Europa (Vrij, Dragt y Koppelaar, 1992). Los estudios llevados a cabo por Vrij y su equipo han mostrado que las conductas típicas de algunos grupos étnicos que viven en Holanda son asociadas a conductas sospechosas por parte de los oficiales de policía nativos, un fenómeno que han denominado “errores de comunicación no verbal transcultural”.

Otros aspectos, como el nivel de extraversión o introversión, también suelen asociarse erróneamente al grado de verosimilitud de la información proporcionada. Así, las personas más expresivas suelen ser más creíbles por su estilo más espontáneo y directo, mientras que los introvertidos o las personas con ansiedad social, suelen crear impresiones menos creíbles. De esta manera, la impresión de tensión, nerviosismo o miedo, propia de las personas con ansiedad social o timidez, suele ser interpretada erróneamente como indicador de engaño (Riggio, 1986). De modo

análogo, las personas con una elevada autoconciencia pública suelen ser menos creíbles ya que son muy sensibles a la evaluación, por lo que en estos contextos suelen presentar unos niveles de activación superiores, independientemente de la veracidad de sus afirmaciones (Kashy y DePaulo, 1996).

Se han probado métodos muy diferentes para detectar a respondientes deshonestos, tales como el polígrafo, la hipnosis, las entrevistas asistidas con drogas y las entrevistas clínicas desestructuradas. No obstante, tal y como se señalaba anteriormente, ninguno de estos métodos se ha mostrado suficientemente fiable y objetivo en la detección de simulación. Por el contrario, los instrumentos psicométricos de autoinforme para la detección de simulación han mostrado una precisión razonable cuando son usados en combinación con otras fuentes de información, incluyendo observaciones conductuales, entrevista con familiares, consulta del historial y entrevista clínica (Capilla Ramírez y González-Ordi, 2009; González-Ordi y Gancedo Rojí, 1999; González-Ordi et al., 2012; Rogers, 2008c; Sellbom y Bagby, 2008).

En relación con la propia concepción que tiene el profesional sobre cómo proceder a detectar la simulación, Rogers y Bender (2003) han establecido una clasificación en tres estadios. El primer estadio es el enfoque “intuitivo”, que hace referencia al profesional que confía en que su criterio clínico es suficiente para la detección de la simulación. Este estadio está fuertemente asociado a la creencia de que la simulación es muy poco frecuente y que, además, cuando se da, es claramente manifiesta. Este es el motivo por el que este grupo considera que su criterio bastará para detectar que una persona está simulando. No obstante, como se ha comentado previamente, la evidencia empírica ha demostrado sistemáticamente la falibilidad del “ojo clínico”, la ausencia de aprendizaje de los errores y el escaso papel de la experiencia (a diferencia de la eficacia en el diagnóstico de otras áreas, aquí no hay retroalimentación del error en la detección de un simulador ya que éste difícilmente informará tras haber conseguido su objetivo). Tampoco es cierto que la simulación sea un hecho espurio o aislado, como se señaló en el capítulo 1 de esta tesis, ni que cuando se presenta es de forma obvia a cualquier evaluador, sino que frecuentemente la simulación se manifiesta de una manera sutil y sofisticada que resulta difícil de detectar al mero ojo clínico.

Continuando con la clasificación de Rogers y Bender (2003), el segundo estadio, más habitual, es el del profesional consciente de la necesidad de evaluar proactiva y sistemáticamente la posible presencia de sesgos de respuestas y de exageración de síntomas. En estos casos, los profesionales suelen hacer uso de tests generales y tradicionales, que incluyen medidas específicas de detección de exageración de síntomas y que les alertan de su presencia.

Finalmente, los profesionales que se encuentran en el tercer estadio, son aquellos que complementan el uso de estos tests globales con otras pruebas específicas de simulación, diseñadas precisamente con este objetivo, permitiendo refinar la interpretación de los resultados cuando es necesario. Este estadio es considerado como el más avanzado y el que permite una mayor precisión diagnóstica, aspecto clave en un área tan delicada y con consecuencias tan devastadoras, como puede ser la clasificación errónea de un paciente real como simulador.

La enorme falibilidad del juicio clínico y, sobre todo, los sesgos presentes en la evaluación de la credibilidad por aspectos ajenos a la propia área evaluada, hacen necesaria la evaluación detallada y protocolizada como pieza fundamental en la detección del engaño o en la determinación del no engaño. En este proceso juegan un papel esencial los autoinformes, como herramienta sistemática de recogida de información y toma de decisiones.

3.1. El papel de los autoinformes en la detección de simulación

Si bien en el proceso de detección de simulación se requiere incluir distintas fuentes de información y métodos (entrevistas, tests, observación y fuentes colaterales) (González-Ordi y Gancedo Rojí, 1999; González-Ordi et al., 2012; McDermott y Feldman, 2007; Rogers et al., 2010), los instrumentos de evaluación mediante autoinforme ocupan un lugar destacado, acaparando una importante parte de la investigación realizada en simulación y situando su capacidad predictiva como una cuestión esencial en contextos científicos y jurídicos (Farkas et al., 2006; Sweet y Guidotti Breting, 2013).

Desde el principio de la evaluación psicológica con autoinforme, se desarrollaron índices y categorías de distorsión de respuestas que pudieran informar sobre la “validez” del protocolo en relación con la exageración de síntomas (Cronbach, 1946). Hathaway y McKinley (1943) fueron pioneros en esta línea, iniciando toda una tendencia de evaluación de estilos de respuestas que se ha reproducido en casi todas las pruebas creadas con posterioridad.

Específicamente, lo que permiten evaluar los autoinformes en relación con la simulación es la presencia de un patrón de exageración de síntomas. El patrón de exageración de síntomas se puede definir como aquel estilo de respuesta en el que el sujeto busca presentarse como alterado o deteriorado, exagerando o fabricando síntomas con este fin y maximizando las características negativas de su supuesto malestar (Baer, Rinaldo y Berry, 2003).

Con el objetivo de detectar este patrón de respuesta, se han ido desarrollando escalas específicas dentro de los autoinformes, que buscan identificar aquellos casos en los que el sujeto

sistemáticamente se presenta de forma poco favorable, exagerando o sobredimensionando deliberadamente sus síntomas psicopatológicos.

Tal y como se señalaba en el primer capítulo, las vías principales empleadas para este fin han sido las basadas en la detección de síntomas inverosímiles y las basadas en la detección de síntomas amplificadas en su intensidad, duración o gravedad (Rogers et al., 2010). Estas dos grandes vías se concretan en estrategias específicas de detección. Dentro de la categoría de inverosimilitud de los síntomas, se encuentran las estrategias de síntomas infrecuentes, síntomas casi-infrecuentes, combinación de síntomas, síntomas improbables y patrones espurios de psicopatología. Dentro de la categoría de amplificación de los síntomas, están las estrategias de asignación indiscriminada de síntomas, severidad de los síntomas, síntomas obvios, problemas psicológicos comunes, estereotipos erróneos y síntomas observados frente a los informados (Rogers et al., 2010). Puede consultarse el apartado 1.4 de esta tesis para una descripción más detallada de cada una de estas estrategias.

En los siguientes apartados se señalarán algunos aspectos a tener en cuenta en el uso de las escalas de exageración de síntomas de los autoinformes y su investigación, tales como la relatividad de los puntos de corte, la importancia del contexto, los factores que afectan a su sensibilidad o la relación entre exageración de síntomas y simulación.

3.1.1. La evaluación de la exageración de síntomas: puntos de corte y contexto

El análisis de las puntuaciones obtenidas en las escalas de exageración de síntomas persigue determinar la presencia o ausencia de exageración de síntomas. En este proceso “diagnóstico” es sumamente importante tener en cuenta distintos aspectos, entre ellos, la concepción de la exageración como un continuo, el uso de puntos de corte y el contexto de evaluación.

La exageración de síntomas debe concebirse como un continuo en el que no hay un punto exacto en el que el sujeto pase de responder sinceramente a exagerar notablemente sus síntomas, sino más bien tenemos que hablar de probabilidades y niveles de intensidad detectados.

En relación con este aspecto, el uso de puntos de corte, si bien es algo que ayuda al profesional en la toma de decisiones, ha de considerarse como una primera aproximación. La aplicación de puntos de corte automáticos para determinar la exageración de síntomas rara vez es una estrategia adecuada (Rogers y Bender, 2012), de hecho Rogers, Gillard, Wooley y Ross (2012) califican la fe ciega en un punto de corte como el *mito de la precisión láser de las puntuaciones de corte*.

Más que la clasificación dicotómica, la investigación ha mostrado que resulta más conveniente graduar en distintos niveles la certeza que podemos tener sobre la presencia de una distorsión de las respuestas. En función de la calidad y la magnitud de la evidencia disponible de exageración de síntomas se podrá informar sobre la probabilidad de que exista (alta, media, baja) y sobre el grado en que aparece (Berry y Nelson, 2010). No en vano, los criterios de simulación de Bianchini et al. (2005) o Slick et al. (1999) emplean gradientes de certeza: posible, probable, definitiva.

Adicionalmente, los puntos de corte también han de considerarse contextualmente ya que numerosas variables, entre ellas el contexto, ayudarán a determinar qué constituye una elevación en la escala. Así, por ejemplo, será más significativa una elevación en una escala de exageración de síntomas psicóticos si aparece en un caso sin patología psiquiátrica previa o relacionada (p. ej., dolor crónico, trastornos afectivos o reclamación de daños tras accidente de tráfico) que si aparece en una unidad psiquiátrica.

El contexto de evaluación determinará, no solo el posible rango de puntos de corte o la significación de una cierta elevación en una escala, sino también qué escalas serán más adecuadas para el caso. Las diferentes escalas de exageración de síntomas disponibles en los distintos autoinformes no son plenamente intercambiables entre sí, ni todas valen para los mismos contextos. Las escalas varían en cuanto a los contextos en que son más efectivas, a su sensibilidad a determinados patrones de exageración de síntomas y a su resistencia al entrenamiento o a la presencia de simulación sofisticada (Santamaría y González-Ordi, 2012).

3.1.2. Factores que afectan a la sensibilidad de las escalas de exageración de síntomas

En general, la evidencia disponible muestra que cuando el sujeto presenta un estilo global de simulación, en el que se ven afectadas diversas áreas sintomáticas y muestra un patrón de afectación severa, no será necesario que el profesional tenga que considerar múltiples escalas ya que habitualmente estarán correlacionadas entre sí y la información será redundante (Greene, 2008). En los casos en los que el sujeto se centre en la simulación de una psicopatología específica, sí es probable que tenga que considerarse aquella escala referida a la sintomatología alegada y valorar desde esta óptica las divergencias entre las puntuaciones de las distintas escalas de exageración.

Este patrón de simulación global o específico también se reflejará en el número y diversidad de elevaciones que se encuentren en las escalas clínicas de la prueba. Un perfil con elevaciones en

determinadas escalas clínicas solo podrá ir acompañado de un perfil selectivo de exageración de síntomas, mientras que un perfil con múltiples elevaciones podrá estar relacionado con un patrón más global de exageración.

Las investigaciones llevadas a cabo parecen indicar que las escalas de exageración de síntomas detectan de forma bastante efectiva la simulación de cuadros psicopatológicos extensos y severos y encuentran más dificultades cuando se trata de síntomas más específicos y menos graves. Esto se halla vinculado al nivel de sofisticación del simulador, que sea más global e indiscriminado o más específico y selectivo. También guarda relación con el contexto de evaluación ya que es más frecuente encontrar un amplio patrón de exageración en contextos forenses criminales que en contextos civiles, donde los sujetos se focalizan más en el trastorno o problema alegado (Wygant et al., 2007, 2010).

El hecho de que los sujetos hayan sido entrenados específicamente para evitar ser detectados por las escalas de exageración afecta de forma sensible a su efectividad. El estudio de Sellbom y Bagby (2010) ilustra perfectamente este punto, mostrando como la capacidad para detectar simuladores análogos puede reducirse prácticamente a la mitad cuando los sujetos han sido entrenados para no ser detectados (sensibilidad escala F-r: 0,78 en sujetos sin entrenamiento; 0,40 en sujetos entrenados). Este mismo resultado se ha hallado en los diversos estudios que han abordado esta cuestión (p.ej., Rogers, Sewell, et al., 1996; Efendov, 2007).

Esta sensibilidad a su entrenamiento alerta de dos cuestiones a tener en cuenta. Por una parte, la necesidad de manejar con extrema confidencialidad toda la información relacionada con las estrategias, escalas e ítems empleados en la exageración de síntomas. Internet puede convertirse en un gran aliado o en un terrible enemigo que inutilice nuestras herramientas, como mostró el estudio de Ruiz et al. (2002) en el contexto específico de las escalas de detección de sesgo de respuesta. Por otra parte, ha de investigarse más en la creación de futuras escalas más resistentes al entrenamiento y a su difusión.

En relación con la instrucción específica sobre los síntomas que componen un determinado cuadro psicopatológico, no parece que esta información ayude a evitar la detección (Efendov, 2007; Rogers, Bagby y Chakraborty, 1993). Esto es, la instrucción a los sujetos sobre qué síntomas deben fingir acordes con el trastorno alegado, no parece afectar de forma notable a la capacidad de las escalas de exageración para detectar a estos sujetos.

Otros factores adicionales parecen influir también en las escalas para detectar la presencia de simulación. Así, ciertos estudios apuntan a la relación entre una mayor inteligencia y una menor probabilidad de ser detectado (Pelfrey, 2004; Steffan, Kroner y Morgan, 2007) o se ha indagado incluso el papel de la inteligencia emocional en la capacidad para distorsionar los resultados en las pruebas (Grieve y Mahar, 2009). No obstante, estas evidencias son todavía preliminares y no existe un suficiente cuerpo empírico que defina claramente su influencia o su tamaño del efecto en la probabilidad de detección.

3.1.3. Exageración de síntomas y simulación

Tal y como afirma Iverson (2010), se puede considerar que la exageración de síntomas es un aspecto central de la simulación, pero no es sinónimo de esta. Es decir, la simulación requiere la existencia de exageración de síntomas, pero puede haber exageración de síntomas sin que ello implique simulación. Tal y como es definida actualmente la simulación en el DSM-5 (véase el capítulo 1), se requiere, aparte de la presencia de la fabricación o exageración de síntomas, que haya una intencionalidad y que esté dirigida a obtener beneficios externos (American Psychiatric Association, 2013; Greene, 2008).

Son múltiples las causas que pueden subyacer a la exageración de síntomas (véase la tabla 3.1), siendo solamente algunas de ellas relativas propiamente al concepto de simulación. Será la integración de los resultados del autoinforme con toda la información disponible sobre el sujeto lo que determine si el patrón de exageración de síntomas está vinculado a simulación o, por el contrario, responde a otras causas. Las entrevistas con el sujeto y familiares, la revisión de los historiales y otras fuentes de información, nos pueden indicar qué aspectos motivacionales pueden estar presentes, tales como culpabilización, ira, necesidad de reconocimiento, resentimiento o victimización.

Los indicadores de exageración de síntomas de los autoinformes no pueden considerarse, por lo tanto, indicadores directos de simulación, sino indicadores de una presentación no realista o magnificada de los síntomas de los que informa el sujeto. Rogers, Jackson y Kaminski (2005) ilustraron el hecho de que pueden aparecer puntuaciones altas en los índices de exageración de síntomas, tanto en participantes instruidos para simular una incapacidad y obtener la baja laboral, como en aquellos que pretendían mostrarse afectados para conseguir la atención de familiares y amigos ocupando el rol de enfermo.

Tabla 3.1
Factores e incentivos relativos a la exageración de síntomas (basado en Iverson, 2007)

Búsqueda de reconocimiento de derechos (p.ej., llamar la atención del profesional a sus molestias)	Elusión de responsabilidades
Dependencia	Búsqueda de la atención de parejas o amigos
Ira o frustración	Obtención de atención médica especial
Somatización	Obtención de medicación
Patrón de reforzamiento del comportamiento de exageración	Elusión de acusaciones criminales
Pensamiento depresivo negativista	Rebaja de sentencias o influencia en las mismas
Características de personalidad (pensamientos catastrofistas) o de los trastornos	Obtención de pensiones de incapacidad
Efecto nocebo	Obtención de compensaciones por lesiones personales
Sesgo de “viejos tiempos mejores”	Venganza hacia la persona causante del daño

Tampoco estos indicadores pueden descartar que exista psicopatología subyacente. Drob et al. (2009) ponen, como ejemplo, la anécdota de un paciente que, de acuerdo a la evaluación con distintas pruebas, claramente estaba fingiendo un trastorno amnésico. No obstante, cuando se analizó su historial, se advirtió que había sido hospitalizado previamente por paranoia en diversas ocasiones y se descubrió que la simulación de la amnesia tenía el objetivo de no comunicar nada a los psicólogos ya que los creía espías del gobierno que le perseguían. Otros ejemplos pueden ser los pacientes con estados depresivos profundos, que tienden a evaluar su situación de manera extremadamente negativa, o los pacientes con tumor parietal derecho, que refieren síntomas raros (indicativos de exageración de síntomas), como que su mano derecha no es suya (Faust y Ackley, 1998).

Las escalas de exageración de síntomas informan sobre la credibilidad de los síntomas psicológicos, somáticos y cognitivos que el sujeto refiere; ahora bien, no determinan cuál es el objetivo del sujeto, si la conducta es intencional o si este patrón de exageración se presenta junto con patología genuina. La presencia de exageración de síntomas nos indica que la presentación que la persona evaluada hace de sí misma y la información que nos proporciona no son fiables ni precisas y que es probable que estén distorsionadas negativamente. Nos alerta, por lo tanto, de que *la información dada es probablemente más negativa que la que daría un observador objetivo externo.*

De hecho, existe una marcada corriente de investigadores y profesionales que consideran que la evaluación debería restringirse a informar sobre la credibilidad de los síntomas, sin determinar su intencionalidad o sus objetivos (Boone, 2007a; Delis y Wetter, 2007). Esta corriente está basada

en la imposibilidad práctica de determinar fehacientemente el concepto de “intencionalidad” o volición, o en su carácter especulativo, cuando menos (Pankratz y Erickson, 1990). Berry y Nelson (2010) sugerían que el diagnóstico de simulación se centrara en la presencia de síntomas fingidos, sin intentar inferir su carácter consciente o intencional o la meta pretendida por el evaluado. Con este cambio, se eliminaría la necesidad de hacer juicios marcadamente subjetivos e imprecisos sobre la intención y los objetivos del evaluado, como requisito para el diagnóstico de simulación. En esta línea, resultan interesantes también las reflexiones de Merten y Merckelbach (2013) sobre la presencia de síntomas no creíbles y el posterior diagnóstico diferencial entre trastornos facticios, somatomorfos o simulación. Desgraciadamente, ninguno de estos cambios han sido incluidos en la nueva edición del DSM-5 (American Psychiatric Association, 2013), que se ha mantenido inmutable con relación a ediciones previas (American Psychiatric Association, 2007).

3.2. Principales autoinformes empleados en la detección de exageración de síntomas

Las pruebas principales que han recabado, a escala internacional, la atención de los investigadores en simulación y que han acumulado un conjunto mayor de evidencia favorable han sido el Inventario Multifásico de Personalidad de Minnesota, MMPI-2 y el Inventario de Evaluación de la Personalidad, PAI (Rogers et al., 2009; Rogers, 2008c), así como la versión más reciente del MMPI, el MMPI-2-RF (Ben-Porath y Tellegen, 2008). Estas pruebas de personalidad son las más empleadas en psicología forense (Archer, Buffington-Vollum, Stredny y Handel, 2006) y son consideradas mayoritariamente por los profesionales como recomendables o aceptables en la evaluación de simulación (Lally, 2003).

El MMPI-2 es una prueba ampliamente utilizada en lengua española (Echeburúa et al., 2003; Muñiz Fernández y Fernández Hermida, 2010). Su versión más reciente, el MMPI-2-RF (Ben-Porath y Tellegen, 2008), ha supuesto una innovación y mejora, que está generando múltiples investigaciones sobre su uso en la detección de simulación (v.g., Burchett y Ben-Porath, 2010; Gervais, Ben-Porath, Wygant y Sellbom, 2010; Gervais, Lees-Haley y Ben-Porath, 2007; Goodwin, Sellbom y Arbisi, 2013; Sellbom et al., 2010; Sellbom, Wygant y Bagby, 2012; Wygant et al., 2009, 2011).

El *Inventario de evaluación de la personalidad* (PAI; Morey, 2007, 2011), es una prueba empleada en la evaluación de psicopatología, personalidad y otros aspectos clínicos relevantes. Durante los últimos años ha llegado a ser habitual su uso en distintos ámbitos, entre ellos, el forense. Se compone de 8 escalas relacionadas con la validez de los síntomas, además de 11 escalas clínicas, 5 escalas relacionadas con el tratamiento y 2 escalas interpersonales. La capacidad de sus escalas en la

detección de exageración de síntomas ha sido estudiada en múltiples investigaciones, respaldando su utilidad y eficiencia (v.g., Bagby, Nicholson, Bacchiochi, Ryder y Bury, 2002; Blanchard et al., 2003; Boccaccini, Murrie y Duncan, 2006; Edens, Poythress y Watkins-Clay, 2007; Gaines, 2009; García Silgo y Robles Sánchez, 2010; Hawes y Boccaccini, 2009; Hill, 2009; Hopwood, Morey, Rogers y Sewell, 2007; Hopwood, Orlando y Clark, 2010; Kucharski, Toomey, Fila y Duncan, 2007; Laffoon, 2009; Lange, Sullivan y Scott, 2010; Mogge, LePage, Bell y Ragatz, 2010; Rogers et al., 2012; Rogers, Sewell, et al., 1996; Rogers, Ornduff y Sewell, 1993; Whiteside, Dunbar-Mayer y Waters, 2009).

En los próximos apartados se describirán los principales autoinformes generales existentes y posteriormente se describirán aquellas pruebas de *screening* específicamente desarrolladas para simulación. Se finalizará, analizando en detalle las investigaciones disponibles sobre el Inventario Estructurado de Simulación de Síntomas, SIMS.

3.2.1. MMPI-2

El MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen y Kaemmer, 1989) es uno de los cuestionarios más ampliamente utilizados para evaluar alteraciones psicopatológicas en el ámbito clínico (Muñiz Fernández y Fernández Hermida, 2010). Asimismo, su prevalencia en el ámbito forense queda manifiesta en las palabras de Archer, Buffington-Vollum, Stredny y Handel (2006): “the literature base is not sufficient to identify a clear consensus regarding test usage patterns, with the exception of the MMPI–2 and the Wechsler Intelligence scales which appear popular across a wide variety of evaluations” (p. 85).

Una de los aspectos que ha influido en su difusión y uso son precisamente sus escalas e índices para evaluar las actitudes de las personas evaluadas, tanto de exageración de síntomas como de minimización. Estos indicadores de validez han demostrado su utilidad de forma sistemática en múltiples investigaciones (puede verse una síntesis en Greene, 2008 y un completo estudio empírico con la versión española en Sánchez Crespo, 2002).

En concreto, en relación con la exageración de síntomas, Rogers et al. (2010) afirman que ha habido más estudios que han examinado el fingimiento de síntomas en el MMPI-2, que en el conjunto de todas las otras medidas combinadas. Esto ha hecho que sea la herramienta más recomendada por los psicólogos y neuropsicólogos forenses en la evaluación de simulación (Lally, 2003).

Son diversas las escalas específicas de exageración de síntomas que permite obtener el MMPI-2. Dentro de ellas, hay escalas de objetivo más genérico y otras desarrolladas para detectar aspectos específicos. En la tabla 3.2 se ofrece un breve resumen de las principales escalas que han

recabado una mayor cantidad de investigación. Para un análisis más detallado de cada una de ellas puede consultarse Santamaría y González Ordi (2012).

3.2.2. MMPI-2-RF

El *Inventario Multifásico de Personalidad de Minnesota-2 Reestructurado* (MMPI-2-RF; Ben-Porath y Tellegen, 2008; Santamaría Fernández, 2009) es una nueva versión del MMPI compuesta por 338 ítems, seleccionados de entre los 567 ítems originales del MMPI-2. Su objetivo fue completar el trabajo que se había iniciado con las escalas clínicas reestructuradas (Tellegen et al., 2003), desarrollando un conjunto sustantivo de escalas que capturaran las variables clínicas subyacentes a los ítems del MMPI, de forma que permitieran una evaluación más eficiente, centrada en el contenido, coherente en su estructura y organización y con mejores propiedades psicométricas de fiabilidad y validez.

Está compuesto por 42 escalas sustantivas que se engloban en cuatro grandes categorías (escalas de segundo orden, escalas clínicas reestructuradas, escalas de problemas específicos, escalas de intereses y las cinco de personalidad psicopatológica) y por 8 escalas de validez que abarcan la detección de patrones inconsistentes de respuestas, exageración de síntomas y minimización o disimulación.

Una amplia cantidad de estudios avalaron su publicación inicial (Tellegen y Ben-Porath, 2008) y diversas investigaciones posteriores han seguido aportando información sobre el funcionamiento de la prueba en relación con múltiples campos, como los trastornos afectivos (Watson, Quilty y Bagby, 2011), la conducta antisocial (Aslan, 2010), el dolor crónico (Gironda y Clark, 2009), los problemas cognitivos (Gervais, Ben-Porath y Wygant, 2009), el abuso de sustancias (Moran, 2011), el trastorno de estrés postraumático (Arbisi, Polusny, Erbes, Thuras y Reddy, 2011) las quejas somáticas (Thomas y Locke, 2010), la epilepsia (Locke y Thomas, 2010; Locke et al., 2010), el maquiavelismo (Ingram, Kelso y McCord, 2011), los trastornos de personalidad (Ayearst, 2011; Smith, 2010), la evaluación en guardia y custodia de hijos (Archer, Hagan, Mason, Handel y Archer, 2012), su equivalencia transcultural (Khouri, 2011; Moran, 2011; Poizner, 2010; Shemer, 2011) o su convergencia con diversos criterios externos en población universitaria (Forbey, Lee y Handel, 2010).

Tabla 3.2
Escalas de exageración de síntomas del MMPI-2

Abrev.	Nombre	Descripción
F	Infrecuencia	<p>La escala F fue creada siguiendo la estrategia de síntomas casi-infrecuentes, seleccionando para ello un conjunto de 60 ítems que eran infrecuentes en los sujetos adultos de la muestra normal de tipificación original. El contenido de sus ítems incluye síntomas extravagantes o inusuales de psicopatología severa, principalmente de tipo psicótico. Aproximadamente una tercera parte de los ítems reflejan ideación paranoide o procesos psicóticos; otros grupos, de 8 a 12 ítems, reflejan apatía esquizoide y falta de implicación, cinismo y actitudes antisociales, hostilidad familiar y síntomas somáticos.</p> <p>Un individuo que presenta puntuaciones altas en F está reconociendo un amplio rango de síntomas psicopatológicos que son infrecuentes en la población general, reflejando con ello una tendencia a la exageración de los problemas o una presencia real y notable de patología grave. Así, esta escala es sensible, tanto a la exageración de síntomas, como a la presencia de psicopatología severa en forma de psicosis o perturbaciones destacadas de la personalidad, presentes con cierta prevalencia en contextos psiquiátricos (Rogers, 2008b) y penitenciarios (Wise, 2009).</p>
Fb	F posterior	<p>La escala Fb consta de 40 ítems seleccionados con similar criterio que en la escala F, esto es, que fueran respondidos por menos de un 10% de la muestra original de adultos normales. A diferencia de la escala F, cuyos ítems aparecen entre los primeros 370 del cuestionario, los ítems de Fb aparecen en la parte final del cuestionario ya que casi la mitad de sus ítems están entre los 100 últimos y todos después del ítem 280. Se quería complementar, de este modo, la medida de F con una segunda escala de infrecuencia centrada en la parte "posterior" de la prueba.</p> <p>El contenido de sus ítems es sustancialmente diferente del de la escala F; en vez de ítems relacionados con la patología psicótica, los ítems de la escala Fb se relacionan con trastornos afectivos graves. Su contenido refleja principalmente pánico o miedo (9 ítems), depresión y baja autoestima (8 ítems), suicidio o autolesión (4 ítems), abuso de alcohol (4 ítems) y separación familiar (3 ítems). Solamente tres ítems hacen referencia a procesos psicóticos.</p>
F-K		<p>Este índice se obtiene substrayéndole a la puntuación directa de la escala F la puntuación directa de la escala K (F menos K). Si el índice es positivo a partir de un determinado punto de corte, el sujeto evidenciaría una tendencia a la exageración deliberada de síntomas; si el índice es negativo a partir de un determinado punto de corte, el individuo mostraría una tendencia a la negación o encubrimiento de síntomas (disimulador).</p>
Fp	F psicopatológica	<p>La escala Fp (F psicopatológica) fue creada con el objetivo de ser aplicada en contextos caracterizados por una alta tasa de patología severa y discriminar aquellos protocolos con exageración de síntomas de los que presentaban psicopatología grave (Arbisi y Ben-Porath, 1995).</p> <p>Para ello, sus 27 ítems fueron seleccionados por su infrecuencia en muestras clínicas de pacientes psiquiátricos (porcentajes de respuestas inferiores al 20%), lo que hace menos sensible esta escala a cuadros psicopatológicos graves.</p> <p>Su contenido es bastante heterogéneo porque incluye amplias áreas de contenido patológico; solo podrían distinguirse dos pequeños subconjuntos, uno referido a ideas persecutorias y otro a grave hostilidad familiar.</p>
FBS		<p>La escala FBS está compuesta por 43 ítems y fue desarrollada originalmente por Lees-Haley, English y Glenn (1991) para identificar aquellos patrones de síntomas no creíbles en evaluados implicados en juicios en los que reclamaban compensaciones por lesiones sufridas o solicitudes de incapacidad.</p>
Ds-r		<p>La escala Ds fue creada con el objetivo de discriminar entre aquellos sujetos que estaban fingiendo un cuadro de tipo neurótico de los que genuinamente lo padecían. Para ello se seleccionó un conjunto de 32 ítems que reflejaran estereotipos erróneamente asignados a las personas neuróticas pero que realmente no les caracterizaban. Este conjunto de estereotipos permitiría, por lo tanto, distinguir a simuladores de pacientes clínicos.</p>

Específicamente, sus escalas de exageración de síntomas -que fueron revisadas con el objeto de mejorar su sensibilidad a la exageración de síntomas psiquiátricos, emocionales, somáticos y cognitivos- han centrado la atención de distintas investigaciones que han analizado su uso y eficacia en contextos variados, tanto civiles como criminales (Gervais et al., 2010; Goodwin et al., 2013; Harp, Jasinski, Shandera-Ochsner, Mason y Berry, 2011; Marion, Sellbom y Bagby, 2011; Rogers, Gillard, Berry y Granacher, 2011; Santamaría, González-Ordi, Capilla, Blasco y Pallardo Durá, 2010; Sellbom y Bagby, 2010; Sellbom et al., 2010; Weiss, Bell y Weiss, 2009; Wygant et al., 2010, 2011, 2009; Youngjohn, Wershba, Stevenson, Sturgeon y Thomas, 2011).

Son diversas las escalas específicas de exageración de síntomas que permite obtener el MMPI-2-RF. En la tabla 3.3. se ofrece un breve resumen de las escalas que lo componen -para un análisis más detallado de cada una de ellas puede consultarse Santamaría y González Ordi (2012)-.

De las cuatro escalas de exageración de síntomas, tres de ellas (F-r, Fp-r y FBS-r) son versiones revisadas de sus originales del MMPI-2. Durante el proceso de elaboración del MMPI-2-RF, estas tres escalas fueron rediseñadas con el objetivo de igualar o mejorar su sensibilidad a la exageración de síntomas de tipo psicopatológico, somático y cognitivo.

A este conjunto de tres escalas se le añadió una cuarta, Fs (Quejas somáticas infrecuentes) diseñada para identificar específicamente aquellos síntomas somáticos que eran raramente indicados por pacientes en tratamiento médico y que, por lo tanto, eran posibles indicadores de exageración de síntomas somáticos.

Dada su complementariedad, este conjunto de cuatro escalas permite cubrir adecuadamente las tres grandes áreas de exageración de síntomas posibles: psicopatológica, cognitiva y somática. Esto permite una evaluación comprehensiva de la presencia de exageración de síntomas (Wygant et al., 2009). Según el contexto de evaluación, podrá resultar más relevante la información proporcionada por una u otra escala.

La escala F-r es un buen punto de partida para la detección general de exageración de síntomas. Posteriormente, dependiendo del contexto de evaluación y los trastornos alegados será recomendable la consulta de las puntuaciones de Fp-r o de Fs y FBS-r.

Tabla 3.3
Escalas de exageración de síntomas del MMPI-2-RF

Abrev.	Nombre	Descripción
F-r	Infrecuencia	La escala F-r sirve como un indicador general de exageración de síntomas psicológicos. Está compuesta por 32 ítems que son respondidos infrecuentemente en la dirección de la escala por la muestra normativa (porcentajes de respuesta en la dirección de la escala inferiores al 10%). Elimina la antigua distinción entre F y Fb, sintetizando ambas en una única escala, F-r, que está compuesta por ítems de las distintas partes del cuestionario.
Fp-r	Psicopatología infrecuente	La escala Fp-r (Psicopatología infrecuente) es un indicador de exageración de síntomas psicopatológicos graves. Esta escala fue desarrollada originalmente por Arbisi y Ben-Porath (1995) para complementar a la escala F, cuyas puntuaciones podían verse elevadas por la presencia de trastornos psicopatológicos graves. La escala Fp pretendía solucionar este problema seleccionando ítems infrecuentes, no en la muestra normativa como la escala F, sino en muestras con problemas psicopatológicos. Así, a diferencia de lo que sucede en F-r, las puntuaciones de Fp-r tienden menos a confundirse con trastornos o problemas graves.
Fs	Quejas somáticas infrecuentes	La escala Fs se utiliza para identificar a sujetos evaluados que informan de un elevado número de quejas somáticas raramente alegadas por pacientes en tratamiento médico. Esta escala fue introducida en el MMPI-2-RF siguiendo la misma estrategia de infrecuencia, pero aplicada, en este caso, a quejas somáticas. Wygant, Ben-Porath y Arbisi (2004) desarrollaron esta escala a partir de la identificación de 16 ítems que describían síntomas somáticos infrecuentes en dos grandes muestras de pacientes en tratamiento médico y en una muestra de pacientes con dolor crónico.
FBS-r	Validez de los síntomas	La escala FBS-r está formada por 30 ítems que forman un subconjunto dentro del conjunto completo de 43 ítems que componían la versión de esta escala en el MMPI-2 (FBS). Diseñada originalmente como complemento a la escala F del MMPI-2, tiene como objetivo identificar individuos con síntomas no creíbles en el contexto de litigios civiles o solicitudes de incapacidad. Esta escala fue construida a partir de criterios empíricos y racionales buscando detectar una doble estrategia de simular un alto grado de afectación debido a la lesión y una minimización de los problemas emocionales y de personalidad previos al accidente o lesión; en suma, una combinación de exageración y minimización presente habitualmente en los contextos de reclamaciones por accidentes o lesiones o solicitudes de incapacidad (Greiffenstein, Fox y Lees-Haley, 2007)

En los últimos años, ha surgido una nueva escala, RBS, que ha evidenciado complementar a F-r, Fs y FBS-r en la evaluación de quejas cognitivas en contextos civiles y criminales (Wygant et al., 2010). La escala RBS (*Response Bias Scale*) está compuesta por 28 ítems y fue creada con la intención de identificar a aquellos sujetos que eran detectados como simuladores de problema de memoria en pruebas específicas (Gervais, Ben-Porath, Wygant y Green, 2007). Mediante diferentes análisis de regresión múltiple se seleccionaron aquellos ítems que mejor predicaban las puntuaciones en los tests de simulación de deterioro cognitivo como el TOMM (*Test de simulación de problemas de memoria*;

(Tombaugh, 1996), el WMT (*Word Memory Test*; Paul Green, Lees-Haley y Allen, 2003) o el *Computerized Assessment of Response Bias* (Allen, Conder, Green y Cox, 1997).

La escala RBS se ha mostrado como el mejor predictor del MMPI-2 de exageración de síntomas cognitivos en diferentes estudios realizados (Dionysus, Denney y Halfaker, 2010; Gervais et al., 2010; Lange et al., 2010; Rogers, Gillard, et al., 2011; Whitney, Davis, Shepard y Herman, 2008; Wygant et al., 2010). Además, las investigaciones disponibles parecen indicar que la existencia de déficits reales de memoria no producen elevaciones en esta escala (Gervais et al., 2010).

3.2.3. PAI

El *Inventario de evaluación de la personalidad* (PAI; Morey, 2007) es un test ampliamente empleado en contextos clínicos, médicos y forenses (Aikman y Souheaver, 2008; Archer et al., 2006; Belter y Piotrowski, 2001; Clark, Oslund y Hopwood, 2010). Está compuesto de 344 ítems y proporciona resultados en 11 escalas clínicas, 5 escalas relacionadas con el tratamiento y 2 escalas interpersonales. Se corresponde, en gran medida, con las agrupaciones psicopatológicas del DSM, subdividiéndose cada escala clínica en distintas subescalas, que se centran en aspectos específicos dentro del amplio cuadro del síndrome.

En los últimos años, se han ido acumulando evidencias que respaldan su uso (Duellman y Bowers, 2004; Edens, Cruise y Buffington-Vollum, 2001; Edens, Ruiz y Staff, 2005; Morey y Quigley, 2002; Morey, Warner y Hopwood, 2007) y su adaptación española ha sido publicada recientemente (Ortiz-Tallo, Santamaría, Cardenal y Sánchez, 2011)

Entre las ventajas potenciales del PAI suelen destacarse las siguientes:

- a) menor tiempo de aplicación;
- b) requiere un menor nivel de comprensión lectora;
- c) ausencia de solapamiento de ítems entre las escalas;
- d) uso de escala Likert de 4 puntos;
- e) escalas coherentes con la actual terminología diagnóstica;
- f) facilidad de interpretación.

El PAI incluye también varias escalas e índices específicos para la evaluación de exageración de síntomas. En la tabla 3.4 se recoge una breve descripción de cada uno de ellos.

Sellbom y Bagby (2008) recomendaron un interesante esquema para determinar la sospecha de simulación, que a continuación detallamos, adaptándolo a los puntos de corte sugeridos por Hawes y Boccaccini (2009) y por la investigación realizada en su adaptación española (Ortiz-Tallo et al., 2011):

1. Baja probabilidad de exageración de síntomas

IMN T < 73, SIM T < 82 y baja puntuación en FDR.

2. Sospecha de exageración de síntomas.

IMN T ≥ 92, SIM T ≥ 94 o alta puntuación en FDR.

3. Alta probabilidad de exageración de síntomas

IMN T ≥ 110, SIM T ≥ 107 o puntuación muy alta en FDR (≥ 1,80)

Tabla 3.4
Escalas de exageración de síntomas del PAI

Abrev.	Nombre	Descripción
IMN	Impresión negativa	<p>La escala IMN (Impresión negativa) suele ser la escala inicial para valorar la presencia de exageración de síntomas. IMN está compuesta por nueve ítems con un contenido altamente improbable, atípico, psicótico, disfórico y orgánico que son raramente respondidos en la dirección de la escala tanto por muestra normal como por muestra clínica.</p> <p>Existen dos grupos de ítems dentro de esta escala: un grupo presenta una visión distorsionada o exagerada de sí mismo y de sus circunstancias (p.ej., “mi destino ha sido ser infeliz desde el día en que nací”), mientras que otro representa síntomas extremadamente raros y extraños (“a veces veo sólo en blanco y negro”, “creo que dentro de mí hay tres o cuatro personalidades completamente diferentes”). Estos dos grupos representan dos grandes estrategias de detección de exageración de síntomas comentadas: amplificación e inverosimilitud de los síntomas.</p>
SIM	Índice de simulación	<p>El índice de simulación (SIM) fue creado por Morey (1996) como un índice más específico de simulación que fuera relativamente independiente del nivel de psicopatología que mostrara el sujeto. El SIM está compuesto por ocho características del perfil que suelen observarse más frecuentemente en simuladores de trastornos mentales (especialmente de trastornos graves).</p>
FDR	Función discriminante de Rogers	<p>La función de discriminación de Rogers (RDF) fue creada para distinguir perfiles del PAI de pacientes clínicos genuinos, de sujetos a los que se les había instruido para simular distintos trastornos psiquiátricos. Para ello, calcula una puntuación basada en la ponderación de 20 escalas del PAI. Se ponderan negativamente ocho escalas del PAI, que suelen ser más altas en pacientes genuinos que en simuladores; por el contrario, se ponderan positivamente otras doce escalas que suelen ser más altas en muestras de simuladores que en pacientes clínicos genuinos. Esta escala comparte con el SIM su relativa independencia del nivel de psicopatología y su resistencia al entrenamiento para evitar ser detectado.</p> <p>No obstante, la investigación realizada no ha confirmado claramente su efectividad en contextos reales, por lo que Hawes y Boccaccini (2009) recomiendan que esta escala no sea usada por el momento hasta que otras nuevas evidencias garanticen su eficacia en estudios de grupos diagnósticos.</p>

3.3. Principales autoinformes de screening en la detección de exageración de síntomas

Si bien es cierto que las pruebas generales de autoinforme han mostrado una excelente capacidad de detección de simulación, no es menos cierto que algunos autores han llamado la atención sobre sus limitaciones, entre ellas el tiempo de aplicación que requieren y la formación especializada que exige su aplicación y posterior interpretación (Schinka y Borum, 1993, 1994).

Este hecho ha motivado la creación de pruebas específicas de simulación que pretenden ser complementos a las anteriores, particularmente, a través de pruebas de *screening* que permitan un acercamiento eficiente en términos de tiempo y esfuerzo.

Su diseño como *screening* (prueba de cribaje o despistaje) se refiere al uso de una estrategia para identificar síntomas y signos clínicos de manera temprana, rápida y eficaz, con el fin de que de que reduzcan sus efectos adversos sobre la población. Sharp y Lipsky (2002) señalaron las principales características de un instrumento de *screening*:

- se trata de pruebas cortas en cuanto al tiempo de aplicación (10-15 minutos o inferior);
- pruebas centradas en la evaluación específica de síntomas o conductas;
- pruebas centradas en la población “diana”, objeto de la evaluación;
- aplicación y corrección fácil y rápida;
- aplicación accesible a profesionales no especializados en el área evaluada para su posterior remisión;
- instrumentos que no implican costes económicos elevados;
- fundamentación psicométrica adecuada, incidiendo especialmente en fiabilidad y validez predictiva;
- su objetivo es la identificación rápida y eficaz de indicadores sintomatológicos significativos que, posteriormente, pueden ser objeto de una evaluación más detallada y exhaustiva.

Algunas de las pruebas de *screening* de simulación que han mostrado un cierto potencial en la investigación han sido el M-Test (Beaber, Marston, Michelli y Mills, 1985), el ADI (Mogge y LePage, 2004) y el SIMS (Smith y Burger, 1997; Widows y Smith, 2005).

El M-Test (Beaber et al., 1985) es un autoinforme de verdadero-falso, de 33 ítems, diseñado originalmente para la detección de simulación de esquizofrenia y compuesto por tres escalas (Confusión, S y M). La escala Confusión, está compuesta de 8 ítems que incluyen creencias y

actitudes que realmente no se dan en personas con trastornos mentales. La escala S se compone de 10 ítems, que incluyen síntomas reales asociados a la esquizofrenia. Finalmente, la escala M incluye 15 ítems que evalúan síntomas simulados de esquizofrenia. Las investigaciones iniciales de Beaber et al. (1985) mostraron una tasa de clasificaciones correctas del 83,7% de los pacientes clínicos con esquizofrenia y del 78,2% de los simuladores. Desde entonces, el M-test ha sido utilizado tanto en diseños de grupos conocidos como en simuladores análogos, mostrando resultados muy divergentes en cuanto a su sensibilidad y capacidad discriminativa entre pacientes clínicos y simuladores. Algunas investigaciones han indicado que los pacientes clínicos con psicopatologías más severas pueden obtener resultados similares en el test, que simuladores de síntomas de esquizofrenia (Hankins, Barnard y Robbins, 1993; Schretlen, Neal y Lesikar, 2000). También, se ha observado variabilidad relacionada con la etnia, el género y el nivel de educación (Gillis, Rogers y Bagby, 1991; Heinze, 2003; Schretlen et al., 2000), lo que hace recomendable investigación adicional que indague en los posibles factores que interfieren con estas variables.

El ADI (*Assessment of Depression Inventory*; Mogge y LePage, 2004) es un autoinforme de 39 ítems, diseñado para evaluar los síntomas depresivos mediante tres subescalas. La primera subescala evalúa síntomas genuinos de depresión, mientras que las otras dos subescalas están destinadas a evaluar síntomas simulados o respuestas inconsistentes o aleatorias. Aunque los resultados obtenidos en las investigaciones realizadas han sido prometedores, todavía es necesaria más investigación adicional que respalde su uso en contextos forenses y clínicos (Smith, 2008).

Finalmente, el SIMS, Inventario Estructurado de Simulación de Síntomas, prueba que ha destacado en la literatura científica como *screening* de simulación y que se abordará en el siguiente apartado más detallada y exhaustivamente por ser el objetivo principal de esta tesis.

3.4. SIMS, Inventario Estructurado de Simulación de Síntomas

El SIMS (Inventario Estructurado de Simulación de Síntomas; Smith y Burger, 1997; Widows y Smith, 2005) es un instrumento de evaluación que consta de 75 ítems, con un formato de respuesta dicotómico verdadero-falso, que puede ser aplicado en aproximadamente 10-15 minutos. Su objetivo es la detección de síntomas simulados de carácter psicopatológico y neuropsicológico y puede resultar útil aplicado como medida de *screening* aislada o formando parte de una batería de evaluación más completa y pormenorizada en un amplio rango de contextos (p. ej., evaluaciones forenses, neuropsicológicas y médico-legales) y condiciones (ingreso, instituciones psiquiátricas o penitenciarias, clínicas, etc.).

Dentro de las estrategias que emplea para la detección de síntomas simulados, incluye síntomas improbables (p. ej., “A veces cuando escribo un número de teléfono, me doy cuenta de que los números salen al revés aunque yo no lo intente.”), síntomas infrecuentes (“Siento tanto dolor en mi cuerpo que parece como si tuviera bichos andando bajo mi piel”), severidad de los síntomas (“Rara vez me río”) y combinación de síntomas inverosímiles (“Tengo dolor de cabeza y mareos justo cuando olvido algo”).

Una de las características relevantes del SIMS es que, aparte de la puntuación Total, ofrece puntuaciones en cinco escalas específicas, tal y como se observa en la tabla 3.5.

Su estructura factorial ha sido analizada en distintos estudios tanto en su versión original americana (Widow y Smith, 2005) como en su adaptación española (González-Ordi y Santamaría, 2009), así como se ha indagado en sus resultados de fiabilidad tanto para la puntuación total como para cada una de las escalas en distintas muestras clínicas y de simuladores tanto de psicopatología como de dolor crónico (González-Ordi, Capilla, Santamaría y Casado Morales, 2012; González-Ordi & Santamaría, 2009; Merckelbach & Smith, 2003; Widows & Smith, 2005).

Tabla 3.5
Descripción de las escalas y la puntuación total del SIMS

Escala	Descripción
Psicosis (Ps) (15 ítems)	Evalúa el grado en que el sujeto presenta síntomas psicóticos inusuales o extravagantes, que no son típicos de la patología psicótica real.
Deterioro neurológico (Dn) (15 ítems)	Evalúa el grado en que el sujeto presenta síntomas de tipo neurológico ilógicos o muy atípicos.
Trastornos amnésicos (Am) (15 ítems)	Evalúa el grado en que el sujeto presenta síntomas relacionados con trastornos de memoria que son inconsistentes con los patrones de deterioro producidos por disfunción o daño cerebral real.
Baja inteligencia (Bi) (15 ítems)	Evalúa el grado en que el sujeto exagera su déficit intelectual al fallar preguntas sencillas de conocimiento general.
Trastornos afectivos (Af) (15 ítems)	Evalúa el grado en que el sujeto informa de síntomas atípicos de depresión y ansiedad.
Total (75 ítems)	Ofrece una puntuación resumen del nivel general de exageración de síntomas.

La utilidad de sus escalas específicas reside en que en la práctica profesional, resulta difícil predecir qué dominios de funcionamiento puede fingir el evaluado como alterados o afectados. Por este motivo, siempre es recomendable cubrir las distintas áreas posibles en una primera evaluación. Así, las escalas específicas del SIMS permiten un amplio barrido de síntomas de las distintas áreas de simulación: aspectos físicos (escala Deterioro neurológico), cognitivos (escala Baja inteligencia y

Trastornos amnésicos) y psicopatológicos (escala Psicosis y Trastornos afectivos). Los resultados obtenidos permiten determinar el grado de sospecha de simulación y las áreas donde se presenta, para poder profundizar posteriormente en la evaluación de las mismas si se considera oportuno.

Con el objetivo de realizar una revisión exhaustiva de las investigaciones que han abordado las propiedades del SIMS en la detección de simulación, se ha realizado una búsqueda en las bases de datos PsycINFO, ProQuest, Google académico y Google, empleando como términos de búsqueda SIMS o “Structured Inventory of Malingered Symptomatology”.

De las cincuenta y dos investigaciones encontradas, se seleccionaron treinta y siete que específicamente aportaban información sobre la capacidad de detección de simulación del SIMS. Para ello, se consideraron aquellos estudios que habían empleado el SIMS en diseños de simulación análoga, de grupos conocidos, autosuficientes o en muestras clínicas sin incentivos de simulación, que permitían estudiar la especificidad de la prueba.

Se descartaron aquellos estudios que empleaban el SIMS en muestras forenses mixtas de clínicos y simuladores, en los que no se había establecido ningún criterio externo de diagnóstico de simulación y, por lo tanto, no resultaban útiles para determinar la especificidad o sensibilidad del SIMS (Beaulieu, 2008; Carter, Tharshini Chandra y Holden, 2009; Cima, Merckelbach, et al., 2003; Cima y Van Oorsouw, 2013; Göbber, Petermann, Piegza y Kobelt, 2012; Harrison, 2008; Leung, 2010; Merten, Friedel y Stevens, 2007; Merten, Thies, Schneider y Stevens, 2009; Naldony y O’Dell, 2010), aquellos en los que no se proporcionaban datos descriptivos sobre el SIMS que permitieran valorar su funcionamiento (Freeman, Powell y Kimbrell, 2008; Merckelbach et al., 2010), aquellos que eran descripciones de casos clínicos o con muestras inferiores a 10 casos (Merckelbach, Peters, Jellic, Brands y Smeets, 2006; Peters, van Oorsouw, Jellic y Merckelbach, 2013) o aquellos a los que no se pudo acceder a su contenido por no estar publicado ni haber respondido los autores a su consulta (Hammond, 2006).

A continuación, se describen brevemente los principales resultados de cada una de las treinta y siete investigaciones seleccionadas, agrupadas por el tipo de diseño empleado (análogos sin entrenamiento, análogos con entrenamiento, grupos diagnósticos o autosuficientes, muestras clínicas) y el contexto (forense civil o forense penal). Posteriormente, se procederá a realizar una síntesis de los principales resultados encontrados y las principales conclusiones que pueden deducirse de la investigación realizada hasta la fecha con esta herramienta. No obstante, dada la diversidad y heterogeneidad de los estudios realizados, se ha considerado conveniente incluir una descripción somera de cada uno de ellos.

3.5. Estudios de simulación análoga con el SIMS

3.5.1. Estudio de simulación análoga de Smith y Burger (1997)

Este estudio fue el primero realizado por los autores originales de la prueba, Smith y Burger, para validar originalmente el SIMS, por lo que es destacable su relevancia e importancia entre el conjunto de investigaciones que se han realizado sobre el SIMS. En este estudio, los autores contaron con una muestra de 476 estudiantes universitarios (edad media 24,43 años, rango 17 – 66 años). Los participantes fueron asignados aleatoriamente a un grupo de control (o respuesta honesta) o a uno de los siguientes grupos experimentales: simulación del mal estado general, simulación de psicosis, simulación de deterioro neurológico, simulación de amnesia, simulación de baja inteligencia y simulación de depresión. La muestra global fue posteriormente dividida aleatoriamente entre muestra de estudio (n=238) y muestra de validación cruzada (n=238).

En la muestra de estudio (n=238), se calcularon las puntuaciones medias en las escalas del SIMS de cada uno de los grupos considerados (tabla 3.6). Las diferencias encontradas fueron significativas, como puede advertirse en la tabla. Además, se realizó la prueba de rango múltiple de Duncan ($\alpha=0,05$), para especificar qué grupos se diferenciaban significativamente entre sí en cada una de las escalas de la prueba, de forma que pudiera estudiarse la concordancia entre la estructura hipotetizada del SIMS y los resultados en los grupos experimentales.

En todos los casos, el grupo control fue el que obtuvo puntuaciones significativamente inferiores al resto de grupos. En general, las puntuaciones más altas en cada escala fueron obtenidas por el grupo de simulación de mal estado general y el grupo experimental específico de dicha escala (p. ej., el grupo de simulación de psicosis en la escala Psicosis). Como excepción a este patrón, los grupos de simulación de psicosis y deterioro neurológico obtuvieron puntuaciones similares en la escala Deterioro neurológico del SIMS y el grupo de simulación de psicosis obtuvo puntuaciones similares en las escalas Psicosis y Trastornos amnésicos, sugiriendo que aquellos individuos que son instruidos para simular síntomas de psicosis o de trastornos mentales graves, tienden a asignarse síntomas de múltiples dominios. Por otro lado, el grupo de simulación de deterioro neurológico obtuvo una mayor puntuación en la escala Trastornos amnésicos que en la escala Deterioro neurológico, indicando que este grupo de simulación enfatizó la pérdida de memoria asociada al daño neurológico y cerebral.

Tabla 3.6

Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y control, en los estudios originales de Smith y Burger (1997)

	Grupo control		Grupo experimental de simulación												F _(7,230)
			Mal estado general		Psicosis		Deterioro neurológico		Amnesia		Baja inteligencia		Depresión		
	M	Dt	M	Dt	M	Dt	M	Dt	M	Dt	M	Dt	M	Dt	
Total	7,7	3,7	53,9	13,8	48,4	14,5	43,2	13,8	36,2	10,2	37,0	14,1	29,9	12,3	39,8***
Psicosis	0,8	1,0	12,7	2,1	11,8	3,7	7,0	5,2	3,0	4,0	4,9	4,2	3,1	3,8	39,3***
Deterioro neurológico	1,0	1,0	10,9	3,9	9,2	4,3	8,7	3,9	5,2	3,3	5,8	3,5	5,5	3,3	24,3***
Trastornos amnésicos	1,2	1,5	11,2	4,0	10,8	4,4	12,0	3,6	13,1	2,8	9,0	4,7	6,3	5,1	29,8***
Baja inteligencia	1,4	1,2	9,4	4,3	8,0	3,6	7,7	3,7	8,6	3,2	10,8	2,0	3,2	2,4	34,9***
Trastornos afectivos	3,3	2,0	9,8	2,9	8,6	2,9	7,9	3,0	6,2	3,1	6,5	3,9	11,8	1,5	26,6***

Nota. N= 238, n= 34 en cada condición. Aparece destacada en negrita la puntuación media de cada grupo en su escala específica (p. ej., del grupo de simulación de psicosis en la escala Psicosis del SIMS). M= Media; Dt= Desviación típica. *** p<0,001

Se calcularon, también, las medias y desviaciones típicas para el grupo general de simuladores (que incluía todas las condiciones experimentales) y el grupo control, tal y como aparecen en la tabla 3.7. Coherentemente con los análisis previos, el grupo global de simulación obtuvo puntuaciones superiores al grupo control en todas las escalas del SIMS.

Tabla 3.7

Puntuaciones medias en las escalas del SIMS del grupo control y del grupo global de simuladores en los estudios de Smith y Burger (1997)

Escalas del SIMS	Grupo control			Grupo experimental de simulación		
	n	M	Dt	n	M	Dt
Total	34	7,7	3,7	204	37,5	3,7
Psicosis	34	0,8	1,0	204	6,0	1,0
Deterioro neurológico	34	1,0	1,0	204	6,7	1,0
Trastornos amnésicos	34	1,2	1,5	204	9,7	1,5
Baja inteligencia	34	1,4	1,2	204	6,9	1,2
Trastornos afectivos	34	3,3	2,0	204	8,2	2,0

A partir de este estudio se establecieron los puntos de corte a considerar en la versión que posteriormente se publicó de la prueba editorialmente (Widows y Smith, 2005). Para su cálculo, se empleó la muestra de estudio (238 casos extraídos aleatoriamente de la muestra total de 476 casos) y se replicaron sus resultados en la muestra de validación cruzada, formada por los 238 casos

restantes no incluidos en la muestra anterior. El grupo de simulación se consideró globalmente, sin separar por trastorno simulado, tal y como es probable que se encontrara un profesional en su práctica cotidiana cuando utilizara los puntos de corte.

El punto de corte de cada escala se determinó de forma que discriminara óptimamente a aquellos que respondían honestamente de los que estaban simulando. Para ello, en cada escala se compararon las distribuciones de puntuaciones del grupo de respuesta honesta y el grupo de simuladores específicos de esa escala. En el caso de que ambas distribuciones de puntuaciones se solaparan, se elegía en su intersección aquel punto de corte que permitiera mayores tasas de acierto de clasificación en ambos grupos. Si las distribuciones no se solapaban, se elegía como puntuación de corte el punto medio entre ambas. De este modo, se maximizaron las tasas de acierto de clasificación en ambos grupos. En las tablas 3.8. y 3.9 se muestran respectivamente los resultados obtenidos en las muestras de estudio y de validez cruzada.

Ha de advertirse, no obstante, la limitada información de estos estadísticos de sensibilidad y especificidad ya que realmente no abordan la diferenciación clave (pacientes con trastorno frente a simuladores de dichos trastornos), sino meramente la discriminación entre personas “normales” (sin sintomatología clínica significativa) y simuladores análogos.

Tabla 3.8
Sensibilidad y especificidad del SIMS en la muestra de estudio de Smith y Burger (1997)

Escalas del SIMS	Puntos de corte	Sensibilidad (%)	Especificidad (%)	Eficiencia (%)
Total	>14	94,6	87,9	95,0
Psicosis	>1	81,8	76,1	75,1
Deterioro neurológico	>2	85,9	90,9	86,6
Trastornos amnésicos	>2	88,8	84,9	88,2
Baja inteligencia	>2	84,4	78,8	83,6
Trastornos afectivos	>5	76,1	90,9	78,2

Nota. n= 238.

Tabla 3.9
Sensibilidad y especificidad del SIMS en la muestra de validación cruzada de Smith y Burger (1997)

Escalas del SIMS	Puntos de corte	Sensibilidad (%)	Especificidad (%)	Eficiencia (%)
Total	>14	95,6	87,9	94,5
Psicosis	>1	79,5	72,8	78,6
Deterioro neurológico	>2	85,4	75,8	84,0
Trastornos amnésicos	>2	88,3	90,9	88,7
Baja inteligencia	>2	85,4	51,5	80,7
Trastornos afectivos	>5	74,6	72,7	74,4

Nota. n= 238.

3.5.2. Estudio de simulación análoga de Rogers y Cruise (1998)

Rogers y Cruise (1998) emplearon el SIMS junto con otras medidas para estudiar el efecto de determinados parámetros de las instrucciones en la simulación análoga. En concreto, estudiaron el impacto del contexto presentado (aplicable o inaplicable a la persona), el incentivo (positivo o negativo) y la relevancia informada en las instrucciones (implicaciones para el mundo real, desafío). Los autores advirtieron que una de las variables que más influía en la simulación análoga era la existencia de incentivos negativos. Cuando se avisaba a los simuladores análogos que habría consecuencias negativas si eran detectados como simuladores, se producía una presentación más controlada de la simulación y más sutil. El hecho de que el contexto resultara aplicable a la persona evaluada, comprensible y real, era otro factor relevante. En esta investigación, por lo tanto, no se indagaba específicamente sobre la capacidad del SIMS para detectar distintos tipos de simulaciones de trastornos, sino el efecto de las instrucciones de la simulación análoga en la presentación que los participantes hacían de sus síntomas.

3.5.3. Estudio de simulación análoga de Edens, Otto y Dwyer (1999)

Edens, Otto y Dwyer (1999) estudiaron la capacidad discriminativa del SIMS para detectar simuladores empleando para ello una muestra de 196 estudiantes universitarios (17% varones, edad media de 23,14, Dt 5,94). Estos fueron asignados aleatoriamente a una de las siguientes condiciones: simulación de psicosis (n=59), simulación de depresión (n=65) o simulación de deterioro cognitivo (n=72). Los participantes completaron tanto el SIMS como el SCL-90-R (*Cuestionario de 90 síntomas*; Derogatis, 1996) en dos ocasiones contrabalanceadas, una de ellas en la que se les pedía que respondieran sinceramente y otra, en que se les pedía que simularan la condición a la que habían sido asignados. A cada participante se le daba una breve descripción de la condición que debía

simular y se le indicaba que obtendría un incentivo económico si conseguía que su simulación no fuera detectada.

En términos de su capacidad discriminativa, tal y como se puede observar en la siguiente tabla, los análisis realizados mostraron que los participantes obtenían puntuaciones notablemente superiores en la condición de simulación en las distintas escalas del SIMS, especialmente, en la escala relacionada con la patología que se les había solicitado que simularan (p. ej., aquellos que se les indicaba que debían simular psicosis obtenían puntuaciones más altas en esta escala del SIMS).

Tabla 3.10
Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y de respuesta sincera en el estudio de Edens et al. (1999)

Escala del SIMS	Condición experimental											
	Depresión ^a				Psicosis ^b				Deterioro cognitivo ^c			
	Respuesta sincera		Simulación		Respuesta sincera		Simulación		Respuesta sincera		Simulación	
	M	Dt	M	Dt	M	Dt	M	Dt	M	Dt	M	Dt
Total	7,8	4,5	32,5	13,4	8,1	3,8	40,7	14,2	7,9	4,7	38,6	11,4
Psicosis	0,5	0,8	3,6	4,0	0,4	0,7	10,0	4,5	0,7	1,4	3,4	3,9
Deterioro neurológico	1,0	1,3	6,6	3,7	0,9	1,0	8,1	4,1	0,9	1,2	8,2	3,8
Trastornos amnésicos	0,9	1,4	8,1	4,7	1,1	1,5	8,8	4,6	0,8	1,6	12,3	3,6
Baja inteligencia	1,7	1,5	4,3	3,1	1,8	1,5	8,2	3,6	2,0	1,5	7,5	3,4
Trastornos afectivos	3,8	1,9	10,0	1,8	3,9	1,8	7,7	2,1	3,5	1,9	7,3	2,8

Nota. Aparece destacada en negrita la puntuación media de cada grupo en su escala específica (p. ej., del grupo de simulación de psicosis en la escala Psicosis del SIMS).

^a n= 65; ^b n= 59; ^c n= 72

Con el objetivo de conocer con más detalle la precisión de las escalas del SIMS, Edens y colaboradores (1999) analizaron su relación con otras medidas de psicopatología, en concreto, con el SCL-90-R. Tal y como aparece en la siguiente tabla, las correlaciones encontradas entre las escalas del SIMS y el índice global de gravedad del SCL-90-R (Índice GSI o Índice sintomático general) pueden considerarse entre bajas y moderadas.

Tabla 3.11
Correlaciones entre las escalas del SIMS y el nivel de sintomatología psicopatológica en el estudio independiente de Edens et al. (1999)

Escala del SIMS	SCL-90-R Índice global de gravedad
Total	0,52
Psicosis	0,23
Deterioro neurológico	0,43
Trastornos amnésicos	0,28
Baja inteligencia	0,11
Trastornos afectivos	0,49

Nota. N= 196.

Los autores decidieron realizar análisis específicos de sensibilidad y especificidad del SIMS con aquellos sujetos que, de acuerdo al SCL-90-R, mostraban un alto nivel de problemas emocionales para evaluar la posible influencia de este tipo de variables en la efectividad de la clasificación realizada con el SIMS. Para ello, se seleccionaron dos submuestras, la primera de ellas, compuesta por aquellos participantes que estaban actualmente en tratamiento psicológico o que lo habían estado en el pasado (n=33); la segunda, formada por aquellos que mostraron un nivel significativo de sintomatología psicopatológica (n=27), de acuerdo a su puntuación en el índice GSI del SCL-90-R (se seleccionaron aquellos con puntuaciones T en GSI superiores a 45). En estos análisis se emplearon los puntos de corte sugeridos por los autores originales (Smith y Burger, 1997). Como puede advertirse en la siguiente tabla, la puntuación total en el SIMS parece ser el mejor indicador de simulación, tal y como cabría esperar, con una clasificación correcta del 97% y el 100% de los simuladores en las submuestras de historial de tratamiento psicológico y presencia de sintomatología psicopatológica, respectivamente. Asimismo, respecto a su especificidad, el 9,1% y el 22,2% de los que respondieron con sinceridad en las muestras de tratamiento psicológico y presencia de psicopatología respectivamente, fueron clasificados erróneamente como simuladores.

Ha de advertirse que ya en muestras subclínicas (pacientes universitarios con cierto nivel de malestar psicopatológico) se eleva notablemente el número de falsos positivos empleando el punto de corte de 14, lo que hace sospechar que, en muestras con niveles mayores de psicopatología, (difícilmente compatibles con la realización de estudios universitarios en el sistema educativo americano) el número de falsos positivos sea mayor.

Tabla 3.12
Sensibilidad y especificidad del SIMS en participantes bajo tratamiento psicológico o con un alto nivel de sintomatología psicopatológica en el estudio de Edens et al. (1999)

Escalas del SIMS	Con historial de tratamiento psicológico ^a		Presencia de sintomatología psicopatológica ^{b,c}	
	Sensibilidad (%)	Especificidad (%)	Sensibilidad (%)	Especificidad (%)
Total > 14	97,0	90,9	100,0	77,8
Psicosis > 1	60,6	90,9	77,8	58,2
Deterioro neurológico > 2	90,9	97,0	92,6	73,1
Trastornos amnésicos > 2	90,9	84,8	100,0	81,5
Baja inteligencia > 2	81,8	75,8	85,2	74,1
Trastornos afectivos > 5	84,8	72,7	92,6	51,9

Nota. n^a = 33; n^b = 27; ^c Puntuación T en GSI > 45 en condición de respuesta honesta.

3.5.4. Estudio de simulación análoga de Theron et al. (2001)

En esta investigación realizada en Sudáfrica, Theron et al. (2001) dispusieron de 74 estudiantes universitarios a los que asignaron aleatoriamente a un grupo de simuladores análogos

(malestar general, $n=36$) o a un grupo control de respuesta honesta ($n=38$). A todos los sujetos se les aplicó un conjunto de pruebas, entre las que se encontraba el SIMS. Aplicando el punto de corte de 20, informaron de una clasificación correcta del 97% de los sujetos. Estos autores proponen, como punto de corte, puntuaciones superiores a 3 en Deterioro neurológico, superiores a 5 en Trastornos afectivos, superiores a 1 en Psicosis y superiores a 4 en Baja inteligencia y Trastornos amnésicos. No obstante, una de las limitaciones de este estudio es la ausencia de información sobre estadísticos descriptivos de las puntuaciones en cada uno de los grupos, las diferencias de medias o tamaños del efecto de las diferencias, así como la ausencia de grupo clínico de comparación de los resultados.

3.5.5. Estudio de simulación análoga de Merckelbach y Smith (2003)

Merckelbach y Smith (2003) estudiaron las propiedades psicométricas de la adaptación holandesa del SIMS, que fue aplicada a un grupo de estudiantes universitarios de Psicología y a un grupo de pacientes psiquiátricos hospitalizados (según los autores, “con diagnósticos variados, desde abuso de sustancias a trastorno de la personalidad límite, pasando por trastorno bipolar y esquizofrenia”).

La muestra 1 estuvo compuesta por 24 universitarias, que fueron instruidas para responder honestamente a todos los ítems del SIMS. A este grupo se le aplicó el SIMS en dos ocasiones diferentes, con un intervalo de tres semanas. La muestra 2 estuvo formada por 182 universitarios instruidos también para responder honestamente al SIMS. A los sujetos de esta muestra se les aplicó adicionalmente pruebas de evaluación de la depresión y la ansiedad. La muestra 3 estuvo compuesta por un grupo de 10 pacientes psiquiátricos hospitalizados, que respondieron sinceramente a la prueba y un grupo de 82 universitarios, que fueron asignados a una de las cuatro condiciones experimentales: grupo control de respuesta sincera ($n=25$), grupo de simulación de amnesia ($n=28$), grupo de simulación de esquizofrenia ($n=14$) y grupo de simulación de problemas neurológicos ($n=15$). Al grupo de universitarios de la muestra 3, se les proporcionó una breve descripción del trastorno a simular en un contexto relevante desde el punto de vista forense (p. ej., litigio), así como advertencias de que tenían que evitar ser detectados como simuladores.

Como puede verse en la siguiente tabla, aquellos que simularon un trastorno psiquiátrico o cognitivo obtuvieron consistentemente las puntuaciones totales más altas en el SIMS, notablemente por encima de los que respondieron sinceramente a la prueba y de los pacientes psiquiátricos. Estas diferencias fueron, además, significativas. Los análisis que realizaron los investigadores mostraron que los simuladores análogos tenían puntuaciones totales mayores que los pacientes [$t(65) = 6,2$, $p<0,01$], quienes a su vez obtenían puntuaciones superiores a los controles [$t(32) = 3,2$, $p<0,01$].

Además, en relación con la validez de cada una de las escalas específicas del SIMS, cada grupo de simuladores obtuvo puntuaciones más altas en la escala del SIMS relacionada con el trastorno simulado (p. ej., en la escala Psicosis aquellos participantes que simulaban esquizofrenia), con una interacción significativa entre el grupo de simulación y la escala.

Tabla 3.13
Puntuaciones medias en la escala total del SIMS en los estudios de la adaptación holandesa del SIMS de Merckelbach y Smith (2003)

Muestra	Total del SIMS	
	M	Dt
Muestra 1		
Respuesta sincera primera evaluación (n=24)	3,8	2,3
Respuesta sincera segunda evaluación (n=24)	3,4	3,4
Muestra 2		
Respuesta sincera (n=182)	5,8	3,9
Muestra 3		
Respuesta sincera (n=25)	5,1	4,5
Simulación amnesia (n=28)	34,8	9,9
Simulación esquizofrenia (n=14)	37,7	14,3
Simulación trastorno neurológico (n=15)	30,3	11,0
Pacientes clínicos (n=10)	11,4	5,8

En el estudio de Merckelbach y Smith (2003), también se aplicaron, a los sujetos de la condición de respuesta sincera, otras dos pruebas: el BDI (*Inventario de depresión de Beck*; Beck y Steer, 1987) y el STAI (*Cuestionario de Ansiedad estado-rasgo*; Spielberger, 1983). La puntuación total del SIMS obtuvo correlaciones significativas con la puntuación de depresión del BDI ($r=0,64$, $p<0,01$) y con la de ansiedad estado del STAI ($r=0,55$, $p<0,01$). No obstante, cuando se analizaron a los sujetos que obtenían puntuaciones más altas en alguna de las dos pruebas (percentiles iguales o superiores a 90), el 84% ($n=19$) en el caso del BDI y el 89% ($n=19$) en el caso del STAI seguían siendo clasificados como no simuladores (usando un punto de corte para total de >16). Esto es, parece que los errores de clasificación de sujetos con altos niveles de ansiedad o depresión como simuladores fueron relativamente bajos en el SIMS, de acuerdo a los datos de este estudio.

En la citada investigación no se proporcionó información de las puntuaciones en cada una de las escalas del SIMS en los distintos grupos, ni estadísticos de sensibilidad o especificidad para las mismas. A este respecto, ha de señalarse que los estudios de sensibilidad y especificidad fueron realizados con la comparación entre el grupo de simuladores y control, que responde honestamente, no entre el grupo de simuladores y pacientes clínicos. Como se comentó anteriormente, esta

comparación resulta poco relevante de cara a la detección de simulación ya que no se ajusta al tipo de discriminación que deberá realizar el profesional en su práctica.

3.5.6. Estudio de simulación análoga de Cima, Hollnack, et al. (2003)

La muestra constó de 266 participantes, de los cuales 204 eran estudiantes universitarios y 62 individuos evaluados en un contexto psiquiátrico forense. Los 204 estudiantes universitarios (98 varones y 106 mujeres) provenían de la Universidad de Bonn, de las carreras de Informática, Medicina, Psicología y Teología y tenían una edad media de 24,4 años ($Dt=6,3$). Los estudiantes fueron asignados aleatoriamente a uno de los siguientes seis grupos: respuesta honesta ($n=69$), baja inteligencia ($n=27$), trastorno afectivo ($n=31$), deterioro neurológico ($n=27$), psicosis ($n=29$) y trastorno amnésico ($n=21$). De acuerdo a la condición asignada, respondían al SIMS siguiendo instrucciones sobre el contexto en que debían simular sus síntomas.

Los 62 individuos evaluados en contexto psiquiátrico forense eran todos varones, tenían una edad media de 38,5 años ($Dt=10,9$) y estaban en prisión por distintos tipos de delitos (20 por homicidio, 22 por delitos sexuales, 10 por lesiones y 10 por robo, fraude o incendios). La mayoría de los pacientes cumplían criterios DSM-IV para uno o más trastornos de personalidad. En el eje I, los trastornos más comunes fueron esquizofrenia y otros trastornos psicóticos (31 pacientes, 50%), seguidos de diagnósticos relacionados con sustancias (27 pacientes, 44%), trastornos del estado de ánimo (17 pacientes, 27%) y trastornos de ansiedad (18 pacientes, 29%). Todos ellos fueron evaluados con el SIMS y el MMPI-2 en condición de respuesta honesta con las instrucciones estándar. Dos profesionales evaluaron independientemente de los resultados del SIMS y del MMPI los historiales de estos casos para determinar su grado de sospecha de simulación.

Tabla 3.14
Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos del estudio de Cima, Hollnack, et al. (2003)

	Universitario respuesta honesta ^a				Universitarios simuladores análogos ^b				Pacientes psiquiátricos forenses ^c			
	M	Dt	Rango	% >16	M	Dt	Rango	% >16	M	Dt	Rango	% >16
Puntuación Total SIMS	8,09	3,84	2-16	0%	31,01	13,62	4-72	86%	14,13	10,97	2-65	27%

Nota. $n^a=69$; $n^b=135$; $n^c=62$.

Del 27% de evaluados en contexto psiquiátrico forense con puntuaciones en el SIMS superiores a 16, un 65% había sido clasificado independientemente por expertos, como con sospecha de simulación. Adicionalmente, los autores estudiaron el perfil de los casos que habían obtenido

puntuaciones superiores a 16 en el SIMS y encontraron una mayor prevalencia en evaluados por delito sexual y con diagnóstico antisocial de la personalidad.

En la siguiente tabla pueden observarse las puntuaciones en las escalas del SIMS de los distintos grupos de simuladores análogos. Los estudios realizados mostraron que había un efecto significativo del grupo específico de simulación en las escalas que obtenían puntuaciones más elevadas, de forma similar a como se observó en el estudio de Smith y Burger (1997) o en Merckelbach y Smith (2003).

Tabla 3.15
Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales de Cima, Hollnack, et al. (2003)

Escalas del SIMS	Psicosis ^a		Deterioro neurológico ^b		Amnesia ^c		Baja inteligencia ^d		Trastorno afectivo ^e	
	M	Dt	M	Dt	M	Dt	M	Dt	M	Dt
Total	39,79	9,06	22,41	16,69	26	10,15	35,22	14,46	30,03	8,88
Psicosis	9,9	3,86	3,04	4,27	1,86	1,8	5,33	4,2	2,48	2,72
Deterioro neurológico	7,41	3,32	6,19	4,4	6,33	3,35	6,89	3,86	8,23	2,87
Trastornos amnésicos	10,21	3,89	4,22	4,2	8,48	4,06	7,74	4,24	6,65	3,91
Baja inteligencia	3,79	2,47	2,78	3,72	2,95	2,96	8,52	3,58	1,87	1,65
Trastornos afectivos	8,31	2,42	6,19	3,8	6,95	2,87	6,78	2,72	10,87	1,64

Nota. n^a= 29; n^b= 27 n^c= 21; n^d=27; n^e= 31.

Los autores no proporcionan en el artículo los descriptivos de las escalas en el grupo de respuesta honesta, ni en el grupo evaluado en contexto psiquiátrico forense. Además, el grupo evaluado en contexto psiquiátrico forense no es descompuesto en respuesta honesta y con sospecha de simulación, de acuerdo a los expertos, para los análisis de sensibilidad, especificidad o para comparar los descriptivos. Por motivos desconocidos, se considera al grupo en su conjunto a todos los efectos, a pesar de que no se pueda considerar como un grupo clínico de referencia (ya que hay un porcentaje de los casos en los que hay sospecha de simulación), ni como un grupo de simuladores. Esto hace que los estudios de sensibilidad o especificidad que realizan con este grupo de referencia tengan una relevancia o utilidad muy limitada.

3.5.7. Estudio de simulación análoga de Rogers et al. (2005)

Rogers et al. (2005) emplearon el SIMS y el PAI en un estudio de simulación análogo para estudiar la capacidad discriminadora de ambas pruebas entre trastorno facticio y simulación. Para ello, reclutaron a 65 estudiantes de doctorado de psicología, que fueron asignados a una de cuatro condiciones: facticio-dependiente, facticio-demandante, simulador u honesto (control).

Los resultados mostraron escasas diferencias entre las tres primeras condiciones, a excepción de la escala Trastornos afectivos del SIMS, la única que mostró diferencias significativas entre las condiciones. Ha de señalarse también, como importante limitación del estudio, la ausencia de grupos clínicos de comparación de los resultados. En la siguiente tabla, se indican las puntuaciones medias y desviaciones típicas halladas en este estudio en cada uno de los grupos y escalas del SIMS.

Tabla 3.16
Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y de respuesta sincera del estudio de Rogers et al., 2005

Escalas del SIMS	Dependiente ^a		Demandante ^b		Simulación ^c		Respuesta honesta (grupo control) ^d	
	M	Dt	M	Dt	M	Dt	M	Dt
Total	15,07	10,03	16,76	8,13	18,00	8,81	7,56	4,97
Psicosis	0,93	1,64	0,65	1,22	1,00	1,41	0,25	0,58
Deterioro neurológico	2,93	3,22	3,24	3,50	5,63	3,86	0,63	1,26
Trastornos amnésicos	2,21	3,45	2,18	3,05	2,56	2,92	0,69	0,87
Baja inteligencia	0,88	0,86	1,29	0,85	1,69	1,66	0,94	0,85
Trastornos afectivos	8,14	3,23	9,41	2,48	7,13	2,87	5,06	3,04

Nota. n^a= 14; n^b= 18; n^c= 17; n^d= 16.

3.5.8. Estudio de simulación análoga de retraso mental (Graue et al., 2007)

El objetivo de esta investigación fue estudiar la capacidad de detección de distintas pruebas de simulación de retraso mental, entre ellas el SIMS, en un diseño de simulación análoga. Para ello, dispusieron de una muestra de 25 casos con retraso mental moderado y una muestra de 25 adultos equiparada en edad, sexo y nivel educativo (menos de 11 años de educación), que fueron instruidos para simular retraso mental moderado, además de una muestra de adultos comunitarios, también equiparados en edad, sexo y nivel educativo, que respondían honestamente. Ambas muestras de adultos tenían un bajo nivel de inteligencia (CI estimado promedio de 82,1 y 80,2 respectivamente).

En la tabla 3.17 se incluyen los descriptivos de la puntuación total y de la escala Baja inteligencia en cada uno de los grupos, así como la *d* de Cohen y su sensibilidad y especificidad, empleando los puntos de corte anteriormente comentados. Los autores no proporcionan en la investigación los descriptivos del resto de escalas del SIMS, ni la sensibilidad o especificidad, considerando puntos de corte alternativos.

Tal y como se puede advertir en la tabla 3.17, los resultados indican la extrema limitación del SIMS, en su puntuación total y en la escala Baja inteligencia, para detectar simulación de retraso

intelectual moderado. Sus niveles tan bajos de especificidad resultan alarmantes y, claramente, ponen en evidencia que no deben usarse con poblaciones con posible déficit cognitivo o retraso intelectual, debido al alto riesgo de falsos positivos (el 50% de las personas con retraso mental, que responden honestamente hubieran sido clasificadas como con sospecha de simulación, empleando el punto de corte de 16 en la puntuación total del SIMS).

Tabla 3.17
Puntuaciones medias y desviaciones típicas en la puntuación total y en la escala Baja inteligencia del SIMS de los grupos de simulación análogo y de personas con retraso mental moderado en el estudio de Graue et al. (2007)

Escalas del SIMS	Muestra de adultos normales de bajo nivel educativo respondiendo honestamente ^a		Muestra de adultos normales de bajo nivel educativo respondiendo simulando retraso mental ^b		Muestra de personas con retraso mental moderado ^c		Comparación simuladores análogos ^b vs. personas con retraso mental ^c				
	M	Dt	M	Dt	M	Dt	t	d	Sen	Esp	% aciertos
Total	18,3	11,0	36,0	28,7	28,7	12,5	1,82	0,51	0,68	0,50	0,59
Baja inteligencia	4,2	1,9	16,3	12,5	5,6	1,6	1,96	0,56	1,00	0,00	0,49

Nota. n^a= 10; n^b= 25; n^c= 26.

3.5.9. Estudio de simulación análoga de trastorno por estrés postraumático de Efendov (2007)

Efendov (2007) realizó un estudio de simulación análoga con una muestra de 87 adultos de la población general, que habían sufrido un evento traumático y que habían presentado síntomas de estrés postraumático, si bien, en el momento de realización del estudio ya no presentaban dichos síntomas. Estos 87 adultos fueron asignados aleatoriamente a una de las siguientes condiciones: simulación de TEPT (trastorno por estrés postraumático) sin entrenamiento, simulación de TEPT con entrenamiento sobre los síntomas a alegar y simulación de TEPT con entrenamiento sobre las escalas existentes de detección de simulación. Adicionalmente, se seleccionó una muestra de trabajadores diagnosticados con TEPT y en proceso de reclamación de indemnizaciones, que fue considerada como muestra clínica de referencia (n=28).

En la tabla siguiente se pueden ver las medias y desviaciones típicas en cada grupo de simulación análogo y en la muestra de reclamantes diagnosticados con TEPT. Como se puede apreciar, hay un efecto muy marcado del entrenamiento específico en las escalas de síntomas frente a los otros dos grupos. De hecho, si se analizan los resultados de la tabla 3.19 se puede advertir cómo

la capacidad de discriminación de la prueba, evaluada con la *d* de Cohen, decae notablemente a bajos niveles en este grupo, mientras que obtiene valores grandes en los otros dos grupos.

En conjunto, los resultados obtenidos mostraron una satisfactoria capacidad del SIMS en la discriminación de simulación de estrés postraumático.

Tabla 3.18
Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y diagnosticados con TEPT en el estudio de Efendov (2007)

Escalas del SIMS	Simuladores análogos						Reclamantes de TEPT ^d	
	Sin entrenamiento ^a		Entrenado sobre los síntomas ^b		Entrenados sobre las escalas de validez ^c		M	Dt
	M	Dt	M	Dt	M	Dt	M	Dt
Total	37,08	17,27	37,96	17,76	23,94	10,71	19,46	10,13
Psicosis	4,81	4,82	4,62	5,48	1,19	1,05	1,43	2,24
Deterioro neurológico	8,96	4,89	8,23	4,96	4,71	4,24	4,64	3,17
Trastornos amnésicos	9,04	5,19	9,65	4,38	5,71	4,43	3,25	2,91
Baja inteligencia	4,27	3,57	4,85	3,79	3,06	2,11	2,88	2,01
Trastornos afectivos	10,00	2,19	10,62	2,23	9,26	2,16	7,50	2,08

Nota. n^a = 26; n^b = 26; n^c = 31; n^d = 28.

Tabla 3.19
Tamaños del efecto (d de Cohen) e intervalo de confianza (95%) de cada uno de los grupos de simulación análoga frente al grupo de reclamantes de TEPT en el estudio de Efendov (2007)

Escalas del SIMS	Simuladores análogos			Promedio
	Sin entrenamiento ^a	Entrenado sobre los síntomas ^b	Entrenados sobre las escalas de validez ^c	
	d	d	d	d
Total	1,24 (0,6-1,83)	1,28 (0,69-1,87)	0,43 (-0,09-0,95)	0,98
Psicosis	0,90 (0,34-1,46)	0,76 (0,21-1,31)	0,14 (-0,37-0,65)	0,60
Deterioro neurológico	1,09 (0,52 - 1,66)	0,91 (0,35 - 1,47)	0,07 (-0,44 - 0,58)	0,69
Trastornos amnésicos	1,38 (0,78-1,97)	1,72 (1,10-2,35)	0,66 (0,13-1,18)	1,25
Baja inteligencia	0,48 (-0,06-1,02)	0,65 (0,10-1,20)	0,09 (-0,42 - 0,60)	0,41
Trastornos afectivos	1,17 (0,59-1,75)	1,45 (0,85-2,05)	0,83 (0,30-1,36)	1,15

Nota. n^a = 26; n^b = 26; n^c = 31; n^d = 28.

3.5.10. Estudio de simulación análoga de amnesia relacionada con la realización de un crimen (Giger, Merten, Merckelbach y Oswald, 2010)

El estudio de Giger et al. (2010) contó con la participación de 60 adultos, que fueron asignados aleatoriamente a una de las siguientes tres condiciones: respuesta honesta, simulación de amnesia relacionada con la comisión de un crimen y simulación de amnesia relacionada con un crimen con advertencia de no exagerar sus síntomas. Los resultados obtenidos mostraron una alta sensibilidad y especificidad del SIMS en la detección de los simuladores frente a la respuesta honesta. Solo tres simuladores de la última condición (advertencia de no exagerar sus síntomas) no fueron detectados. En la siguiente tabla, se incluyen las puntuaciones medias en el total y en cada escala en cada uno de los grupos considerados. No obstante, una limitación del estudio supone la ausencia de grupo clínico de comparación de los resultados.

Tabla 3.20
Puntuaciones medias en el SIMS del grupo de respuesta sincera, simuladores de amnesia no advertidos y simuladores de amnesia advertidos en el estudio de Giger et al. (2010)

Escala del SIMS	Respuesta sincera ^a		Simuladores de amnesia sin advertencia ^b		Simuladores de amnesia con advertencia ^b	
	M	Dt	M	Dt	M	Dt
Total	5,8	3,7	37,9	13,8	22,3	12,6
Psicosis	0,8	1,1	5,4	4,4	2,9	3,6
Deterioro neurológico	0,9	0,8	7,8	4,0	3,5	2,5
Trastornos amnésicos	0,8	0,8	12,2	3,5	8,3	4,2
Baja inteligencia	1,2	1,0	5,3	4,2	2,9	3,1
Trastornos afectivos	2,2	2,2	7,3	3,3	4,8	3,1

Nota. n^a = 20; n^b = 20; ; n^c = 20

3.5.11. Estudio de simulación análoga de TDAH (Buijs, 2011)

Buijs (2011) realizó una investigación con 54 pacientes adultos diagnosticados de TDAH y 98 estudiantes universitarios sin diagnóstico de TDAH, ni ningún trastorno psicológico conocido. Los participantes fueron asignados aleatoriamente a dos condiciones experimentales: respuesta sincera o simulación de TDAH. La tabla 3.21 muestra las características demográficas de cada grupo.

Tabla 3.21
Características demográficas de los grupos de estudio de Buijs (2011)

Grupo	N	% varones	Edad (Dt)
Universitarios – respuesta sincera	51	17,6%	21,8 (6,9)
Pacientes TDAH – respuesta sincera	27	51,9%	31,3 (9,7)
Universitarios – simulación TDAH	47	12,8%	22,7 (7,4)
Pacientes TDAH – simulación TDAH	27	55,6%	31,9 (9)

Al total de la muestra se le aplicó una batería completa de evaluación, que incluía pruebas de atención sostenida (CPT de Conners; Conners, 2000), pruebas psicopatológicas, como el SCL-90-R o el BDI y pruebas de simulación de síntomas, como el Word Memory Test (Green, 2005), el Amsterdam Short Test Memory (Schmand y Lindeboom, 2005) o el SIMS.

En la siguiente tabla, pueden observarse las puntuaciones medias en cada una de las escalas del SIMS para cada condición. Los datos evidencian que las escalas que más diferencias mostraron entre el grupo clínico de TDAH y el grupo de simuladores análogos universitarios de TDAH, fueron las de Trastornos amnésicos, Psicosis y Baja inteligencia. Por el contrario, la escala con menos capacidad discriminatoria fue la de Trastornos afectivos. En este caso, el autor no proporcionó los descriptivos de la puntuación total del SIMS en los distintos grupos considerados.

Tabla 3.22
Puntuaciones medias en el SIMS del grupo de respuesta sincera de universitarios y pacientes con TDAH y del grupo de simulación de universitarios y de pacientes con TDAH en el estudio de Buijs (2011)

Escalas del SIMS	Universitarios respuesta sincera		Pacientes TDAH respuesta sincera		Universitarios simulación TDAH		Pacientes TDAH simulación TDAH	
	M	Dt	M	Dt	M	Dt	M	Dt
Psicosis	0,5	0,7	0,5	0,7	2,2	2,9	0,7	1,1
Deterioro neurológico	0,9	0,7	2,2	1,7	3,1	2,5	2,2	1,7
Trastornos amnésicos	1,5	0,9	3,3	2,2	6,2	3,5	4,5	2,5
Baja inteligencia	6	1	5,9	1,1	7,2	2,3	5,5	1,4
Trastornos afectivos	3,4	1,9	5,3	1,9	5,5	3,1	5,7	1,9

Nota. El autor de la investigación no proporcionó los estadísticos descriptivos de la puntuación total del SIMS

3.5.12. Estudio de simulación análoga de Dandachi-FitzGerald y Merckelbach (2012)

Esta investigación se compuso de tres estudios, empleándose el SIMS junto con el MVST (*Medical Symptom Validity Test*; Green, 2004) en los estudios 2 y 3. El estudio 2 consistió en un diseño de análogos con 62 estudiantes universitarios (14 varones, edad media de 22 años, Dt de 3,7, rango de 18 a 40 años), con tres grupos asignados aleatoriamente: control de respuesta honesta (n=20), simulador en contexto civil (n=21) y simulador en contexto criminal (n=21). Las instrucciones para el grupo de simuladores en contexto civil, consistían en pedir a los participantes que se imaginaran que acababan de ver la muerte de su sobrino en un accidente de tráfico y que querían solicitar, por vía judicial, una compensación económica por daños psicológicos al conductor, para lo cual tenían que mostrar los síntomas que consideraran convenientes a este objetivo en la evaluación que se les iba a realizar. Las instrucciones para el grupo de simuladores en el contexto criminal, consistían en indicar a los participantes que habían matado a un pedófilo, que había acosado a su joven sobrino y que tenían que alegar los síntomas que consideraran oportunos en la evaluación, para evitar la condena y eludir su responsabilidad criminal.

Como se puede observar en la siguiente tabla, ambos grupos de simulación obtenían puntuaciones sensiblemente superiores al grupo de respuesta honesta. Es interesante señalar que el grupo de simulación en contexto criminal obtuvo puntuaciones ligeramente superiores al grupo de simulación en contexto civil, aunque dichas diferencias no fueron estadísticamente significativas. El SIMS obtuvo mejores tasas de sensibilidad que el MSVT en este estudio, si bien, como limitación fundamental, cabe indicar la ausencia de un grupo clínico de referencia para establecer la especificidad de los resultados. Adicionalmente, no se informa de los resultados descriptivos o de capacidad discriminativa de las escalas del SIMS, ni de posibles puntos de corte alternativos y su sensibilidad.

Tabla 3.23
Puntuaciones medias y desviaciones típicas en la puntuación total de los grupos de respuesta honesta, simulación en contexto civil y simulación en contexto criminal en el estudio de Dandachi-FitzGerald y Merckelbach (2013)

SIMS	Grupo control ^a		Grupo de simuladores análogos en contexto civil ^b		Grupo de simuladores análogos en contexto criminal ^b		F	p	Eta cuadrado	Scheffé post-hoc
	M	Dt	M	Dt	M	Dt				
Total	4,5	3,5	27,9	7,7	30,1	10,3	69,77	<0,05	0,84	1 < 2,3

Nota. n^a= 20; n^b= 21; n^c= 21.

En esta misma investigación, se dispuso un tercer estudio en el que se contó con una muestra de 21 personas sometidas a evaluación en un centro de salud mental (13 varones, media de edad de 33 años, Dt de 11,8, rango de 19 a 55 años), de los cuales 14 pacientes alegaban trastorno por déficit de atención con hiperactividad y 6 un posible trastorno del espectro autista. En la tabla 3.24 se detallan los descriptivos de la puntuación total del SIMS obtenidos en esta muestra.

Tabla 3.24
Puntuación media y desviación típica en el grupo mixto que accedían a evaluación en el estudio de Dandachi-FitzGerald y Merckelbach (2013)

Escalas del SIMS	Muestra mixta clínica y exageración de síntomas	
	M	Dt
Total	13,4	11,8

Nota. n= 21.

3.5.13. Estudio de simulación análoga con adolescentes con diagnóstico dual en tratamiento psicológico (Rogers, Hinds y Sewell, 1996)

Rogers, Hinds y Sewell investigaron, por primera vez, el SIMS con adolescentes. Para ello contaron con una muestra de 53 adolescentes con diagnóstico dual de edades comprendidas entre los 14 y los 17 años, que habían sido remitidos por el juzgado para recibir tratamiento psicológico en un centro (79,2% varones, edad media de 15,5 años y Dt de 1,01). Los adolescentes cumplimentaron el SIMS junto con otras medidas de simulación en dos condiciones: respuesta sincera y simulación. Posteriormente, los participantes fueron asignados aleatoriamente a una de las siguientes condiciones de simulación: a) esquizofrenia, b) depresión mayor y c) trastorno de ansiedad generalizada. A cada uno de los grupos se les proporcionó una descripción del trastorno a simular junto con sus principales síntomas y se les informó que tendrían una compensación económica, tanto por su participación en el estudio como si conseguían evitar ser detectados como simuladores.

Los autores analizaron el valor predictivo positivo y negativo del SIMS, obteniendo valores moderados, tal y como se pueden observar en la tabla 3.25, en un contexto donde la tasa de simulación era del 50%, proponiendo puntos de corte alternativos para las diversas escalas y la puntuación total.

Tabla 3.25
Sensibilidad de las escalas del SIMS en adolescentes en el estudio de Rogers, Hinds y Sewell (1996)

Escalas del SIMS	Criterios propuestos			Criterios óptimos		
	Punto de corte	Valor predictivo positivo	Valor predictivo negativo	Punto de corte	Valor predictivo positivo	Valor predictivo negativo
Total	> 16	0,87	0,62	> 40	0,49	0,94
Psicosis	> 2	0,77	0,70	> 9	0,43	0,91
Deterioro neurológico	> 3	0,83	0,62	> 8	0,57	0,94
Trastornos amnésicos	> 4	0,74	0,68	> 10	0,41	0,94
Baja inteligencia	> 4	0,58	0,81	> 7	0,28	0,93
Trastornos afectivos	> 5	0,91	0,70	> 7	0,68	0,91

Nota. N=57.

Los autores no proporcionaron las puntuaciones medias en el SIMS de los distintos grupos o los resultados de sensibilidad y especificidad de cada punto de corte considerado.

3.5.14. Estudio de simulación análoga con adolescentes con grupo clínico de referencia (Dearth, 2007)

El estudio de Dearth (2007) dispuso de tres muestras de adolescentes: dos muestras procedentes de la población general y la otra muestra de población clínica. Al primer grupo, compuesto por 18 adolescentes, se le solicitó que respondiera de forma honesta y sincera a la prueba (33% varones, edad media de 16 años, Dt de 1,46); al segundo grupo, compuesto por 37 adolescentes, se le solicitó que respondiera simulando un determinado trastorno (40,5% varones, edad media de 15,8 años, Dt de 1,13). El tercer grupo estuvo formado por 31 adolescentes clínicos previamente diagnosticados con TDAH, trastorno bipolar o trastorno depresivo (48,6% varones, edad media de 15,71 y Dt de 1,37).

Todos los participantes fueron evaluados con pruebas de personalidad con indicadores de validez incluidos en ella (*The Personality Inventory for Youth PIY*; Lachar y Gruber, 1995), pruebas específicas de simulación de psicopatología (SIRS y SIMS) y pruebas de simulación cognitiva (TOMM y *The Letter Memory Test LMT*, Inman et al., 1998). Al grupo clínico se le solicitó que respondiera honestamente a la prueba, de acuerdo a los síntomas y problemas que mostraban. A los dos grupos de adolescentes no clínicos se les presentó una situación por escrito que tenían que imaginarse como contexto de respuesta a la prueba. En ambos casos se les indicaba que habían tenido una pelea, que ellos no habían empezado y que se les había denunciado por ello e iban a ser evaluados previamente a un juicio por faltas. En el caso de respuesta honesta, se les indicó que respondieran honestamente

a la prueba en este contexto; al grupo de simulación análogo se les instruyó para que simularan una serie de síntomas afectivos, psicopatológicos y cognitivos (problemas de concentración o atención) con el objetivo de ser derivados a tratamiento en vez de a algún tipo de reformatorio. Se les indicaba también a este último grupo que debían ser precavidos en sus respuestas sin que resultara obvio que estaban simulando sus síntomas. En las dos tablas siguientes se puede observar la capacidad discriminativa del SIMS.

Si bien la autora no incluye correlaciones del SIMS con el resto de pruebas, si realiza una regresión logística para determinar qué subconjunto de pruebas o escalas permitía una mejor clasificación de los participantes en el grupo de simulador análogo o clínico. En este análisis por pasos la primera escala que se incluyó fue Trastornos afectivos del SIMS, consiguiendo una clasificación por si sola del 70,6%, aumentando al 73,5% cuando se incluyó a la puntuación RO del SIRS. Sus resultados mostraban que la escala más potente en su capacidad discriminativa de ambos grupos era Trastornos afectivos del SIMS.

Tabla 3.26
Puntuaciones medias en el SIMS de los grupos del estudio de Dearth (2007)

Escala del SIMS	Respuesta honesta ^a		Muestra clínica ^b		Simulador análogo ^c		T	d
	M	Dt	M	Dt	M	Dt		
Total	8,72	5,52	12,23	8,21	25,43	17,07	3,94***	0,99
Psicosis	0,50	1,2	0,94	1,57	4,22	4,67	3,73***	0,94
Deterioro neurológico	1,22	1,22	2,39	2,47	4,84	3,96	2,99**	0,74
Trastornos amnésicos	1,67	2,00	2,23	2,16	5,30	4,76	3,32***	0,83
Baja inteligencia	1,67	1,14	2,39	2,04	3,73	3,19	2,02	0,50
Trastornos afectivos	3,67	2,43	4,19	1,96	7,35	3,39	4,59***	1,14

Nota. n^a = 20; n^b = 20; ; n^c = 20.

p<0,01 * p < 0,001.

Tabla 3.27
Sensibilidad, especificidad y poder predictivo positivo y negativo de la puntuación Total del SIMS (Dearth, 2007)

Punto de corte	Sensibilidad	Especificidad	Prevalencia 0,54		Prevalencia 0,25		Prevalencia 0,10	
			PPP	PPN	PPP	PPN	PPP	PPN
Total >16	0,65	0,77	0,77	0,65	0,48	0,87	0,26	0,96
Total > 40	0,22	1,00	1,0	0,52	1,0	0,80	1,0	0,92
Af > 5	0,62	0,81	0,79	0,64	0,52	0,87	0,25	0,94
Af > 7	0,46	0,94	0,80	0,59	0,73	0,84	0,43	0,93

3.6. Estudios de simulación análoga y resistencia al entrenamiento con el SIMS

3.6.1. Estudio de detección de simulación de psicosis en simuladores análogos con diferente nivel de entrenamiento (Jelicic, Hessels y Merckelbach, 2006)

El estudio de Jelicic et al. (2006) tenía por finalidad analizar la capacidad del SIMS para detectar a simuladores que habían sido entrenados para simular y eludir las medidas de detección habituales, particularmente, en el contexto criminal y en la simulación de psicosis (alucinaciones, delirios...), como vía para reducir o modificar la condena por homicidio o crímenes. Como argumentan los autores, es frecuente en estos casos que los acusados sean entrenados para su posterior evaluación y resultar convincentes en los síntomas alegados. Dado que se desconoce la eficacia del SIMS bajo estas circunstancias, decidieron acometer este primer estudio análogo.

Para la investigación se dispuso de 60 estudiantes universitarios, que fueron asignados aleatoriamente a una estas cuatro condiciones: respuesta honesta (control), simulador análogo de psicosis sin información dada (simulador análogo ingenuo, solo se le informaba que había cometido un homicidio y que debía alegar psicosis para evitar la condena), simulador análogo de psicosis con cierta información sobre la psicosis (simulador análogo informado) y simulador análogo de psicosis, con información acerca de la psicosis y advertido de que no debía exagerar sus síntomas ya que había tests con índices específicos para detectarlo (simulador análogo entrenado). Sus resultados mostraron que, conforme mayor era el conocimiento de los simuladores análogos de la patología y de las técnicas de detección, había una ligera tendencia a reducir sus puntuaciones en el Total del SIMS y a desarrollar un estilo menos flagrante y global de simulación, frente a aquellos con menor conocimiento y entrenamiento. No obstante, tomado en su conjunto, los autores consideraron que el efecto del entrenamiento y la información era relativamente pequeño y el SIMS seguía manteniendo altas tasas de detección (considerando un punto de corte >16). No obstante, debe tenerse en cuenta que el “entrenamiento” proporcionado a los participantes consistió en una charla de escasos minutos, en la que se les proporcionó información general sobre los síntomas clave en un trastorno psicótico.

Otras limitaciones del estudio son la falta de información sobre las puntuaciones medias y desviaciones típicas obtenidas por cada grupo en cada escala y en el total, la ausencia de estudios

específicos de sensibilidad o especificidad en relación con otros puntos de corte alternativos, así como la falta de grupo clínico de comparación de los resultados.

3.6.2. Estudio de detección de simulación de psicosis en simuladores análogos con diferente nivel de conocimiento de psicopatología (Jelicic, Peters, Leckie y Merckelbach, 2007)

De forma complementaria al estudio anterior, Jelicic, Peters, et al. (2007) quisieron profundizar en la resistencia del SIMS al entrenamiento de los simuladores, complementando el estudio anterior con otro que tuviera en cuenta el nivel de conocimiento del simulador de la psicopatología bajo análisis.

Para ello, 61 estudiantes universitarios fueron seleccionados y clasificados por su nivel de conocimiento en psicopatología, en dos grupos (n=30 y n=31, respectivamente). El primer grupo, estaba formado por estudiantes de Psicología, que habían cursado al menos una asignatura específica de trastornos psicológicos (grupo con conocimiento en psicopatología). El segundo grupo, se compuso de estudiantes de otras carreras, que nunca habían cursado ninguna asignatura relacionada con la psicopatología (grupo sin conocimientos de psicopatología). A todos los participantes se les aplicó el SIMS en dos ocasiones, en un primer momento, se les pidió que lo contestaran honestamente y en un segundo momento, fueron instruidos para simular un trastorno psicótico como eximente de condena, tras un crimen que supuestamente habían cometido. En la tabla 3.28, pueden observarse los resultados obtenidos. Si bien se advierte la presencia de puntuaciones más bajas en el grupo de participantes con mayor conocimiento de psicopatología, reflejando un efecto de esta variable en la simulación análoga del SIMS, en ambos casos siguen siendo puntuaciones muy superiores a los puntos de corte establecidos habitualmente (total > 16), lo que hace que la tasa de detección o sensibilidad fuera del 100% en el primer caso (simuladores con bajo conocimiento de psicopatología) y del 97% en el segundo (simuladores con alto conocimiento de psicopatología).

Tabla 3.28
Puntuaciones medias y desviaciones típicas en la puntuación Total en el SIMS en el estudio de (Jelicic, Peters, et al., 2007) de simulación análoga de psicosis

	Participantes con bajo conocimiento de psicopatología (n=30)	Participantes con alto conocimiento de psicopatología (n=30)
Condición control	5,3 (3,5)	3,5 (2,7)
Condición análoga de psicosis	45,9 (14,2)	37,0 (12,8)

3.6.3. Estudio de detección de simulación de psicosis en simuladores análogos expertos en psicopatología (Jelicic, van Gaal y Peters, 2013)

Profundizando más en esta línea de investigación, Jelicic et al. (2013) estudiaron el efecto del conocimiento de la psicopatología en la capacidad de simular síntomas. En concreto, estos autores seleccionaron una muestra de 24 estudiantes sin conocimiento o experiencia de trato con pacientes psiquiátricos (9 varones, edad media de 21,3 años y Dt de 2,2) y 23 profesionales de salud mental, que trabajaban como clínicos en un hospital psiquiátrico o como investigadores (2 varones, edad media de 28,7 años y Dt de 4,1). Todos los profesionales habían tenido experiencia directa con pacientes psicóticos o bien investigaban específicamente en el área de la psicosis. Ninguno de los participantes en el estudio tenía un historial previo de problemas psiquiátricos o psicopatológicos.

Al total de la muestra se le pidió que respondiera al SIMS primero de forma honesta y, en un segundo momento, simulando psicosis para evitar una condena por un delito grave (homicidio). Los resultados que obtuvieron mostraron que ambos grupos obtuvieron puntuaciones más altas en la condición de simulación, que en la de respuesta honesta, ($F(1,45)=235,41, p<0,001$). Los datos evidenciaron, además, un efecto del grado de experiencia psicopatológica ($F(1,45)=13,10, p<0,01$) asociado a los expertos, obteniendo puntuaciones más bajas en el SIMS y mostrándose menos proclives a ser detectados como simuladores por la prueba.

Por lo tanto, los resultados indican unas menores puntuaciones en el SIMS en aquellas personas con experiencia y conocimiento de las alteraciones psiquiátricas, si bien en ambos casos la sensibilidad del SIMS fue alta, empleando el punto de corte de 16. Entre las limitaciones del estudio, cabe señalar, al igual que en los anteriores, la ausencia de grupos clínicos de referencia, que permitan establecer la especificidad de la medida con dicho punto de corte. Igualmente, el estudio no ofrece los estadísticos del SIMS, ni ninguna información sobre sus escalas, ni siquiera sobre la escala Psicosis, así como tampoco considera otros posibles puntos de corte.

3.6.4. Estudio de detección de simulación de daño cerebral en simuladores análogos con diferente nivel de entrenamiento (Jelicic, Merckelbach, Candel y Geraerts, 2007)

Jelicic, Merckelbach, et al. (2007) llevaron a cabo un estudio similar al de Jelicic et al., (2006), pero, en esta ocasión, se centraron en la simulación de daño cerebral como consecuencia de un accidente. La muestra se compuso de 90 estudiantes universitarios, que fueron asignados aleatoriamente a una de las siguientes condiciones: control de respuesta honesta ($n=30$), simulador ingenuo -sin información extra aparte de que debía simular daño cerebral tras un accidente ($n=30$)- y

simulador entrenado -con información sobre los síntomas que caracterizaban la lesión, sobre la existencia de escalas para detectar su simulación y advertencias sobre que fuera creíble en sus respuestas (n=30)-. En la tabla 3.29, se recogen las puntuaciones totales en cada uno de los grupos, que muestran que el grupo de simuladores entrenados obtiene puntuaciones sensiblemente inferiores al otro grupo, si bien siguen siendo superiores al punto de corte establecido. En comparación con el otro instrumento incluido en el estudio, el AMST (Schmand y Lindeboom, 2005), el SIMS mostró una mayor resistencia al entrenamiento en simulación, viéndose menos afectada su capacidad de detección en estos casos.

Tabla 3.29
Puntuaciones medias y desviaciones típicas en la puntuación Total en el SIMS en el estudio de Jellic, Merckelbach, et al. (2007) de simulación análoga de daño cerebral

	Control (n=30)	Simuladores ingenuos (n=30)	Simuladores entrenados (n=30)
Puntuación total SIMS	5,0 (3,3)	30,3 (11,2)	25,5 (10,4)

3.6.5. Estudio de detección de simulación de daño cerebral en simuladores análogos con diferente nivel de entrenamiento de Jellic, Ceunen, Peters y Merckelbach (2011)

El estudio de Jellic et al. (2011) tenía por finalidad estudiar la resistencia de las pruebas de simulación TOMM y SIMS al efecto del entrenamiento de los simuladores en los síntomas de daño cerebral y en las estrategias de detección de la simulación. Para esta investigación se contó con la participación de 90 estudiantes universitarios a los que se les asignó aleatoriamente a una de tres condiciones experimentales, para aplicárseles posteriormente el TOMM y el SIMS. Las tres condiciones experimentales fueron: a) grupo control (n=30); b) grupo de simulador análogo entrenado con síntomas (n=30), instruido para simular daño cerebral tras un accidente, al que se le informaba de los síntomas principales y se le advertía que, al menos, uno de los tests buscaría detectar si mentía; y c) grupo de simulador análogo entrenado con síntomas y estrategias de detección de simulación (n=30), instruido para simular daño cerebral tras un accidente, con la misma información que el grupo anterior pero al que, adicionalmente, se le proporcionaban las siguientes instrucciones: “Ten en cuenta que los instrumentos diseñados para detectar simulación de problemas cognitivos aparentemente son como los otros tests y cuestionarios. No obstante, en estos tests los pacientes con daño cerebral y las personas sanas, que no están simulando obtienen

puntuaciones muy similares. Por lo tanto, trata de responder normalmente a estos tests que pueden estar diseñados para detectar que estás mintiendo”.

Los resultados, que se muestran en la tabla 3.30, muestran que, a pesar de las instrucciones y el entrenamiento, se siguen obteniendo puntuaciones notablemente superiores a los puntos de corte habitualmente considerados. El 93% de los participantes del grupo b eran correctamente identificados como simuladores por el SIMS y el 86% del grupo c (Total > 16), lo que sugiere una satisfactoria resistencia del SIMS al entrenamiento en estas técnicas de detección de simulación.

Tabla 3.30
Puntuaciones medias y desviaciones típicas en la puntuación Total en el SIMS en el estudio de Jelacic et al. (2011) de simulación análoga de daño cerebral tras accidente

	Control (n=30)	Simuladores entrenados con síntomas (n=30)	Simuladores entrenados con síntomas y estrategias específicas de detección de la simulación (n=30)
Puntuación total SIMS	4,6 (2,7)	25,8 (6,8)	27,2 (9,8)

3.6.6. Estudio de detección de simulación de lesiones por esguince cervical en simuladores análogos con diferente nivel de entrenamiento (Merten, Diederich y Stevens, 2008)

El estudio de Merten et al. (2008) investigó el efecto del entrenamiento en simuladores análogos de esguince cervical. El objetivo del estudio fue determinar el efecto del entrenamiento en los simuladores, en la capacidad de detección de distintas pruebas, entre ellas el SIMS. Para ello, se seleccionó una muestra de 32 adultos de la población general que fueron asignados aleatoriamente a una de dos condiciones: a) simulación de esguince cervical sin entrenamiento; y b) simulación de esguince cervical con advertencia sobre la existencia de escalas de detección de la simulación. Los resultados pueden observarse en la tabla 3.31. El SIMS con un punto de corte de 16 presentaba una sensibilidad de 0,94, en el grupo de simuladores ingenuos (sin advertencias previas sobre la existencia de escalas de detección) y de 0,78, en el de los simuladores entrenados y advertidos sobre la existencia de escalas de detección. Los autores concluyen señalando que, si bien las pruebas manifiestan una cierta resistencia al entrenamiento (en ambos casos tienen unos valores aceptables de sensibilidad), no es menos cierto que afectan a su capacidad discriminativa. Por lo tanto, en estos contextos, aconsejan la inclusión de una amplia batería de pruebas de detección de la simulación que, en su conjunto, permitan un cribado adecuado de los casos.

Tabla 3.31

Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales de simuladores análogos de esguince cervical en el estudio de Merten et al. (2008)

Escala del SIMS	Simuladores análogos ingenuos ^a	Simuladores análogos entrenados ^a
	M (Dt)	M (Dt)
Total	31,6 (11,3)	20,1 (8,7)
Psicosis	2,4 (3,2)	1,2 (1,8)
Deterioro neurológico	6,7 (2,8)	3,4 (2,1)
Trastornos amnésicos	8,2 (3,8)	5,1 (3,2)
Baja inteligencia	4,6 (3,2)	2,9 (2,3)
Trastornos afectivos	9,6 (2,0)	7,4 (2,3)

Nota. ^a n=16 ^b n=16.

3.6.7. Estudio de detección de simulación de estrés postraumático en simuladores análogos con diferente nivel de entrenamiento (Merten, Lorenz y Schlatow, 2010)

El estudio de Merten et al. (2010) investigó el efecto del entrenamiento en simuladores análogos de trastorno por estrés postraumático (TEPT). Estudios previos de este autor, en los que se había incluido el SIMS (Merten et al., 2009), habían mostrado que la simulación de este trastorno en el contexto forense civil es muy elevada (con valores entre un 25% y un 70%), por lo que el objetivo de esta investigación fue determinar el efecto del entrenamiento en los simuladores en la capacidad de detección de distintas pruebas, entre ellas el SIMS. Para ello, seleccionó una muestra de 80 adultos que fueron asignados aleatoriamente a una de las siguientes condiciones: a) simulación de TEPT sin entrenamiento previo; b) simulación de TEPT con información sobre sus principales características y síntomas; c) simulación de TEPT con advertencia sobre la existencia de escalas de detección de la simulación; d) simulación de TEPT con información sobre sus principales características y síntomas y advertencia sobre la existencia de escalas de detección. En la tabla siguiente aparecen los descriptivos obtenidos en el SIMS en cada uno de los grupos. Los autores concluyen que el empleo de un conjunto de pruebas eleva la capacidad discriminativa y permite obtener altos porcentajes de detección de simuladores (específicamente, del 97%). No obstante, la ausencia de grupos clínicos de referencia o de estudios con grupos diagnósticos limita sensiblemente el alcance de las conclusiones de este estudio.

Tabla 3.32
Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales de simuladores análogos de trastorno por estrés postraumático. Estudio independiente de Merten et al. (2010)

Escalas del SIMS	Simuladores análogos ingenuos ^a	Simuladores análogos informados de los síntomas ^a	Simuladores análogos advertidos de escalas de detección ^a	Simuladores análogos informados de los síntomas y advertidos de escalas de detección ^a
	M (Dt)	M (Dt)	M (Dt)	M (Dt)
Total	29,60 (9,48)	30,55 (10,56)	26,55 (10,34)	22,95 (7,40)
Psicosis	3,35 (2,11)	3,60 (3,50)	2,40 (2,23)	1,75 (1,83)
Deterioro neurológico	6,10 (2,95)	4,80 (3,02)	4,75 (2,15)	4,15 (2,13)
Trastornos amnésicos	8,00 (3,04)	7,95 (3,97)	6,65 (4,22)	5,50 (3,22)
Baja inteligencia	2,95 (1,99)	4,10 (3,31)	3,50 (3,05)	2,60 (2,33)
Trastornos afectivos	9,20 (2,14)	10,10 (2,13)	9,25 (2,12)	8,95 (1,96)

Nota. ^a n=20.

3.7. Estudios de grupos diagnósticos o diseños autosuficientes en contexto forense penal con el SIMS

3.7.1. Estudio de simulación de psicopatología en presos de Poythress et al., (2001) y Edens et al. (2007)

En los artículos de Poythress et al. (2001) y Edens et al. (2007) se presentan los resultados de una investigación en la que incluyeron el SIMS en un estudio más amplio sobre evaluación de la simulación y la psicopatía y en la que emplearon conjuntamente diseños de simulación análoga con grupos diagnósticos. Para ello, contaron con una muestra de 116 reclusos varones que fueron subdivididos en las siguientes submuestras: a) reclusos que respondieron en la condición de sinceridad a la prueba (n=30); b) reclusos que respondieron en la condición de simulación a la prueba alegando un trastorno mental relevante, sin mayores precisiones por parte de las instrucciones análogas (n=30); c) reclusos que se encontraban en la unidad psiquiátrica de la prisión y que respondieron sinceramente (n=30); y d) reclusos que se encontraban en la unidad psiquiátrica de la prisión y que fueron clasificados por el equipo clínico como simuladores (n=26). Para la clasificación de los reclusos de la unidad psiquiátrica de la prisión como simuladores, los clínicos de la unidad psiquiátrica emplearon un inventario basado en la SIRS (Rogers, Bagby y Dickens, 1992).

Todos los participantes fueron evaluados en este estudio con distintas medidas de simulación, junto con escalas generales de psicopatía y psicopatología. Con el objetivo de hacer lo

más reales posibles las condiciones de simulación de los reclusos de la muestra b, se les informó de que había un incentivo económico si conseguían no ser detectados como simuladores.

Los resultados obtenidos en este estudio mostraron que, de todas las medidas empleadas, la puntuación total del SIMS (punto de corte > 14) fue la que mejores resultados de clasificación obtuvo, tal y como se muestra en la tabla 3.33.

Tabla 3.33
Correlación de la puntuación total del SIMS con las escalas de validez del PAI y la puntuación total en el SIRS (Edens et al., 2007)

	Porcentaje global de clasificación correcta ^f	ABC		Sensibilidad		Especificidad	
		ABC (muestra total)	ABC (muestras clínicas c y d)	Simuladores análogos ^b	Sospechosos de simulación ^c	Controles ^a	Pacientes ^d
Total SIMS>14	0,77	0,81 (ET: 0,04)	0,56 (ET: 0,08)	0,90	0,85	0,97	0,40

Nota. ABC: Área bajo la curva ^a n=30; ^b n=29; ^c n=30; ^d n=26; ^f N=115

Son reseñables dos aspectos de estos resultados. Por un lado, su adecuada especificidad en presos respondiendo en condición de sinceridad (por lo tanto, no parece que haya un sesgo tendente a clasificar a las personas reclusas como simuladores) y, por otro, la relativa alta sensibilidad, junto con la baja especificidad, cuando se refiere a pacientes genuinos con problemas psicopatológicos y se emplea el punto de corte de 14.

Es interesante también mencionar las altas correlaciones halladas entre la puntuación total del SIMS y las puntuaciones en las escalas de simulación del PAI y la puntuación total del SIRS en el conjunto de la muestra, tal y como se advierte en la tabla 3.34. Estos valores respaldan la validez concurrente de la puntuación total del SIMS en contexto penal.

No obstante, se echa en falta la realización de análisis complementarios ya que no se proporcionan las puntuaciones medias de cada uno de los grupos en las distintas escalas del SIMS, ni en el total, así como tampoco se especifican los resultados de sensibilidad o especificidad de cada una de las escalas de la prueba.

Tabla 3.34

Correlación de la puntuación total del SIMS con las escalas de validez del PAI y la puntuación total en el SIRS (Edens et al., 2007)

	PAI			Puntuación total en el SIRS
	Escala IMN	Índice SIM	Índice FDR	
SIMS – Total	0,84	0,68	0,45	0,81

Nota. IMN: Impresión negativa, SIM: Índice de simulación, FDR: Función discriminante de Rogers; SIRS: Structured Interview of Reported Symptoms (Rogers, Bagby y Dickens, 1992)

3.7.2. Estudio de detección de simulación de psicopatología en presos de nuevo ingreso (Laffoon, 2009)

El objetivo de este estudio fue determinar la eficacia del SIMS junto con otras pruebas en la detección de simulación en contextos carcelarios en un diseño de grupos diagnósticos. Para ello, Laffoon seleccionó una muestra de 100 presos (50 varones y 50 mujeres) de recién ingreso en un centro penitenciario, dentro de su fase de evaluación inicial. A partir de sus puntuaciones en el SIRS (Rogers et al., 1992), los participantes fueron clasificados como probable respuesta honesta o probable respuesta simulada, estableciendo una prevalencia de simulación de acuerdo a este criterio del 12%. Los autores no proporcionaron las medias o desviaciones típicas de la puntuación total o de las escalas del SIMS en cada uno de los grupos clasificados por el SIRS, sino en la muestra total y solo de la puntuación total (media = 10,07; Dt = 5,89). Tampoco proporcionaron estudios de sensibilidad y especificidad para cada una de las escalas de la prueba y solamente analizaron estos datos para la puntuación total empleando el punto de corte de 14, sin considerar puntos de corte alternativos. Los resultados muestran una sensibilidad de 0,75 y una especificidad de 0,88. Los estudios correlacionales de la puntuación total del SIMS con la puntuación total en el SIRS ($r=0,71$) y con las puntuaciones del PAI IMN ($r=0,59$), SIM ($r=0,29$) y FDR ($r=0,13$), respaldaron la validez convergente del SIMS en muestras de presos.

3.7.3. Estudio de simulación de incapacidad para ser juzgados de Lewis, Simcox y Berry (2002)

Lewis, Simcox y Berry (2002) evaluaron la eficacia de distintas pruebas habitualmente empleadas para la detección de la simulación, entre ellas el SIMS, en una muestra de 55 varones remitidos a evaluación psicológica, para valorar si eran competentes para ser juzgados y su estado mental en el momento del delito. Para ello, emplearon un diseño de grupos diagnósticos, clasificando a los participantes -de acuerdo a los resultados del SIRS (Rogers et al., 1992)- como simuladores ($n=24$; edad media de 32,5 años y Dt de 10,1) o de respuesta sincera ($n=31$; edad media de 43.5 y Dt de 14,6). Dado que en uno de los lugares de evaluación solo se aplicaron todas las

pruebas a aquellos con sospechas de simulación, esos casos fueron excluidos de los análisis de las tasas de prevalencia de simulación. A partir del resto de los casos, se calculó una tasa de prevalencia del 31,4% que fue considerada en los siguientes análisis de utilidad, de acuerdo a la detección de 16 simuladores de los 51 sujetos evaluados.

Los resultados del análisis multivariante de la varianza (MANOVA) indicaron que los simuladores obtuvieron puntuaciones significativamente superiores a las de los no simuladores en todas las escalas del SIMS (Lambda de Wilks=0,062, $F(5,49) = 149,31$, $p < 0,01$). Tal y como se puede observar en la tabla 3.35, las pruebas de t realizadas a continuación con cada escala del SIMS mostraron que, en todas ellas, había diferencias significativas entre las puntuaciones de ambos grupos. El tamaño del efecto de dicha diferencia (d de Cohen) fue mayor para la puntuación total que para las escalas de la prueba, tal y como cabría esperar, con un valor máximo de 3,0. La escala Psicosis obtuvo valores similares ($d=2,9$), mientras que las escalas Baja inteligencia y Trastornos afectivos fueron las que obtuvieron valores más bajos ($d=1,1$ y $d=1,8$ respectivamente). Este patrón de diferencias parece razonable en el contexto de evaluación y en los perfiles de simulación que cabría esperar.

Tabla 3.35
Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores. Estudio independiente de Lewis, Simcox y Berry (2002)

Escalas del SIMS	Respuesta sincera ^a		Simuladores ^b		T	d
	M	Dt	M	Dt		
Total	14,5	8,8	42,8	10,6	-10,8**	3,0
Psicosis	1,2	2,1	9,0	3,4	-10,0**	2,9
Deterioro neurológico	2,4	2,1	9,5	3,6	-9,2**	2,5
Trastornos amnésicos	2,5	3,5	9,5	3,6	-7,3**	2,0
Baja inteligencia	3,2	2,1	5,8	2,8	-3,9**	1,1
Trastornos afectivos	5,2	2,6	9,0	1,6	-6,7**	1,8

Nota. ^a $n = 31$; ^b $n = 24$ ** $p < 0,01$.

El análisis de regresión logística realizado señaló que la puntuación total del SIMS predecía significativamente el nivel de simulación, tal y como era establecida por la SIRS ($R^2=0,61$, $odds-ratio=1,27$, $p < 0,05$). Por otra parte, los autores también realizaron análisis de sensibilidad y especificidad de la puntuación total del SIMS (tabla 3.36), con un punto de corte de 16. Todos los simuladores fueron detectados por el SIMS como tales (sensibilidad = 100%), si bien se clasificaron erróneamente a un 39% de sujetos no simuladores (respuesta sincera) como simuladores. Si tenemos en cuenta las tasas de prevalencia, el SIMS mostró un valor predictivo positivo atenuado ya que solo

el 54% de los sujetos identificados como simuladores por el SIMS eran realmente simuladores. Por el contrario, todos los individuos clasificados como no-simuladores eran realmente no-simuladores (valor predictivo negativo = 100%). No se proporcionó información sobre datos de las escalas del SIMS en cuanto a sensibilidad o especificidad.

Tabla 3.36
Precisión predictiva de la puntuación total del SIMS. Estudio independiente de Lewis, Simcox y Berry (2002)

Total SIMS \geq 16	Precisión
Sensibilidad	1,00
Especificidad	0,61
Valor predictivo positivo	0,54
Valor predictivo negativo	1,00
Tasa de acierto	0,73

Nota. N=51.

3.7.4. Estudio de simulación de incapacidad para ser juzgados de Heinze y Purisch (2001)

Heinze y Purisch estudiaron la utilidad del SIMS en una muestra de 57 varones identificados como simuladores mediante un diseño de grupos diagnósticos (edad media de 36,5, Dt de 9,2 años). Estos sujetos habían sido identificados como tales en una muestra más amplia de 438 varones, de acuerdo a una evaluación clínica que no incluía el SIMS entre sus instrumentos. La tasa de prevalencia de simulación en esta muestra fue del 13%. El acuerdo interjueces entre el psicólogo evaluador y el equipo multidisciplinar de tratamiento en la clasificación de estos sujetos como simuladores fue alto (0,92).

Las puntuaciones medias en el SIMS de esta muestra fueron comparadas con las medias hipotéticas de poblaciones clínicas y de simuladores en el SIMS, según una revisión bibliográfica de la información disponible a este respecto en el momento en que se escribió el artículo. Como puede advertirse en la tabla 3.37, las puntuaciones del grupo de simuladores de este estudio fueron notable y significativamente superiores a las de la hipotética muestra clínica. No obstante, estas comparaciones se encuentran seriamente limitadas por la falta de muestras específicas de control y referencia y por el uso de medias hipotéticas basadas en la revisión de la bibliografía existente en ese momento, bastante limitada. Los estudios posteriores realizados parecen apuntar, de hecho, a

puntuaciones superiores en las muestras de pacientes clínicos genuinos, tal y como se muestra en las diversas investigaciones recogidas en este capítulo.

Tabla 3.37
Puntuaciones medias en el SIMS del grupo de simuladores de incapacidad para ser sometidos a juicio y medias estimadas de la población clínica y de simuladores en el estudio de Heinze y Purisch (2001)

Escala del SIMS	Simuladores		Muestra clínica hipotética	t
	M	ET	M	
Total	33,66	2,30	8,09	11,12**
Psicosis	6,84	1,06	0,67	5,82**
Deterioro neurológico	6,97	0,62	1,00	9,63**
Trastornos amnésicos	9,26	1,04	1,05	7,89**
Baja inteligencia	6,45	0,89	1,97	5,03**
Trastornos afectivos	8,16	0,63	3,94	6,70**

Nota. n^a= 57; ET=Error típico

** p < 0,01; *** p < 0,005

Dado que el eje central de este estudio fue conocer la sensibilidad del SIMS para reconocer a una muestra previamente clasificada como simuladores por jueces independientes, en la tabla 3.38 se muestran estas estimaciones de sensibilidad, empleando unos puntos de corte específicos para su estudio (aunque los autores no informan sobre cómo fueron establecidos) y que son inferiores en un 1 punto a los recomendados por los autores originales (Smith y Burger, 1997; Widow y Smith, 2005).

Tabla 3.38
Sensibilidad de las escalas del SIMS para la detección de simuladores en el estudio de Heinze y Purisch (2001)

Escala del SIMS	Sensibilidad
Total > 13	0,87
Psicosis > 0	0,89
Deterioro neurológico > 1	1,00
Trastornos amnésicos > 1	1,00
Baja inteligencia > 1	1,00
Trastornos afectivos > 4	0,89

Nota. N=57.

3.7.5. Estudio de simulación de incapacidad para ser juzgados de Vitacco, Rogers, Gabel y Munizza (2006)

El objetivo de esta investigación fue determinar la eficacia diagnóstica de distintas herramientas de *screening* de simulación, en el contexto de la evaluación de la competencia de un individuo para someterse a un juicio, mediante un diseño de grupos diagnósticos. La muestra estaba compuesta por 100 participantes varones procedentes de un contexto forense criminal, con diagnósticos de trastornos psicóticos (68%), trastornos del estado de ánimo (17%), abuso de sustancias (8%), trastornos de la personalidad (6%) y trastornos cognitivos (1%), siendo frecuente la existencia de comorbilidad entre los distintos trastornos. El 80% habían sido juzgados por delitos graves, mientras el 20% por faltas menores. La mayoría de la muestra (70%) estaba acusada de un crimen violento y se enfrentaba a largas condenas de cárcel. De acuerdo a sus resultados en el SIRS, 79 fueron clasificados como no simuladores, mientras que 21 fueron considerados probables simuladores. En la siguiente tabla se pueden observar sus puntuaciones en el SIMS, en cada una de sus escalas y en el total en cada grupo, así como el tamaño del efecto (*d* de Cohen). Como se puede advertir, se obtienen valores muy destacados, siendo la puntuación total en el SIMS la variable que mayor tamaño del efecto mostró de las distintas medidas consideradas en el estudio (M-FAST, ECST-R ATP).

Tabla 3.39
Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores en el estudio de Vitacco et al. (2006)

Escala del SIMS	Respuesta sincera ^a		Simuladores ^b		<i>F</i>	<i>d</i>
	M	Dt	M	Dt		
Total	14,94	9,35	45,33	11,77	156,49	3,07
Psicosis	2,46	2,58	9,57	2,96	118,49	2,67
Deterioro neurológico	2,51	2,61	9,43	3,59	98,15	2,43
Trastornos amnésicos	2,46	3,16	10,29	4,01	90,46	2,34
Baja inteligencia	2,57	1,79	7,52	3,19	88,16	2,30
Trastornos afectivos	4,94	2,38	8,52	2,32	38,18	1,51

Nota. n^a = 79; n^b = 21

En la tabla 3.40 se aborda la sensibilidad, especificidad, poder predictivo positivo, poder predictivo negativo y tasa de acierto del SIMS, tanto en la puntuación total como en las escalas. Se puede advertir los altos resultados de sensibilidad, a la par que los más modestos e, incluso bajos, de especificidad, con su correspondencia en el caso del poder predictivo positivo y negativo y en la tasa de acierto.

Tabla 3.40
Sensibilidad y especificidad de las escalas del SIMS para la detección de simuladores en el estudio de Vitacco et al. (2006)

Escalas del SIMS	Sensibilidad	Especificidad	Poder predictivo positivo	Poder predictivo negativo	Tasa de acierto
Total > 14	1,00	0,65	0,43	1,00	0,72
Psicosis > 1	1,00	0,51	0,35	1,00	0,61
Deterioro neurológico > 2	0,95	0,66	0,43	0,98	0,72
Trastornos amnésicos > 2	0,90	0,67	0,42	0,96	0,72
Baja inteligencia > 2	0,90	0,57	0,36	0,96	0,64
Trastornos afectivos > 5	0,90	0,71	0,45	0,97	0,75

3.8. Estudios de grupos diagnósticos conocidos o diseños autosuficientes en contexto forense civil con el SIMS

3.8.1. Estudio de simulación de lesiones para la obtención de indemnizaciones o compensaciones económicas de Alwes (2006)

El estudio de Alwes (2006) de grupos diagnósticos contó con una muestra de 308 personas que estaban siendo evaluadas en un contexto legal, para la determinación de indemnizaciones o compensaciones económicas por lesiones (contexto forense civil). El objetivo de esta investigación fue evaluar la precisión del SIMS en la clasificación de simuladores, tanto de psicopatología como de trastorno neurocognitivo. La muestra, con una edad media de 41,3 años ($Dt=11,2$, rango 18 – 71), estaba formada por un 65,9% de varones.

Los participantes fueron clasificados como simuladores o de respuesta honesta en dos áreas: psicopatología y problemas neurocognitivos. Para la determinación de la simulación, en el primer caso, se empleó el SIRS y los puntos de corte recomendados en la literatura (Rogers et al., 1992); para la determinación de la simulación de problemas neurocognitivos, se emplearon tres pruebas estándar en el campo, con sus puntos de corte establecidos según las investigaciones previas TOMM (Tombaugh, 1996), VSVT (Slick, Hopp, Strauss y Thompson, 1997, p. 199) y LMT (Inman et al., 1998). De acuerdo a su evaluación, cada caso fue clasificado como simulador, respuesta honesta o indeterminado.

En concreto, en el área de la psicopatología, de los 308 casos, 23 fueron clasificados como simuladores, 172 con respuesta honesta y 113 indeterminados; en el área de problemas neurocognitivos, de los 308 casos, 75 fueron clasificados como simuladores, 178 como de respuesta honesta y 55 como indeterminados. Finalmente, se hicieron dos clasificaciones adicionales, una que comparaba a los que simulaban tanto en psicopatología como en problemas neurocognitivos (n=13) con los que habían respondido honestamente en ambas áreas (n=121) y otra que comparaba a aquellos individuos que habían simulado en alguna de las dos áreas, psicopatología o problemas neurocognitivos (n=85), frente a los que habían respondido honestamente a ambas (n=121). A continuación, se puede ver el número de casos agrupados en cada una de las categorías estipuladas.

Tabla 3.41
Descripción de los grupos de la investigación de Alwes (2006)

Tipo de simulación	Grupo simulador	Grupo de respuesta honesta	Grupo indeterminado
Psicopatológica	23	172	113
Neurocognitiva	75	178	55
Ambas	13	121	174
Alguna de las dos	85	121	102

Las tablas de 3.42 a 3.45 incluyen las puntuaciones del SIMS para cada uno de los grupos, en los distintos contrastes realizados en la investigación.

Tabla 3.42
Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores de psicopatología en el estudio de Alwes (2006)

Escala del SIMS	Respuesta sincera ^a		Simuladores ^b		t	d
	M	Dt	M	Dt		
Total	14,41	6,47	31,57	8,85	11,398**	2,55
Psicosis	0,31	0,64	3,35	3,16	4,605**	3,27
Deterioro neurológico	3,42	2,53	8,13	2,72	8,331**	1,85
Trastornos amnésicos	2,53	2,54	7,17	3,17	7,975**	1,77
Baja inteligencia	2,15	1,60	3,70	1,96	4,226**	0,94
Trastornos afectivos	5,95	2,37	9,17	1,47	9,064**	1,42

Nota. n^a = 172; n^b = 23 ** p < 0,01.

Tabla 3.43***Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores de problemas neurocognitivos en el estudio de Alwes (2006)***

Escala del SIMS	Respuesta sincera ^a		Simuladores ^b		<i>t</i>	<i>d</i>
	M	Dt	M	Dt		
Total	15,55	7,07	24,76	10,75	6,821**	1,13
Psicosis	0,46	0,90	1,61	2,61	3,726**	0,82
Deterioro neurológico	3,75	2,83	6,21	3,63	5,235**	0,80
Trastornos amnésicos	2,53	2,55	5,97	3,38	7,920**	1,23
Baja inteligencia	2,31	1,67	3,04	2,08	2,682**	0,41
Trastornos afectivos	6,44	2,42	7,88	2,21	4,424**	0,61

Nota. n^a= 178; n^b= 75 ** p < 0,01.**Tabla 3.44*****Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores de psicopatología y de problemas neurocognitivos (ambos) en el estudio de Alwes (2006)***

Escala del SIMS	Respuesta sincera ^a		Simuladores ^b		<i>t</i>	<i>d</i>
	M	Dt	M	Dt		
Total	13,63	6,03	34,85	8,96	11,438**	3,37
Psicosis	0,31	0,68	4,38	3,69	3,974**	4,23
Deterioro neurológico	3,12	2,40	8,57	3,10	7,521**	2,20
Trastornos amnésicos	2,11	2,35	8,46	2,70	9,153**	2,67
Baja inteligencia	2,18	1,64	4,31	1,93	4,370**	1,28
Trastornos afectivos	5,84	2,33	9,08	1,38	7,386**	1,45

Nota. n^a= 121; n^b= 13 ** p < 0,01.**Tabla 3.45*****Puntuaciones medias en el SIMS del grupo de respuesta sincera y de simuladores de psicopatología o de problemas neurocognitivos (alguno de los dos) en el estudio de Alwes (2006)***

Escala del SIMS	Respuesta sincera ^a		Simuladores ^b		<i>t</i>	<i>d</i>
	M	Dt	M	Dt		
Total	13,63	6,03	25,06	10,38	9,126**	1,46
Psicosis	0,31	0,68	1,66	2,51	4,807**	0,94
Deterioro neurológico	3,12	2,40	6,38	3,51	7,429**	1,14
Trastornos amnésicos	2,11	2,35	5,92	3,33	9,089**	1,39
Baja inteligencia	2,18	1,64	3,02	2,035	3,282**	0,47
Trastornos afectivos	5,84	2,33	8,05	2,19	6,859**	0,97

Nota. n^a= 121; n^b= 85 ** p < 0,01.

En la tabla 3.46 se puede advertir una perspectiva global de los tamaños del efecto para cada una de las escalas del SIMS según los grupos de contraste. A este respecto, los datos muestran que la escala Psicosis es la que mejor discrimina en simuladores de psicopatología, mientras que la escala Trastornos Amnésicos en simuladores de problemas neurocognitivos y la escala Total cuando no se combinan indistintamente cualquiera de los dos simuladores. También puede observarse como la escala Baja inteligencia es la que peor capacidad discriminativa presenta en cualquiera de las condiciones.

Tabla 3.46
Tamaño del efecto (d de Cohen en valor absoluto) de las escalas del SIMS en la detección de simuladores en el estudio de Alwes (2006)

	Psicopatología Honestos vs. Simuladores	Problemas neurocognitivos Honestos vs. simuladores	Psicopatología y Problemas neurocognitivos Honestos vs. simuladores	Psicopatología O problemas neurocognitivos Honestos vs. simuladores
Total	2,55	1,13	3,37	1,46
Psicosis	3,27	0,82	4,23	0,94
Deterioro neurológico	1,85	0,80	2,20	1,14
Trastornos amnésicos	1,77	1,23	2,67	1,39
Baja inteligencia	0,94	0,41	1,28	0,47
Trastornos afectivos	1,42	0,61	1,45	0,97

El estudio también analizó las áreas bajo la curva para la puntuación Total del SIMS en cada una de las condiciones, obteniéndose valores satisfactorios (para simulación de psicopatología ABC=0,899, ET=0,026; para simulación de problemas neurocognitivos ABC=0,755, ET=0,035; para simulación de psicopatología y problemas neurocognitivos ABC=0,983, ET=0,010; y para simulación de alguno de los dos ABC=0,822; ET = 0,031).

Tabla 3.47
Sensibilidad y especificidad de la puntuación total en el SIMS en los distintos grupos en el estudio de Alwes (2006)

	Psicopatología Honestos vs. simuladores	Problemas neurocognitivos Honestos vs. simuladores	Psicopatología y Problemas neurocognitivos Honestos vs. simuladores	Psicopatología O problemas neurocognitivos Honestos vs. simuladores
Total > 14				
Sensibilidad	0,957	0,80	1,00	0,812
Especificidad	0,541	0,472	0,57	0,57
Tasa de acierto	0,568	0,569	0,612	0,67
Total > 16				
Sensibilidad	0,957	0,747	1,00	0,765
Especificidad	0,669	0,601	0,703	0,703
Tasa de acierto	0,687	0,644	0,731	0,728

Es interesante también mencionar el análisis realizado sobre la validez incremental del SIMS y del MMPI-2, consistente en el cálculo de la capacidad predictiva de estas pruebas sobre cada una de las pruebas criterio (SIRS, TOMM, VSVT y LMT). En primer lugar, se analizó si las escalas de validez del MMPI-2 aportaban validez incremental a la predicción de los criterios, una vez considerada la puntuación total del SIMS. Los resultados obtenidos mostraron diferencias según la prueba criterio considerada: las escalas F y Fb (pero no Fp) añadían validez incremental a la predicción de la puntuación en el SIRS, la escala F a la predicción de la puntuación en el VSVT y en el LMT y ninguna escala del MMPI-2 añadía validez incremental a la hora de predecir la puntuación en el ensayo 2 del TOMM, una vez tenida en cuenta la puntuación Total en el SIMS.

Resultó relevante también el estudio inverso, esto es, el análisis de si la introducción de la puntuación Total en el SIMS añadía validez incremental, una vez tenida en cuenta en cada regresión las escalas de infrecuencia del MMPI-2 (F, Fb y Fp). Los resultados obtenidos pusieron de manifiesto que, para todas las pruebas criterio, la validez incremental en la predicción se veía aumentada al añadir el SIMS al MMPI-2 dentro de las variables predictoras.

Tabla 3.48
Tamaño del efecto (*d* de Cohen en valor absoluto) de las escalas del MMPI-2, SIMS, SIRS, TOMM, VSVT y LMT en la detección de simuladores en el estudio de Alwes (2006)

	Psicopatología Honestos vs. simuladores	Problemas neurocognitivos Honestos vs. Simuladores	Psicopatología y Problemas neurocognitivos Honestos vs. Simuladores	Psicopatología O problemas neurocognitivos Honestos vs. simuladores
MMPI-2 F	2,70	1,16	3,13	1,59
MMPI-2 Fb	2,92	0,94	3,21	1,53
SIMS Total	2,55	1,13	3,37	1,46
SIRS	nc	0,99	nc	nc
TOMM	1,12	Nc	nc	nc
VSVT	1,01	nc	nc	nc
LMT	1,25	nc	nc	nc

Nota. Nc: no calculable porque fue la prueba criterio empleada para determinar los grupos de simuladores vs. respuesta honesta en esa área de simulación.

Cabe mencionar finalmente que si bien la autora señala la necesidad de estudios que aborden la capacidad diagnóstica de las escalas del SIMS, indica también que dicho aspecto no fue explorado en su investigación y que deberá ser abordado en un futuro.

3.8.2. Estudio de simulación de síntomas depresivos para obtención de incapacidad de Clegg, Fremouw y Mogge (2009)

El estudio de Clegg et al. (2009) de grupos diagnósticos contó con una muestra de 64 personas que estaban siendo evaluadas para la concesión de discapacidad en la seguridad social por deterioro funcional. La mayoría de los casos (88,4%) fueron diagnosticados principalmente de trastornos del estado de ánimo, mientras que el resto de participantes presentaban trastornos de ansiedad, a excepción de un adulto con trastorno de déficit de atención e hiperactividad. La edad media de la muestra fue de 42,5 años ($Dt=11,9$), siendo el 38% de los casos varones. Además, se seleccionó una muestra de 63 personas procedente de la población general, que fueron asignadas aleatoriamente a un grupo de respuesta sincera o de simulación de depresión. Su edad media fue de 39,9 años ($Dt=15,2$), siendo el 68% de los casos mujeres.

Todos los individuos fueron evaluados con el SIMS y el ADI (Assessment of Depression Inventory; Mogge y LePage, 2004). A la muestra evaluada para la concesión de discapacidad se le aplicó, además, la SIRS (*Structured Interview of Reported Symptoms*; Rogers et al., 1992), siendo clasificado cada participante, a partir de sus resultados en esta entrevista, como simulador o de

respuesta honesta. De los 64 casos, 36 personas fueron clasificadas como de respuesta honesta y 20 como posibles simuladores. Los 8 casos restantes no pudieron ser clasificados claramente por la SIRS como simuladores o de respuesta honesta, por lo que fueron excluidos de los análisis.

La muestra comunitaria fue previamente evaluada con el SCL-90-R (Derogatis, 1996) para descartar la presencia de sintomatología psicopatológica. Tres casos fueron excluidos por este motivo, quedando una muestra final de 30 casos en la condición de respuesta honesta y 30 en la de simulación.

En la tabla 3.49 aparecen las medias y desviaciones típicas de cada grupo en el SIMS. Los datos evidencian notables diferencias entre las puntuaciones de los grupos de respuesta honesta (tanto de la muestra comunitaria como de la procedente de la valoración de discapacidad) y los grupos de simuladores, apuntando a la satisfactoria capacidad del SIMS para discriminar entre ambos grupos. La puntuación total del SIMS y la escala Trastornos afectivos fueron aquellas que mayores diferencias establecieron entre ambos grupos, tal y como cabría esperar, puesto que este era el ámbito de simulación (los análogos precisamente fueron instruidos para simular síntomas de depresión mayor bajo el supuesto de obtener una pensión por discapacidad).

Tabla 3.49
Puntuaciones medias y desviaciones típicas en las escalas del SIMS de los grupos experimentales y de respuesta sincera en el estudio de Clegg et al. (2009)

Escala del SIMS	Respuesta honesta				Simulación				Tamaño del efecto d de Cohen	
	Muestra comunitaria (n=30)		Muestra valoración discapacidad (n=35)		Muestra comunitaria (n=30)		Muestra valoración discapacidad (n=20)		Muestra comunitaria (hon. vs. sim.)	Muestra valoración discapacidad (hon. vs. sim.)
	M	Dt	M	Dt	M	Dt	M	Dt	d	d
Total	7,7	2,9	17,1	6,6	26,6	14,6	30,2	9,1	1,8	1,6
Psicosis	0,2	0,4	0,5	1,3	2,7	4,5	3,4	3,5	0,8	1,1
Deterioro neurológico	1,3	1,6	4,2	2,7	5,6	4,6	6,0	3,1	1,3	0,6
Trastornos amnésicos	0,5	0,8	3,3	2,7	5,8	4,9	7,4	2,9	1,5	1,5
Baja inteligencia	1,4	1,5	2,6	2,4	2,8	2,3	3,6	2,4	0,7	0,4
Trastornos afectivos	4,4	1,5	6,6	1,8	9,6	1,7	9,7	1,6	3,3	1,8

Clegg (2007) estudió también la utilidad de la prueba en la clasificación en ambos grupos, obteniendo sus valores de sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo para diversos puntos de corte alternativos del total del SIMS. En la tabla 3.50 se detallan los

valores hallados para los puntos de corte de >14, >16 y >19 ya que son los propuestos específicamente por Clegg para el uso del SIMS, de acuerdo a sus resultados.

Para justificar su recomendación del punto de corte de 19 en el contexto de su estudio, Clegg afirma que ese punto de corte permite mejorar notablemente la especificidad (y por lo tanto el valor predictivo positivo), lo que conlleva descartar a un mayor número de casos para una siguiente fase de evaluación, sin que esto afecte dramáticamente a los niveles de sensibilidad y valor predictivo negativo. En palabras de Clegg (2007):

Ideally a screening instrument cut-score would yield nearly perfect sensitivity and NPP, but it must have reasonable specificity and PPP to be useful. To look these results another way, using a SIMS total cut-score of > 14 with the actual disability seekers in this study would have excluded only 20%, or 11 out of 55, of them from requiring more extensive malingering testing (...) However, using a cut-score of > 19 with this sample would have eliminated nearly half (47.3%) of these individuals from needing more extensive testing, and would have resulted in only two false-negatives. (p.24)

Tabla 3.50

Precisión predictiva de la puntuación total del SIMS al clasificar a los participantes como del grupo de respuesta honesta o del grupo de simuladores. Muestra de evaluación de discapacidad. Estudio de Clegg (2007)

Total SIMS	Punto de corte >14	Punto de corte >16	Punto de corte >19
Sensibilidad	1,00	0,95	0,90
Especificidad	0,37	0,51	0,69
Poder predictivo positivo	0,48	0,53	0,62
Poder predictivo negativo	1,00	0,95	0,92

Clegg (2007) estudió también la utilidad de los puntos de corte de cada una de las escalas, sin sugerir, en este caso, puntos de corte alternativos. Según los resultados obtenidos, que aparecen en la tabla 3.51. y tal y como cabría esperar, las escalas que mayor sensibilidad presentan son Trastornos afectivos y Trastornos amnésicos, si bien obtienen también valores bajos de especificidad. Cabría plantearse, tal y como hizo el autor de este estudio con la puntuación total, aumentar el valor de punto de corte para obtener una mayor especificidad.

Tabla 3.51

Precisión predictiva de la puntuación en cada una de las escalas del SIMS al clasificar a los participantes como del grupo de respuesta honesta o del grupo de simuladores. Muestra clínica. Estudio independiente de Clegg (2007)

Escalas del SIMS	Punto de corte	Sensibilidad	Especificidad	Valor predictivo positivo	Valor predictivo negativo
Psicosis	> 1	0,65	0,94	0,87	0,83
Deterioro neurológico	> 2	0,80	0,40	0,43	0,78
Trastornos amnésicos	> 2	1,00	0,54	0,56	1,00
Baja inteligencia	> 2	0,70	0,63	0,52	0,79
Trastornos afectivos	> 5	1,00	0,29	0,44	1,00

3.8.3. Estudio de simulación de lesiones para la obtención de indemnizaciones o compensaciones económicas de Wisdom et al. (2010)

El estudio de Wisdom et al. (2010) de grupos diagnósticos contó con una muestra de 33 personas que alegaban lesiones tras un traumatismo craneal o una reacción física o psicológica adversa ante un determinado traumatismo. Independientemente de la causa de evaluación, un 60% se quejaba de problemas físicos, cognitivos y emocionales, mientras que un 10% informó sobre problemas cognitivos, junto con problemas físicos o emocionales. El 30% restante refirió quejas físicas o emocionales pero no en el área cognitiva.

A partir de una extensa batería de pruebas y siguiendo los criterios de (Slick et al., 1999) o (Bianchini et al., 2005), según el motivo de evaluación, se clasificó a los evaluados como de respuesta honesta (n=11), probable simulación (n=17) o clara simulación (n=5). En la tabla 3.52 se pueden observar las puntuaciones totales del SIMS en cada uno de los grupos.

Tabla 3.52

Puntuaciones medias y desviaciones típicas en la puntuación total del SIMS de los grupos de respuesta honesta, probable simulación y clara simulación en el estudio de Wisdom et al. (2010)

	Respuesta honesta ^a		Probable simulación ^b		Clara simulación ^c	
	M	Dt	M	Dt	M	Dt
Puntuación Total - SIMS	15,3	4,7	26,6	8,9	26	6,5

Nota. n^a = 11; n^b = 17; n^c = 5.

La tabla 3.53 muestra la sensibilidad y especificidad de los distintos puntos de corte alternativos valorados. Los autores de la investigación recomiendan el uso del punto de corte de 24 para asegurar una alta especificidad y la comisión de un bajo número de falsos positivos. En este sentido, enfatizan la

utilidad del SIMS como evidencia diagnóstica, más que como prueba de *screening*, priorizando la especificidad frente a la sensibilidad.

Tabla 3.53
Sensibilidad y especificidad de los distintos puntos de corte alternativos para la puntuación total del SIMS en el estudio de Wisdom et al. (2010)

Punto de corte en Total SIMS	Sensibilidad	Especificidad
Total ≥ 14	0,96	0,64
Total ≥ 21	0,68	0,73
Total ≥ 22	0,64	0,82
Total ≥ 23	0,55	0,91
Total ≥ 24	0,55	1,00
Total ≥ 25	0,50	1,00

3.8.4. Estudio con pacientes sometidos a evaluación en un centro de valoración (Merten et al., 2007)

El estudio de Merten et al. (2007) analizó retrospectivamente todos los casos evaluados en un centro de valoración neurológico-psiquiátrico, entre 2004 y 2006, a los que se les había aplicado alguna prueba de exageración de síntomas ya fuera el WMT (Green, 2005) o el SIMS, las dos pruebas disponibles con este objetivo en el protocolo de valoración del centro. No se dispuso de criterios externos que determinaran qué casos eran pacientes genuinos y cuáles simuladores, por lo que se analizaron exclusivamente las puntuaciones de los evaluados en estas pruebas. La muestra total considerada fue de 378 casos, de los cuales en 339 se disponía del protocolo del WMT, en 198 del SIMS y en 159 casos los datos de ambas pruebas. Un 37% de las personas evaluadas con el WMT obtuvieron puntuaciones superiores a los puntos de corte recomendados, siendo clasificadas como con sospecha de simulación; en el caso del SIMS, el porcentaje de personas evaluadas con puntuaciones superiores al punto de corte de 16 fue del 36%. En los casos evaluados con ambas pruebas, la concordancia de ambas pruebas en la clasificación de la persona evaluada como con o sin sospecha de simulación se estableció en un 77% de los casos..

En la siguiente tabla puede verse las puntuaciones en el SIMS en los grupos clasificados por el WMT como con o sin sospecha de simulación.

Tabla 3.54

Puntuaciones medias y desviaciones típicas del SIMS y prueba t de Student entre los participantes clasificados como con y sin sospecha de simulación de acuerdo a sus resultados en el WMT (Green, 2005)

Escala del SIMS	Participantes clasificados por el WMT como SIN sospecha de simulación ^a		Participantes clasificados por el WMT como CON sospecha de simulación ^a		g.l.	t	p	d
	M	Dt	M	Dt				
Puntuación Total - SIMS	11,6	6,7	21,2	11,1	47,3	5,15	<0,05	1,05
Psicosis	0,6	1,0	1,5	2,6	157	4,18	<0,05	0,46
Deterioro neurológico	2,7	2,1	5,6	3,0	51	5,66	<0,05	1,12
Trastornos amnésicos	1,8	2,1	4,3	3,2	47,3	5,16	<0,05	0,92
Baja inteligencia	1,8	1,2	2,5	1,9	49,1	2,11	<0,05	0,44
Trastornos afectivos	4,8	3,2	7,3	3,5	61	4,03	<0,05	0,75

Nota. n^a= 120; n^b= 39.

3.8.5. Estudio con pacientes con sintomatología somatoforme y de baja laboral (González-Ordi, Santamaría y Fernández Marín, 2010)

En este estudio se analizó la capacidad del SIMS para discriminar la presencia de simulación en pacientes que presentaban sintomatología consistente con trastornos de tipo somatoforme y cursaban baja por AT y por ITCC.

La muestra comprendía 61 usuarios con este tipo de patología que acudieron a la mutua ASEPEYO y que fueron remitidos a evaluación. De acuerdo a criterios de sospecha clínica (véase Samuel y Mittenberg, 2006), una jueza externa clasificó a los pacientes en el grupo de probable paciente clínico o de probable simulador. De esta manera, quedaron configurados dos grupos:

- Grupo de probable paciente clínico genuino. Compuesto por 27 participantes que oscilaban en un rango de edad desde los 18 a 59 años (media = 38,3 años, Dt = 11,2 años).
- Grupo de probable simulador. Compuesto por 34 participantes que oscilaban en un rango de edad desde los 23 a 59 años (media = 37,9 años, Dt= 9,7 años)

En la siguiente tabla se pueden observar los estadísticos descriptivos del SIMS en ambos grupos, así como los resultados del ANOVA que fue llevado a cabo y el tamaño del efecto de la diferencia (d de Cohen) y a continuación el análisis de sus estadísticos de utilidad diagnóstica con los puntos de corte de 14 y 16.

Tabla 3.55
Estadísticos descriptivos del SIMS del grupo de probable pacientes clínico genuino y de probable simulador en el estudio de González Ordi et al. (2010)

Escala del SIMS	Probable paciente clínico ^a		Probable simulador ^b		F	d
	M	Dt	M	Dt		
Total	11,7	4,6	23,2	8,9	37,0***	1,6
Psicosis	0,4	0,6	2,2	2,3	15,5***	1,1
Deterioro neurológico	2,8	2,6	5,6	3,6	11,5***	0,9
Trastornos amnésicos	1,2	1,4	4,4	2,9	27,5***	1,4
Baja inteligencia	1,8	1,3	2,9	1,8	6,9*	0,7
Trastornos afectivos	5,5	1,9	8,2	2,5	21,21***	1,2

Nota. n^a= 27; n^b= 34 *** p < 0,001 *p<0,05

Tabla 3.56
Precisión predictiva de la puntuación total del SIMS al clasificar a los participantes como del grupo de probable paciente clínico o de probable simulador en el estudio de González Ordi et al. (2010)

Total SIMS	Punto de corte >14	Punto de corte >16
Sensibilidad	0,88	0,82
Especificidad	0,81	0,93
Poder predictivo positivo	0,66	0,83
Poder predictivo negativo	0,94	0,92

3.8.6. Estudio con pacientes diagnosticados de trastorno mixto ansioso-depresivo y adaptativo en el contexto médico-legal (Blasco Saiz y Pallardó Durá, 2013)

El objetivo de este trabajo preliminar fue el estudio del patrón de exageración de síntomas en una pequeña muestra de pacientes diagnosticados de trastorno mixto ansioso-depresivo y adaptativo en contexto médico-legal. Todos los participantes fueron evaluados con el SIMS y el Inventario Multifásico de Personalidad de Minnesota 2 Forma Reestructurada (MMPI-2-RF), siguiendo el protocolo establecido en González-Ordi et al. (2012) y Santamaría y González-Ordi (2012) y la metodología indicada en Santamaría (2012).

La muestra del estudio estuvo compuesta por 26 participantes, con edades comprendidas entre los 18 y 65 años, que cumplían criterios diagnósticos para un trastorno mixto ansioso-depresivo y adaptativo reactivo a agentes estresores identificables, de más de un mes de duración, remitidos por mutuas de accidente de trabajo para tratamiento psicológico. Se excluyeron para este

estudio aquellos pacientes remitidos que cumplieran criterios diagnósticos de retraso mental, abuso o dependencia de sustancias, trastornos psiquiátricos severos (trastorno bipolar, trastorno psicótico, etc.) y que presentasen enfermedades médicas que pudieran cursar con sintomatología ansioso-depresiva.

Posteriormente, se asignó a los pacientes a dos grupos, con o sin sospecha de simulación, en función de que presentasen una serie de siguientes indicadores clínicos establecidos por González-Ordi et al. (2012) y adaptados a la patología mixta ansioso-depresiva objeto de este estudio. En función de estos criterios de asignación, se formaron dos grupos: el grupo de “pacientes con sospecha de simulación”, compuesto por 14 individuos (12 mujeres y 2 varones), con una edad media de 41,3 años (Dt = 11,6) y el grupo de “pacientes sin sospecha de simulación”, compuesto de 12 pacientes (8 mujeres y 4 varones), con una edad media de 33,3 años (Dt = 6,7).

Con el objetivo de determinar la existencia de diferencias significativas entre ambos grupos en las puntuaciones del SIMS y del MMPI-2-RF, se realizaron diversos análisis no paramétricos para dos grupos independientes (U de Mann-Whitney). Los resultados referidos a las escalas del SIMS se muestran en la siguiente tabla y permiten observar la existencia de diferencias significativas, tanto para la puntuación Total, como para las escalas Deterioro neurológico, Trastornos amnésicos y Trastornos afectivos.

Tabla 3.57
Medias, desviaciones típicas y estadístico U de Mann-Whitney del grupo con sospecha de simulación y del grupo sin sospecha de simulación del estudio de Blasco Saiz y Pallardó Durá (2013)

Escalas del SIMS	Sin sospecha de simulación (n=12)	Con sospecha de simulación (n=14)	U
	M (Dt)	M (Dt)	
Total SIMS	11,25 (3,44)	26,43 (8,64)	2,5***
Psicosis (Ps)	0,25 (0,45)	2,07 (3,52)	43,5
Deterioro neurológico (Dn)	1,67 (1,61)	7,00 (2,45)	8,0***
Trastornos amnésicos (Am)	0,67 (1,23)	5,36 (3,25)	8,5***
Baja inteligencia (Bi)	2,17 (1,53)	3,00 (1,41)	59,0
Trastornos afectivos (Af)	6,50 (1,88)	9,00 (2,42)	34**

p<0,01 *p<0,001

3.9. Estudios en muestras de pacientes clínicos

3.9.1. Estudio con muestras de pacientes psiquiátricos (Dandachi-FitzGerald et al., 2011)

El objetivo de la investigación fue determinar la prevalencia de exageración de síntomas en una muestra mixta de pacientes psiquiátricos ambulatorios. Para ello, los investigadores contaron con 183 pacientes psiquiátricos ambulatorios adultos (66,7% varones, media = 34,7 años; Dt = 12,52) con distintas psicopatologías (trastorno por déficit de atención, trastornos del espectro autista, trastornos afectivos, trastornos psicóticos, trastornos de la personalidad...). La evaluación fue realizada como parte de la intervención y no dentro de un contexto forense, por lo que no se sospechaba la existencia de incentivos que motivaran la presencia de simulación. Se aplicaron dos medidas de simulación de síntomas, el *Amsterdam Short-Term Memory Test* (ASTM; Schmand y Lindeboom, 2005) y el SIMS. Un total del 34% de los casos obtuvieron puntuaciones indicativas de exageración de síntomas en el ASTM, el SIMS o en ambos (ambas pruebas, ASTM y SIMS mostraron una correlación modesta de $-0,22$). En concreto, un 22% de los pacientes obtuvieron puntuaciones superiores al punto de corte de 16 en el SIMS. En la tabla 3.58 se pueden observar las puntuaciones medias en el SIMS en este grupo que, si bien no aparecían originalmente en el artículo, fueron proporcionadas por la autora principal tras contacto por correo electrónico (Dandachi-FitzGerald, comunicación personal, 30 de enero de 2013).

Tabla 3.58
Descriptivos en el SIMS del grupo de pacientes psiquiátricos ambulatorios. Estudio independiente de Dandachi-FitzGerald et al. (2011)

Escalas del SIMS	M	Dt
Total	11,78	8,17
Psicosis	0,92	1,57
Deterioro neurológico	2,02	2,36
Trastornos amnésicos	2,89	2,78
Baja inteligencia	1,36	1,33
Trastornos afectivos	4,64	2,58

Nota. N=183 pacientes psiquiátricos ambulatorios.

3.9.2. Estudio con muestras de pacientes con esquizofrenia (Peters, Jelacic, Moritz, Hauschildt y Jelinek, 2013)

Peters, Jelacic, et al. (2013) tuvieron como objetivo analizar cuál era la especificidad del SIMS cuando se empleaba en pacientes con esquizofrenia. Para ello, seleccionaron dos muestras, una de 18 pacientes (edad M=36,8, Dt=13,7, 22% mujeres) y otra de 23 pacientes (edad M=36,7, Dt=12,5, 35% mujeres), a los que aplicaron el SIMS junto con una batería de pruebas neuropsicológicas y de valoración clínica. En ambas muestras se excluyeron aquellos pacientes que pudieran estar en situación de litigancia o que presentaran algún incentivo para exagerar sus síntomas, cribando, por lo tanto, la muestra para poder considerarla de tipo netamente clínico.

Aplicando el uso del punto de corte de 16 para la puntuación total del SIMS, un 28% de la primera muestra (5 casos) y un 30% de la segunda muestra (7 casos) superó dicho punto de corte, siendo estos individuos clasificados como con sospecha de simulación y, por lo tanto, falsos positivos. La especificidad de la medida en ambas muestras fue del 72% y el 70%, respectivamente. En la tabla 3.59 se pueden ver las puntuaciones medias de ambos grupos. Adicionalmente, los autores seleccionaron muestras de adultos normales, sin diagnóstico psicopatológico o neuropsicológico, equiparadas en edad y sexo, cuyas puntuaciones en el SIMS también se pueden observar en la tabla.

Tabla 3.59
Descriptivos en la puntuación total del SIMS del grupo de pacientes con esquizofrenia en el estudio de Peters et al. (2012)

	n	M	Dt
Muestra 1 pacientes con esquizofrenia	18	13,72	8,15
Muestra 1 control sin trastorno	19	3,68	2,21
Muestra 2 pacientes con esquizofrenia	23	12,22	5,80
Muestra 2 control sin trastorno	24	3,75	3,69

Los autores calcularon la especificidad del SIMS usando otros puntos de corte alternativos, obteniendo los valores que aparecen en la tabla 3.60. El uso del punto de corte cercano a 24 (sugerido por Wisdom et al., 2010), obtiene tasas deseables de especificidad cercanas al 95%, tal y como puede observarse

Tabla 3.60

Especificidad de puntos de corte alternativos para la puntuación total del SIMS en la muestra global de pacientes con esquizofrenia (n=41) en el estudio de Peters et al. (2012)

Punto de corte	Especificidad
>16	70,7%
>17	73,2%
>18	75,6%
>19	80,5%
>20	83,0%
>21	90,2%
>22	95%
>25	97,6%

3.9.3. Estudio de prevalencia de exageración de síntomas en pacientes con trastornos neurológicos de origen psicológico (van Beilen, Griffioen, Gross y Leenders, 2009)

En este estudio los investigadores querían indagar en la presencia de exageración de síntomas en pacientes con trastornos neurológicos sin causa orgánica y considerados de origen psicológico. En concreto, emplearon tres muestras: 26 pacientes con diferentes trastornos motores o parálisis de origen psicológico (sin base anatómica o física, 28% varones, edad media de 47,15, Dt 14,4 años), 26 pacientes con diversos trastornos neurológicos del movimiento (42% varones, edad media de 49,4 años y Dt de 13,2) y una muestra control de 16 participantes equiparada en edad y sexo (28% varones, edad media de 50,9 y Dt de 8,1). Ninguno de los evaluados se encontraba implicado en procesos legales, así como tampoco se evidenciaban incentivos externos manifiestos, que pudieran hacer sospechar la presencia de simulación. Dentro de las pruebas empleadas en el estudio, se incluyó el Amsterdam Short-Term Memory test (ASTM; Schmand y Lindeboom, 2005) y el SIMS, empleando como punto de corte de este último la obtención de puntuaciones en Total mayores de 16. Del conjunto de pacientes evaluados, obtuvieron puntuaciones por encima del punto de corte del ASTM, el 37,5% de los pacientes con trastornos motores de origen psicológico, el 23,1% de los pacientes con trastornos neurológicos del movimiento y un 6% del grupo control. En el caso del SIMS, obtuvieron puntuaciones totales superiores a 16, siendo clasificados como sospecha de exageración de síntomas, un 23% de los pacientes con trastornos motores de origen psicológico, un 3,8% de los pacientes con trastornos neurológicos del movimiento y un 0% del grupo control. En la tabla 3.61 se incluyen los estadísticos descriptivos de la puntuación total del SIMS en cada uno de los grupos considerados.

Tabla 3.61
Puntuaciones medias y desviaciones típicas en la puntuación total de los grupos de trastornos neurológicos con y sin origen psicológico en el estudio de van Beilen et al. (2009)

Escala del SIMS	Muestra de adultos normales respondiendo honestamente ^a		Muestra de pacientes con diversos trastornos neurológicos del movimiento ^b		Muestra de pacientes con diversos trastornos del movimiento de origen psicológico ^c	
	M	Dt	M	Dt	M	Dt
Total	4,58	2,27	7,8	3,9	11,48	7,44

Nota. n^a= 18; n^b= 26; n^c= 26.

3.9.4. Estudio de prevalencia de exageración de síntomas en pacientes con epilepsia de origen psicológico (Benge et al., 2012)

Benge et al. (2012) analizaron los datos de 120 casos clínicos que habían sido sometidos a evaluación en un hospital en los últimos años y dentro de cuyo protocolo se había aplicado el SIMS. Estos 120 casos habían sido clasificados como con epilepsia (n=29) o epilepsia de origen psicológico (n=91), de acuerdo a diversas pruebas.

En la tabla 3.62 se pueden ver los resultados en el SIMS y en cada de sus escalas en ambos grupos. Como se puede advertir existen diferencias significativas y relevantes tanto en el Total como en las escalas Deterioro neurológico, Trastornos afectivos y Trastornos Amnésicos. Por el contrario, no existen diferencias en Psicosis y Baja inteligencia.

Tabla 3.62
Descriptivos en el SIMS del grupo de epilepsia y epilepsia de origen psicológico. Estudio independiente de Benge et al. (2012)

Escala del SIMS	Epilepsia ^a		Epilepsia de origen psicológico ^b		t	d
	M	Dt	M	Dt		
Total	14,41	7,16	22,42	10,57	-3,81***	0,82
Psicosis	1,31	1,98	1,95	2,6	-1,2	0,26
Deterioro neurológico	3,03	2,38	6,02	3,24	-4,58***	0,99
Trastornos amnésicos	2,9	2,57	5,43	3,8	-3,34***	0,72
Baja inteligencia	2,38	1,52	2,16	1,73	0,6	0,13
Trastornos afectivos	4,79	2,37	6,86	2,2	-4,31***	0,93

Nota. n^a= 29; n^b= 91 *** p < 0,001.

A partir del estudio de regresión logística y curvas COR (véase resultados en la tabla 3.63) los autores determinaron que las escalas más útiles en la discriminación de la epilepsia de origen psicológico eran la puntuación Total, Deterioro neurológico y Trastornos afectivos. En las tablas 3.64 a 3.66 se indican los valores de sensibilidad, especificidad y poder predictivo positivo y negativo para

cada una de las escalas. Los autores sugirieron el uso de un punto de corte de 16 en la puntuación Total, de 3 en Deterioro neurológico y de 5 en Trastornos afectivos.

Tabla 3.63
Área bajo la curva de Total, Deterioro neurológico y Trastornos afectivos del SIMS para diferencia entre epilepsia y epilepsia de origen psicológico. Estudio independiente de Benge et al.(2012)

	Área bajo la curva	ET	P
Total SIMS	0,77	0,05	<0,001
Deterioro neurológico	0,77	0,06	<0,001
Trastornos afectivos	0,74	0,05	<0,001

Tabla 3.64
Sensibilidad, especificidad y poder predictivo positivo y negativo de la puntuación Total del SIMS para diferencia entre epilepsia y epilepsia de origen psicológico (Benge et al., 2012)

Punto de corte	Sensibilidad	Especificidad	Prevalencia 0,75		Prevalencia 0,50		Prevalencia 0,25		Prevalencia 0,10	
			PPP	PPN	PPP	PPN	PPP	PPN	PPP	PPN
10	0,90	0,35	0,8	0,54	0,58	0,78	0,31	0,91	0,13	0,97
11	0,85	0,41	0,81	0,47	0,59	0,73	0,32	0,89	0,14	0,96
12	0,85	0,41	0,81	0,47	0,59	0,73	0,32	0,89	0,14	0,96
13	0,78	0,48	0,82	0,42	0,6	0,69	0,33	0,87	0,14	0,95
14	0,76	0,55	0,84	0,43	0,63	0,7	0,36	0,87	0,16	0,95
15	0,73	0,62	0,85	0,43	0,66	0,69	0,39	0,87	0,18	0,95
16	0,71	0,69	0,87	0,45	0,7	0,71	0,43	0,88	0,2	0,96
17	0,68	0,72	0,88	0,43	0,71	0,69	0,45	0,87	0,22	0,95
18	0,63	0,79	0,9	0,41	0,75	0,68	0,5	0,86	0,25	0,95
19	0,59	0,83	0,91	0,4	0,78	0,67	0,53	0,86	0,28	0,95
20	0,52	0,86	0,92	0,37	0,79	0,64	0,55	0,84	0,29	0,94
21	0,48	0,86	0,91	0,36	0,78	0,63	0,54	0,83	0,28	0,94
22	0,43	0,86	0,9	0,33	0,76	0,6	0,51	0,82	0,26	0,93
23	0,39	0,9	0,92	0,33	0,79	0,59	0,55	0,81	0,29	0,93
24	0,34	0,9	0,91	0,31	0,77	0,58	0,52	0,8	0,27	0,92
25	0,32	0,93	0,93	0,31	0,82	0,58	0,61	0,8	0,34	0,92
26	0,26	0,93	0,92	0,3	0,79	0,56	0,56	0,79	0,3	0,92
27	0,23	0,97	0,95	0,3	0,87	0,56	0,69	0,79	0,43	0,92
28	0,22	0,97	0,95	0,29	0,87	0,55	0,68	0,79	0,42	0,92
29	0,2	0,97	0,95	0,29	0,85	0,55	0,66	0,78	0,39	0,92
30	0,18	0,97	0,94	0,28	0,84	0,54	0,63	0,78	0,37	0,91
35	0,13	0,97	0,92	0,27	0,8	0,53	0,56	0,77	0,3	0,91
40	0,09	1	1	0,27	1	0,52	1	0,77	1	0,91

Tabla 3.65

Sensibilidad, especificidad y poder predictivo positivo y negativo de la puntuación Deterioro neurológico del SIMS para diferencia entre epilepsia y epilepsia de origen psicológico (Benge et al., 2012)

Punto de corte	Sensibilidad	Especificidad	Prevalencia 0,75		Prevalencia 0,50		Prevalencia 0,25		Prevalencia 0,10	
			PPP	PPN	PPP	PPN	PPP	PPN	PPP	PPN
1	0,97	0,28	0,80	0,74	0,57	0,89	0,31	0,96	0,13	0,99
2	0,86	0,52	0,84	0,55	0,64	0,78	0,37	0,92	0,16	0,97
3	0,75	0,62	0,86	0,45	0,66	0,71	0,40	0,88	0,18	0,96
4	0,64	0,76	0,89	0,41	0,73	0,68	0,47	0,86	0,23	0,95
5	0,55	0,83	0,91	0,38	0,76	0,65	0,52	0,85	0,26	0,94
6	0,40	0,93	0,95	0,34	0,85	0,61	0,66	0,82	0,39	0,93
7	0,28	0,93	0,92	0,30	0,80	0,56	0,57	0,79	0,31	0,92
8	0,23	0,97	0,95	0,30	0,87	0,56	0,69	0,79	0,43	0,92
9	0,14	1,00	1,00	0,28	1,00	0,54	1,00	0,78	1,00	0,91
10	0,12	1,00	1,00	0,27	1,00	0,53	1,00	0,77	1,00	0,91

Tabla 3.66

Sensibilidad, especificidad y poder predictivo positivo y negativo de la puntuación Trastornos afectivos del SIMS para diferencia entre epilepsia y epilepsia de origen psicológico (Benge et al., 2012)

Punto de corte	Sensibilidad	Especificidad	Prevalencia 0,75		Prevalencia 0,50		Prevalencia 0,25		Prevalencia 0,10	
			PPP	PPN	PPP	PPN	PPP	PPN	PPP	PPN
1	1,00	0,03	0,76	1,00	0,51	1,00	0,26	1,00	0,10	1,00
2	0,97	0,17	0,78	0,63	0,54	0,84	0,28	0,94	0,11	0,98
3	0,93	0,38	0,82	0,66	0,60	0,85	0,33	0,95	0,14	0,98
4	0,82	0,55	0,85	0,51	0,65	0,76	0,38	0,90	0,17	0,97
5	0,73	0,62	0,85	0,43	0,66	0,69	0,39	0,87	0,18	0,95
6	0,58	0,72	0,86	0,37	0,68	0,63	0,41	0,84	0,19	0,94
7	0,42	0,76	0,84	0,30	0,63	0,57	0,37	0,80	0,16	0,92
8	0,26	0,97	0,96	0,30	0,89	0,57	0,72	0,80	0,46	0,92
9	0,12	1,00	1,00	0,27	1,00	0,53	1,00	0,77	1,00	0,91
10	0,02	1,00	1,00	0,25	1,00	0,51	1,00	0,75	1,00	0,90

3.9.5. Estudio de Geraerts, Jelicic y Merckelbach (2006) sobre abuso sexual y simulación de síntomas

Geraerts et al. (2006) estudiaron en una muestra de personas que habían sufrido abuso sexual si había diferencias en sus puntuaciones en las pruebas de exageración de síntomas, en virtud de si sus recuerdos de abuso habían estado siempre presentes o, por el contrario, habían sido olvidados durante un determinado periodo de tiempo para emerger posteriormente, como ocurre frecuentemente tras algunas terapias. Para el estudio, se seleccionó una muestra de 268 casos, de

los cuales 119 habían sufrido abuso y guardaban recuerdo del mismo desde que ocurrió, 66 sufrieron abuso y habían eliminado su recuerdo durante un amplio periodo de tiempo, para volver a recordarlo posteriormente y 83 era un grupo control, que no habían sufrido ningún tipo de abuso y no contaban con historial psicopatológico.

En la tabla 3.67 pueden observarse los resultados obtenidos, que muestran escasas diferencias entre ambos grupos de abuso sexual (recuerdo continuo y recuerdo discontinuo), a la par que puntuaciones significativamente superiores en estos grupos frente al grupo control. Estas diferencias estaban vinculadas, particularmente, a las escalas de Deterioro cognitivo, Problemas de memoria y Trastornos afectivos, que fueron las escalas del SIMS en las que hubo diferencias significativas entre el grupo de abuso sexual (recuerdo continuo y recuerdo recuperado) y el grupo control.

Tabla 3.67
Puntuaciones medias y desviaciones típicas en la puntuación total en el SIMS de los grupos de abuso sexual y de respuesta sincera en el estudio de Geraerts et al., 2006

	Abuso sexual Recuerdo continuo (n=119)	Abuso sexual Recuerdo recuperado (n=66)	Control (n=83)
Edad	42,3 (13,5)	41,7 (12,6)	41,3 (12,1)
Puntuación total en el SIMS	10,4 (6,6)	10,3 (7,8)	4,6 (3,8)
Rango de puntuación total en el SIMS	1-37	1-35	1-29

3.10. Visión general de las investigaciones realizadas con el SIMS

A partir de la revisión realizada de la literatura científica se han encontrado 37 investigaciones que han sido descritas previamente. Tal y como se puede advertir en la tabla 3.68, la mayoría de estas investigaciones han sido con diseños de simulación análoga (21 investigaciones, un 57%), siendo menor el número de estudios en contextos reales de evaluación con diseños de grupos diagnósticos o diseños autosuficientes (11 investigaciones, un 30%) o los estudios con muestras exclusivamente clínicas (5 investigaciones, un 14%).

En relación a las investigaciones con diseños de simulación análoga, la mayoría han empleado universitarios (13 investigaciones), contando en menor medida con adultos (6 estudios) o adolescentes (2 investigaciones). Ha de destacarse que 14 de las 21 investigaciones (un 66%) no incluían muestras clínicas de referencia con la que comparar los resultados de los simuladores análogos. Solo 7, apenas un tercio del total, disponían de estas muestras. En el resto de los estudios,

se comparaban las respuestas de los simuladores análogos con las dadas por universitarios o adultos normales que respondían honestamente, estableciendo así la capacidad discriminativa del test. Esto supone una seria limitación ya que el grupo de comparación relevante debe ser una muestra clínica para valorar la distinción entre quienes realmente tienen trastornos y quienes los simulan, no entre quienes no lo tienen (universitarios o adultos normales) y quienes los simulan.

De las 11 investigaciones realizadas con el SIMS con diseños de grupos conocidos o diseños autosuficientes, 6 han sido realizadas en contextos forenses civiles (solicitudes de indemnizaciones, compensaciones o discapacidades) y 5 en contextos forenses penales. En la mitad de los estudios (5 de las 11 investigaciones) se empleó el SIRS como criterio diagnóstico para la asignación de los grupos, mientras que en los otros casos se empleó el criterio de expertos (4 investigaciones), los criterios de Slick et al. (1999) y Bianchini et al. (2005) (1 investigación) o los resultados en otra prueba (1 investigación). Es llamativo que, aunque la literatura científica recalca la recomendación de los criterios diagnósticos de Slick et al. (1999) y Bianchini et al. (2005), en la práctica suelen emplearse minoritariamente a la hora de realizar investigaciones para identificar a los simuladores, utilizándose otras pruebas como criterio. Este hecho puede guardar cierta concordancia con la crítica realizada por Rogers, Bender, et al. (2011a), que señalaba el escaso y limitado uso empírico de los criterios de Slick y colaboradores y puede estar relacionado con su difícil implementación de forma operativa de cara a la realización de investigaciones.

Del total de investigaciones, la mayoría ha estudiado la simulación de trastornos mentales (28 de las 37 investigaciones), siendo menor el número de investigaciones en que se ha abordado también la simulación de deterioro cognitivo (10 de 37) o de síntomas neurológicos o somáticos (7 de 37). Solo una investigación se ha centrado en aspectos de simulación de dolor (en concreto, sobre esguince cervical), si bien era un diseño de simulación análoga, con universitarios y sin muestras clínicas de comparación.

Tabla 3.68*Características de las investigaciones realizadas sobre el SIMS en la literatura científica*

Investigación	Tipo de estudio	Tipo de muestral	¿Muestra clínica de referencia?	¿Simulación de trastornos mentales?	¿Simulación de deterioro cognitivo?	¿Simulación de síntomas físicos, somáticos o neurológicos?
Smith y Burger, 1997	Simulación análoga	Universitarios	--	Sí	Sí	Sí
Rogers y Cruise, 1998	Simulación análoga	Universitarios	--	Sí	--	--
Edens et al., 1999	Simulación análoga	Universitarios	--	Sí	--	--
Rogers et al., 2005	Simulación análoga	Universitarios	--	Sí	--	--
Theron et al., 2001	Simulación análoga	Universitarios	--	Sí	--	--
Jelicic et al., 2006	Simulación análoga	Universitarios	--	Sí	--	--
Jelicic, Peters et al., 2007	Simulación análoga	Universitarios	--	Sí	--	--
Jelicic et al., 2013	Simulación análoga	Universitarios	--	Sí	--	--
Dandachi et al., 2013	Simulación análoga	Universitarios	--	Sí	--	--
Jelicic et al., 2007	Simulación análoga	Universitarios	--	--	Sí	--
Jelicic et al., 2011	Simulación análoga	Universitarios	--	--	Sí	--
Merckelbach y Smith, 2003	Simulación análoga	Universitarios	Sí	--	Sí	--
Cima et al., 2003	Simulación análoga	Universitarios	Sí	Sí	Sí	Sí
Rogers et al., 1996	Simulación análoga	Adolescentes	Sí	Sí	--	--
Dearth et al., 2007	Simulación análoga	Adolescentes	Sí	Sí	--	--
Efendov et al., 2007	Simulación análoga	Adultos	Sí	Sí	--	--
Graue et al., 2007	Simulación análoga	Adultos	Sí	--	Sí	--
Buijs, 2010	Simulación análoga	Adultos	Sí	--	Sí	--
Merten et al, 2008	Simulación análoga	Adultos	--	--	--	Sí
Giger et al., 2010	Simulación análoga	Adultos	--	Sí	--	--
Merten et al., 2010	Simulación análoga	Adultos	--	Sí	--	--
Alwes et al., 2006	Grupos diagnósticos	Adultos en litigio civil	Sí	Sí	--	--
Clegg et al. 2009	Grupos diagnósticos	Adultos en litigio civil	Sí	Sí	--	--
Wisdom et al., 2010	Grupos diagnósticos	Adultos en litigio civil	Sí	Sí	Sí	Sí
Merten et al., 2007	Grupos diagnósticos	Adultos en litigio civil	Sí	--	Sí	--
González Ordi et al., 2010	Grupos diagnósticos	Adultos en litigio civil	Sí	Sí	--	Sí
Blasco y Pallardo, 2013	Grupos diagnósticos	Adultos en litigio civil	Sí	Sí	--	--
Edens et al., 2007	Grupos diagnósticos	Adultos en contexto penal	Sí	Sí	--	--

Tabla 3.68 (continuación)

Investigación	Tipo de estudio	Tipo de muestra	¿Muestra clínica de referencia?	¿Simulación de trastornos mentales?	¿Simulación de deterioro cognitivo?	¿Simulación de síntomas físicos, somáticos o neurológicos?
Laffoon, 2009	Grupos diagnósticos	Adultos en contexto penal	Sí	Sí	--	--
Lewis et al., 2002	Grupos diagnósticos	Adultos en contexto penal	Sí	Sí	--	--
Heinze y Purish, 2001	Grupos diagnósticos	Adultos en contexto penal	--	Sí	--	--
Vitacco et al., 2006	Grupos diagnósticos	Adultos en contexto penal	Sí	Sí	--	--
Dandachi et al., 2011	Muestras clínicas	Adultos pacientes clínicos	--	Sí	Sí	--
Peters et al., 2013	Muestras clínicas	Adultos pacientes clínicos	--	Sí	--	--
Van Beilen et al., 2009	Muestras clínicas	Adultos pacientes clínicos	--	--	--	Sí
Benge et al., 2012	Muestras clínicas	Adultos pacientes clínicos	--	--	--	Sí
Geraerts et al., 2006	Muestras clínicas	Adultos pacientes clínicos	--	Sí	--	--

En la tabla 3.69 se describen específicamente los análisis realizados y de los que se informan en cada una de las investigaciones. La mayoría de los estudios proporciona datos de sensibilidad y especificidad, empleando como punto de corte 16 (en un 70% de los casos) o 14 (en un 19% de los casos), si bien son escasos los estudios que informan también del poder predictivo positivo y negativo (solo 15 investigaciones) o que presentan los valores de sensibilidad y especificidad para otros puntos de corte alternativos (11 investigaciones, apenas un tercio).

Respecto a estos datos es interesante señalar varios aspectos. En primer lugar, el punto de corte originalmente propuesto por los autores de la prueba para la puntuación total, 14 (Smith y Burger, 1997; Widows y Smith, 2005), apenas es utilizado en la actualidad, mientras que el punto de corte alternativo de 16 se ha impuesto, sugiriéndose también puntos de corte más elevados. Asimismo, resulta de interés el hecho de que existiendo diversos puntos de corte recomendados, la mayoría de las investigaciones no ofrezca los valores de sensibilidad y especificidad para estos, más aún, cuando varias de esas investigaciones señalan la limitada especificidad del SIMS y, precisamente, la variación en los puntos de corte podría permitir un mayor equilibrio entre la sensibilidad y la especificidad o un aumento de esta última. A este respecto, el punto de corte de 16 ha mostrado sistemáticamente una moderada especificidad con valores alrededor de 0,70 en diversas muestras clínicas (Dandachi-FitzGerald et al., 2011; Edens et al., 2007; Lewis, Simcox y Berry, 2002; Peters, Jelcic, et al., 2013; Wisdom et al., 2010).

En segundo lugar, destaca la escasa atención prestada a las escalas del SIMS, a pesar de que cada una de ellas está específicamente diseñada para detectar la simulación en un área de sintomatología concreta. Sólo algo más de la mitad de las investigaciones (22 de 37) informa de los estadísticos descriptivos de dichas escalas y, de ellas, solo 13 estudian la sensibilidad y especificidad de sus puntos de corte. Si bien el punto de corte de la puntuación total se ha visto modificado desde los autores originales a la práctica actual, en el caso de las escalas, no se barajan de forma consistente otros puntos de corte alternativos, algo que, probablemente, fuera necesario.

Tabla 3.69

Estudios estadísticos empleados en las investigaciones realizadas sobre el SIMS

Investigación	¿Qué punto corte para Total?	¿Datos de sensibilidad y especificidad?	¿Datos de PPP y PPN?	¿Estudio de puntos de corte alternativos para Total?	¿Descriptivos de las escalas?	¿Estudio de los puntos de corte de las escalas?
Smith y Burger, 1997	>14	Sí	--	--	Sí	Sí
Rogers y Cruise, 1998	>16	--	--	--	Sí	--
Edens et al., 1999	>14	Sí	--	--	Sí	Sí
Rogers et al. 2005	--	Sí	Sí	Sí	Sí	Sí
Theron et al., 2001	>20	Sí	--	Sí	--	Sí
Jelicic et al., 2006	>16	Sí	--	--	--	--
Jelicic et al., 2007	>14	Sí	--	--	--	--
Jelicic et al., 2013	>16	Sí	--	--	--	--
Dandachi et al., 2013	>16	Sí	--	--	--	--
Jelicic et al., 2007	>16	Sí	--	--	--	--
Jelicic et al., 2011	>16	Sí	--	--	--	--
Merckelbach y Smith, 2003	>16	Sí	Sí	--	--	--
Cima et al., 2003	>16	Sí	Sí	--	Sí	--
Rogers et al., 1996	>16	--	Sí	Sí	--	Sí
Dearth, 2007	>16	Sí	Sí	Sí	Sí	Sí
Efendov, 2007	--	--	--	--	Sí	--
Graue et al., 2007	>16	Sí	--	--	Sí	Sí
Buijs, 2011	--	--	--	--	Sí	--
Merten et al, 2008	>16	Sí	--	--	Sí	--
Giger et al., 2010	>16	Sí	--	--	Sí	Sí
Merten et al., 2010	>16	Sí	--	--	Sí	--
Alwes et al., 2006	>16	Sí	Sí	Sí	Sí	--
Clegg et al. 2009	>19	Sí	Sí	Sí	Sí	Sí
Wisdom et al., 2010	>24	Sí	Sí	Sí	--	--
Merten et al., 2007	>16	--	--	--	Sí	--
González Ordi et al., 2010	>16	Sí	Sí	Sí	Sí	--

Tabla 3.69 (continuación)

Investigación	¿Qué punto corte para Total?	¿Datos de sensibilidad y especificidad?	¿Datos de PPP y PPN?	¿Estudio de puntos de corte alternativos para Total?	¿Descriptivos de las escalas?	¿Estudio de los puntos de corte de las escalas?
Blasco y Pallardo, 2013	>16	--	--	--	Sí	Sí
Edens et al., 2007	>14	Sí	--	--	--	--
Laffoon, 2009	>14	Sí	Sí	--	--	--
Lewis et al., 2002	>16	Sí	Sí	--	Sí	--
Heinze y Purish, 2001	>13	Sí	--	Sí	Sí	Sí
Vitacco et al., 2006	>14	Sí	Sí	--	Sí	Sí
Dandachi et al., 2011	>16	Sí	-	--	--	--
Peters et al., 2013	>16	Sí	-	Sí	--	--
Van Beilen et al., 2009	>16	Sí	-	--	--	--
Benge et al., 2012	>16	Sí	Sí	Sí	Sí	Sí
Geraerts et al., 2006	>16	-	-	--	Sí	--

Nota. PPP: Poder predictivo positivo, PPN: Poder predictivo negativo.

De aquellas investigaciones que informaban de los descriptivos del SIMS, se procedió a calcular el tamaño del efecto de las diferencias entre la condición de respuesta honesta y respuesta simulada. Solo en 19 estudios pudo realizarse este cálculo y los resultados aparecen agrupados en las tablas de la 3.70 a la 3.72, en función del diseño empleado (análogos sin muestra clínica de referencia, análogos con muestra clínica de referencia, diseño de grupos diagnósticos o autosuficientes en contexto civil o penal). Para el cálculo del tamaño del efecto, se calculó la g de Hedges como estimador insesgado de la d de Cohen ya que resulta más apropiada cuando se realiza una revisión sistemática de distintos estudios con tamaños muestrales diversos y, en algunos casos, pequeños (Chalmers, Hedges y Cooper, 2002; Hedges y Olkin, 1985; Turner y Bernard, 2006).

Del análisis de estos estudios, puede observarse que los mayores valores de la d de Cohen aparecen, lógicamente, en los diseños análogos que comparan simuladores análogos con participantes normales respondiendo honestamente (promedio d de Cohen = 3,17, tabla 3.70). Cuando el grupo de comparación es clínico, los valores disminuyen sensiblemente, tanto en los diseños análogos con muestra clínica de referencia (promedio d de Cohen = 1,56, tabla 3.71), como en los diseños de grupos diagnósticos o autosuficientes (promedio d de Cohen = 1,94, tabla 3.72). Otro factor relevante que modula el valor del tamaño del efecto obtenido es que haya habido formación o entrenamiento de los simuladores. La introducción de advertencias sobre la credibilidad y, particularmente, la información sobre la existencia de escalas de validez destinadas a detectar la simulación, disminuye sensiblemente los tamaños del efecto obtenidos, tal y como se puede observar en la tabla 3.70.

Asimismo, en aquellos estudios que se ofrecen datos sobre las escalas para cada condición de simulación (psicosis, deterioro neurológico...), se suelen obtener valores mayores de la d de Cohen en la escala específica sobre los síntomas alegados, lo cual es una evidencia del ajuste de las escalas a cada área de simulación, así como de su utilidad discriminativa. La puntuación Total es, además, sistemáticamente la que mayores valores obtiene para la d de Cohen en los diversos estudios, lo que refuerza su uso como puntuación global de referencia para la evaluación con el SIMS.

Si se analizan los resultados de la tabla con diseños de grupos diagnósticos o autosuficientes (tabla 3.72), se advierte que los mayores tamaños del efecto aparecen en el contexto penal, donde se suele alegar psicopatología más grave, más global y más incapacitante, frente a los contextos civiles, que presentan valores ligeramente inferiores. Esto es coherente con lo encontrado en otras investigaciones al comparar simulación en contexto penal y civil (Wygant et al., 2007, 2010). En el contexto civil, además, se han obtenido mayores valores de la d de Cohen cuando los participantes

alegaban problemas de tipo psicopatológico que cuando eran de índole cognitivo. Si bien el SIMS obtiene valores satisfactorios en ambos casos, es mayor su capacidad discriminativa cuando el área de simulación es psicopatológica que cuando es cognitiva. Las escalas que mayores d de Cohen obtienen son diferentes también según los síntomas alegados, obteniendo mayores valores las escalas de Deterioro neurológico o Trastornos amnésicos cuando el motivo principal alegado son síntomas cognitivos o neurológicos.

A efectos meramente ilustrativos, en la figura 3.3 se han representado las puntuaciones medias en el total del SIMS según el tipo de muestra de las distintas investigaciones mencionadas en este capítulo (se han considerado solo los estudios realizados con mayores de 18 años). Así, se pueden observar cuáles son las puntuaciones medias que se han obtenido con personas de la población normal, respondiendo sinceramente (población general), en clínicos en contexto no litigante, en clínicos en contexto litigante, en simuladores en contexto litigante y en simuladores análogos. Los datos muestran un progresivo aumento de las puntuaciones en los diferentes grupos, siendo las puntuaciones más bajas en el grupo de población general con respuesta sincera, como cabría esperar y las más elevadas en simuladores análogos instruidos para alegar determinados síntomas o trastornos, tal y como también han observado sistemáticamente otras investigaciones en esta área (p.ej. Viglione et al., 2001).

Tabla 3.70

Tamaño del efecto (d Cohen) de las investigaciones que han empleado diseños de simulación análoga sin grupo clínico de referencia: comparación de simuladores análogos frente a respuesta honesta normales.

Investigación	Síntomas simulados	n _{honesto}	n _{simulador}	Entrenamiento	d					
					Total	Ps	Dn	Am	Bi	Af
Malestar general (global)										
Smith y Burger,1997	Mal estado general	34	34	--	4,52	7,17	3,43	3,28	2,50	2,59
Psicosis										
Smith y Burger,1997	Psicosis	34	34	--	3,80	4,01	2,59	2,89	2,43	2,12
Edens et al., 1999	Psicosis	59	59	--	3,12	2,96	2,40	2,24	2,31	1,93
Jelicic et al., 2007	Psicosis	30	30	--	3,87	--	--	--	--	--
Jelicic et al., 2007	Psicosis	31	31	Síntomas	3,58	--	--	--	--	--
Dandachi et al., 2013	Psicosis	20	21	--	3,23	--	--	--	--	--
Deterioro cognitivo y neurológico										
Smith y Burger,1997	Deterioro neurológico	34	34	--	3,48	1,63	2,67	3,87	2,26	1,80
Edens et al., 1999	Deterioro cognitive	72	72	--	3,50	0,92	2,58	4,11	2,08	1,58
Jelicic et al., 2007	Daño cerebral tras accidente	30	30	--	3,02	--	--	--	--	--
Jelicic et al., 2007	Daño cerebral tras accidente	30	30	Síntomas y validez	2,62	--	--	--	--	--
Jelicic et al., 2011	Daño cerebral tras accidente	30	30	Síntomas	4,04	--	--	--	--	--
Jelicic et al., 2011	Daño cerebral tras accidente	30	30	Validez	3,10	--	--	--	--	--
Trastornos amnésicos										
Smith y Burger,1997	Amnesia	34	34	--	3,67	0,74	1,70	5,23	2,95	1,11
Giger et al., 2010	Amnesia tras crimen	20	20	--	3,11	1,41	2,34	4,40	1,32	1,78
""	Amnesia tras crimen	20	20	Credibilidad	1,74	0,77	1,37	2,43	0,72	0,95
Baja inteligencia										
Smith y Burger,1997	Baja inteligencia	34	34	--	2,81	1,32	1,84	2,22	5,66	1,03

Tabla 3.70 (continuación)

Investigación	Síntomas simulados	n _{honesto}	n _{simulador}	Entrenamiento	d					
					Total	Ps	Dn	Am	Bi	Af
Trastornos afectivos										
Smith y Burger, 1997	Depresión	34	34	--	2,42	0,81	1,82	1,35	0,93	4,77
Edens et al., 1999	Depresión	65	65	--	2,46	1,07	2,01	2,06	1,06	3,33
Rogers et al., 2005	Simulación baja laboral	16	17	--	1,41	0,67	1,68	0,84	0,55	0,68
Dandachi et al., 2013	TEPT	20	21	--	3,80	--	--	--	--	--
Otros										
Rogers et al., 2005	Facticio dependiente	16	14	--	0,94	0,55	0,94	0,61	-0,07	0,96
""	Facticio demandante	16	18	--	1,31	0,40	0,95	0,63	0,40	1,54

Nota. Todos los estudios dispusieron de muestra de universitarios, a excepción del estudio de Giger et al. (2010) que fueron adultos.

Tabla 3.71

Tamaño del efecto (d Cohen) de las investigaciones que han empleado diseños de simulación análoga con grupo clínico de referencia: comparación de simuladores análogos frente a respuesta honesta clínicos.

Investigación	Síntomas simulados	n _{clínico}	n _{simulador}	Total	Ps	Dn	Am	Bi	Af
Psicosis									
Cima et al., 2003	Psicosis	69	29	2,42	--	--	--	--	--
Merckelbach y Smith, 2003	Esquizofrenia	10	14	2,19	--	--	--	--	--
Deterioro cognitivo y neurológico									
Merckelbach y Smith, 2003	Trastorno neurológico	10	15	1,96	--	--	--	--	--
Cima et al., 2003	Deterioro neurológico	69	27	0,63	--	--	--	--	--
Trastornos amnésicos									
Buijs, 2011	TDAH	27	47	--	0,71	0,40	0,93	0,66	0,07
Merckelbach y Smith, 2003	Amnesia	10	28	1,80	--	--	--	--	--
Cima et al., 2003	Amnesia	69	21	1,08	--	--	--	--	--
Baja inteligencia									
Cima et al., 2003	Baja inteligencia	69	27	1,72	--	--	--	--	--
Trastornos afectivos									
Efendov, 2007	TEPT			1,25	0,74	0,84	1,68	0,63	1,42
Cima et al., 2003	Trastornos afectivos	69	31	1,50	--	--	--	--	--
Dearth, 2007	Trastornos afectivos	37	31	1,00	0,97	0,75	0,85	0,50	1,15

Nota. Todos los estudios dispusieron de muestra de universitarios en el grupo de simulación análoga, a excepción de los estudios de Dearth (2007), Efendov (2007) y Buijs (2011) que incluyeron adultos no estudiantes universitarios.

Tabla 3.72

Tamaño del efecto (d Cohen) de las investigaciones que han empleado diseños de grupos conocidos o autosuficientes: comparación de simuladores frente a clínicos en contextos reales.

Investigación	Síntomas simulados	N _{clínico}	n _{simulador}	Total	Ps	Dn	Am	Bi	Af
Contexto penal - Competencia para ser juzgado									
Lewis et al., 2002	Presos	31	24	2,90	2,81	2,46	1,95	1,06	1,69
Vitacco et al., 2006	Presos	79	21	3,05	2,65	2,42	2,32	2,28	1,50
Contexto civil - alegación psicopatología indemnización, discapacidad									
Alwes et al., 2006	Adultos litigantes	172	23	2,52	2,47	1,84	1,76	0,94	1,40
Clegg et al. 2009	Adultos litigantes	36	20	1,71	1,23	0,62	1,46	0,41	1,76
Wisdom et al., 2010	Adultos litigantes	11	17	1,45	--	--	--	--	--
González Ordi et al., 2010	Adultos litigantes	27	34	1,60	1,10	0,90	1,40	0,70	1,20
Blasco y Pallardo, 2013	Adultos litigants	12	14	2,48	0,79	3,79	2,23	0,59	1,89
Contexto civil - alegación problema cognitivo y solicitud compensación									
Merten et al., 2007	Adultos litigantes	120	39	1,20	0,58	1,23	1,03	0,50	0,76
Alwes et al., 2006	Adultos litigantes	178	75	1,10	0,71	0,79	1,22	0,40	0,61

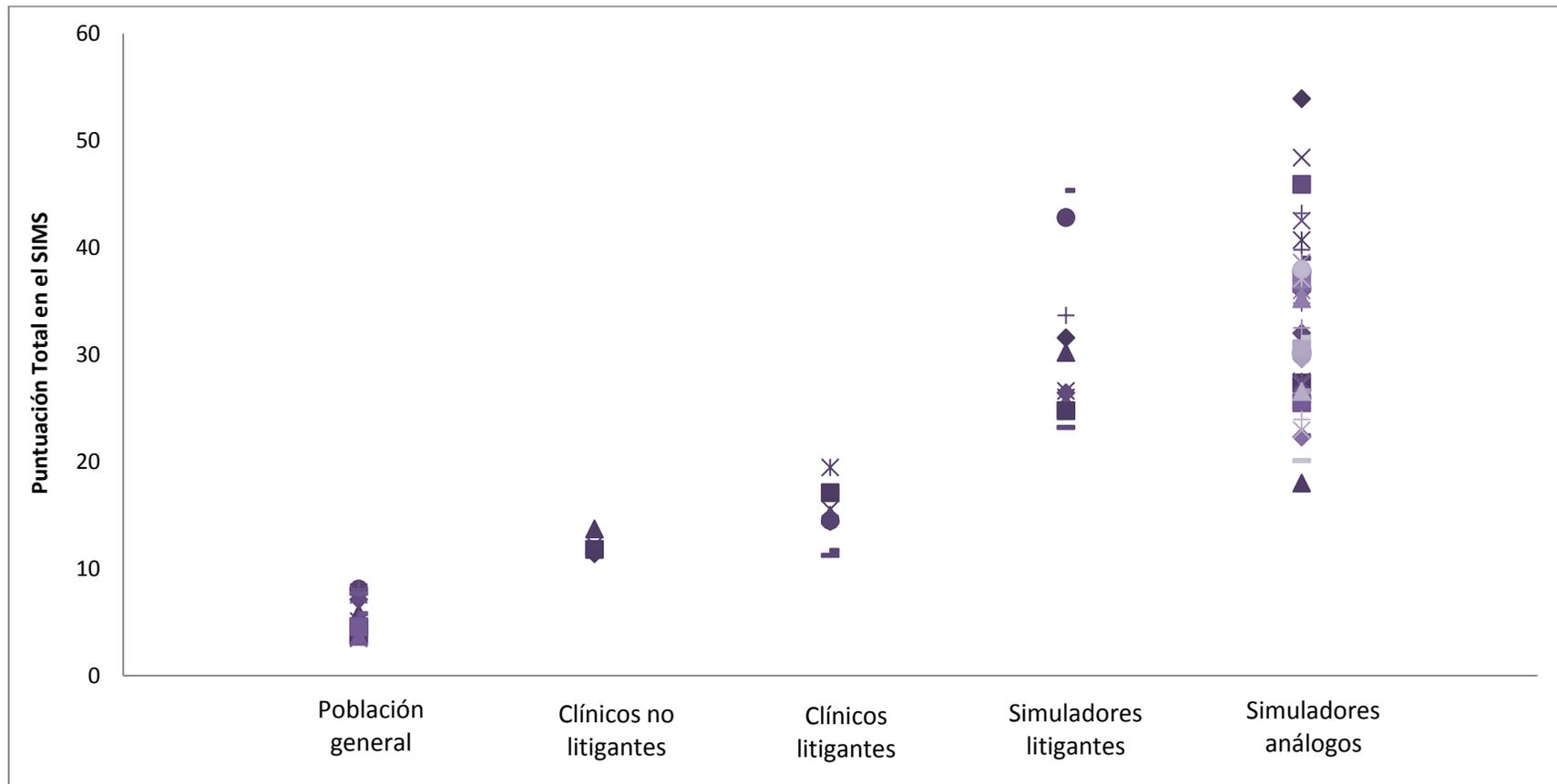


Figura 3.1. Puntuaciones medias en el total del SIMS por tipo de muestras en las distintas investigaciones existentes en la literatura científica

Población general: muestra normal, sin sospecha de patología, que responden a la prueba en condición de sinceridad; **Clínicos no litigantes:** muestra clínica, con diagnósticos de psicopatología, que responden a la prueba en condición de sinceridad y en un contexto sin incentivos externos para la simulación; **Clínicos litigantes:** muestra clínica que responde en un contexto de litigancia en el cual existen incentivos externos para la simulación pero que han sido clasificados como de respuesta sincera de acuerdo a un criterio externo al SIMS; **Simuladores litigantes:** muestra que responde en un contexto de litigancia en el cual existen incentivos externos para la simulación y que han sido clasificados como simuladores de acuerdo a un criterio externo al SIMS; **Simuladores análogos:** muestra instruida para simular trastornos específicos.

En la tabla 3.73 se detallan las medias ponderadas en cada una de las condiciones. Tal y como puede observarse, las investigaciones realizadas hasta el momento con el SIMS sitúan la media de la población general de respuesta sincera alrededor de los 6 puntos, la de las muestras clínicas respondiendo honestamente en 12 o 15 puntos, según si existe un contexto de litigancia o no y la de los grupos de simuladores alrededor de los 32 o 33 puntos.

Tabla 3.73
Puntuaciones medias en Total SIMS en diferentes grupos en las investigaciones encontradas en la literatura científica

Condición	n muestras	n participantes	M TOTAL SIMS ponderada
Población general	23	958	6,2
Clínicos no litigantes	4	234	12,0
Clínicos litigantes	8	545	15,2
Simuladores litigantes	9	256	31,4
Simuladores análogos	45	1463	32,7

Nota. Población general: muestra normal, sin sospecha de patología, que responden a la prueba en condición de sinceridad; Clínicos no litigantes: muestra clínica, con diagnósticos de psicopatología, que responden a la prueba en condición de sinceridad y en un contexto sin incentivos externos para la simulación; Clínicos litigantes: muestra clínica que responde en un contexto de litigancia en el cual existen incentivos externos para la simulación pero que han sido clasificados como de respuesta sincera de acuerdo a un criterio externo al SIMS; Simuladores litigantes: muestra que responde en un contexto de litigancia en el cual existen incentivos externos para la simulación y que han sido clasificados como simuladores de acuerdo a un criterio externo al SIMS; Simuladores análogos: muestra instruida para simular trastornos específicos.

Si desglosamos los resultados en los clínicos y simuladores litigantes por contexto (civil o penal, tabla 3.74), o por condición simulada en los simuladores análogos (tabla 3.53), se obtienen resultados interesantes. Así, se advierte que los simuladores litigantes en contexto penal, frecuentemente inmersos en evaluaciones de su competencia mental para ser juzgados y los simuladores análogos instruidos para simular trastornos más graves e incapacitantes, como son los trastornos psicóticos, son los que obtienen puntuaciones medias más altas en el total del SIMS. Por el contrario, los simuladores litigantes en contexto civil y los simuladores análogos de psicopatología afectiva o somática (trastornos afectivos, TEPT, deterioro neurológico o esguince cervical), frecuentemente más vinculados al contexto civil de indemnizaciones y bajas laborales, obtienen puntuaciones relativamente más bajas.

Tabla 3.74
Puntuaciones medias en Total SIMS según contexto civil o penal en muestras de simuladores y clínicos litigantes encontradas en la literatura científica

Condición	n muestras	n participantes	M ponderada TOTAL SIMS
Clínico litigante			
Contexto civil	6	435	15,3
Contexto penal	2	110	14,8
Simulador litigante			
Contexto civil	6	154	26,9
Contexto penal	3	102	38,2

Tabla 3.75
Puntuaciones medias en Total SIMS según trastorno instruido en diferentes grupos de simulaciones análogas encontradas en la literatura científica

Condición simulada	n muestras	n participantes	M ponderada
			TOTAL SIMS
Psicosis	12	310	39,2
Deterioro neurológico	8	226	29,0
Trastornos amnésicos	3	83	33,1
Baja inteligencia	4	158	37,3
Trastornos afectivos	4	160	30,4
TEPT	7	163	30,0
Amnesia tras crimen	2	40	30,1
Esguince cervical	2	32	25,9

Al analizar los resultados en función de las escalas del SIMS, tabla 3.76., se puede observar el mismo gradiente de aumento de puntuaciones conforme avanzamos de población general respondiendo de forma honesta, hasta los simuladores litigantes o análogos, pasando por los clínicos no litigantes o los clínicos litigantes.

Tabla 3.76
Puntuaciones medias en las escalas del SIMS en diferentes condiciones de respuesta encontradas en la literatura científica

Condición	n muestras	n participantes	M ponderada				
			Ps	Dn	Am	Bi	Af
Población general	23	958	0,54	0,94	0,89	1,6	3,73
Clínicos no litigantes	1	183	0,92	2,02	2,89	1,36	4,64
Clínicos litigantes	8	545	0,80	3,41	2,56	2,39	6,06
Simuladores litigantes	7	256	4,71	7,24	7,72	4,67	8,47
Simuladores análogos	32	1139	4,77	6,41	7,87	5,07	8,20

Nota. Población general: muestra normal, sin sospecha de patología, que responden a la prueba en condición de sinceridad; Clínicos no litigantes: muestra clínica, con diagnósticos de psicopatología, que responden a la prueba en condición de sinceridad y en un contexto sin incentivos externos para la simulación; Clínicos litigantes: muestra clínica que responde en un contexto de litigancia en el cual existen incentivos externos para la simulación pero que han sido clasificados como de respuesta sincera de acuerdo a un criterio externo al SIMS; Simuladores litigantes: muestra que responde en un contexto de litigancia en el cual existen incentivos externos para la simulación y que han sido clasificados como simuladores de acuerdo a un criterio externo al SIMS; Simuladores análogos: muestra instruida para simular trastornos específicos. Ps: Psicosis; Dn: Deterioro neurológico; Am: Trastornos amnésicos; Bi: Baja inteligencia; Af: Trastornos afectivos.

El análisis por contexto de simulación en muestras litigantes (tabla 3.77) pone en evidencia que la escala Psicosis aparece sistemáticamente más elevada en el contexto penal que en el contexto civil, tanto con muestras clínicas como con simuladores. En particular, en el contexto penal, es la escala que más diferencias marca entre los clínicos y los simuladores, de forma coherente al objetivo de simulación en el contexto penal de aparentar ser incompetente mentalmente para ser juzgado.

Tabla 3.77

Puntuaciones medias en las escalas del SIMS según contexto civil o penal en muestras de simuladores y clínicos litigantes encontradas en la literatura científica

Condición y contexto	n muestras	n participantes	M ponderada				
			Ps	Dn	Am	Bi	Ta
Clínico litigante							
Contexto civil	6	435	0,46	3,65	2,59	2,30	6,33
Contexto penal	2	110	2,1	2,48	2,47	2,75	5,01
Simulador litigante							
Contexto civil	4	132	2,23	6,60	6,33	3,24	8,50
Contexto penal	3	102	7,91	8,07	9,53	6,52	8,43

El análisis de las puntuaciones medias en las escalas según la condición instruida para simular en los estudios análogos, tabla 3.76, muestra sistemáticamente que en cada condición se eleva de forma más destacada aquella escala del SIMS específicamente destinada a evaluar dicha área de simulación. Esto es un importante refrendo de las escalas del SIMS y un apoyo a su utilidad en la interpretación del SIMS y en la evaluación con la prueba.

Tabla 3.78

Puntuaciones medias en las escalas del SIMS según trastorno instruido en diferentes grupos de simulaciones análogas encontradas en la literatura científica

Condición simulada	n muestras	n participantes	M ponderada				
			Ps	Dn	Am	Bi	Ta
Psicosis	4	136	10,5	8,11	9,31	6,67	8,29
Deterioro neurológico	3	76	4,61	8,26	8,45	5,62	6,92
Trastornos amnésicos	3	83	2,04	5,76	11,9	6,97	6,32
Baja inteligencia	4	158	4,18	7,32	10,5	8,26	6,99
Trastornos afectivos	4	160	3,11	6,49	7,01	3,31	10,5
TEPT	7	163	3,09	6,07	7,52	3,65	9,65
Amnesia tras crimen	2	40	4,15	5,65	10,3	4,1	6,05
Esguince cervical	2	32	1,8	5,05	6,65	3,75	8,5

Si analizamos la sensibilidad y especificidad en aquellos estudios que específicamente han abordado esta cuestión con diseños de grupos conocidos o diseños autosuficientes en adultos, se obtienen los datos que aparecen en la tabla 3.79 para la puntuación total. A este respecto, el punto de corte de 14 obtiene una alta sensibilidad pero una moderada especificidad (entre el 0,47 y el 0,81) y conforme vamos aumentando en el punto de corte se disminuye la sensibilidad y va aumentando su especificidad. Dos autores han defendido, específicamente, un punto de corte más elevado para el SIMS, Clegg et al. (2009), que argumentaba el uso del punto de corte de 19 y Wisdom et al. (2010),

que recomendaba el punto de corte mayor de 23, en ambos casos con el claro objetivo de obtener valores de especificidad más satisfactorios, que redujeran el número de falsos positivos.

Tabla 3.79
Sensibilidad y especificidad de la puntuación total en el SIMS en los diferentes estudios de grupos diagnósticos conocidos o diseños autosuficientes en simulación de psicopatología o trastornos cognitivos en adultos

Punto de corte y estudio	Tipo de estudio	Área	Sensibilidad	Especificidad
>14				
Alwes, 2006	Contexto litigante civil	Psicopatología	0,96	0,54
Alwes, 2006	Contexto litigante civil	Cognitivo	0,80	0,47
Alwes, 2006	Contexto litigante civil	Psicopatología o cognitivo	0,81	0,57
Clegg et al., 2009	Contexto litigante civil	Psicopatología	1,00	0,37
Wisdom et al., 2010	Contexto litigante civil	Psicopatología o cognitivo	0,96	0,64
González-Ordi et al., 2010	Contexto litigante civil	Psicopatología	0,88	0,81
Vitacco et al., 2006	Contexto litigante penal	Psicopatología	1,00	0,65
>16				
Dandachi-FitzGerald et al., 2011	Muestras clínicas	Psicopatología	--	0,78
Peters et al., 2013	Muestras clínicas	Psicopatología	--	0,71
Clegg et al., 2009	Contexto litigante civil	Psicopatología	0,95	0,51
Alwes, 2006	Contexto litigante civil	Psicopatología	0,96	0,67
Alwes, 2006	Contexto litigante civil	Cognitivo	0,75	0,60
Alwes, 2006	Contexto litigante civil	Psicopatología o cognitivo	0,77	0,70
González-Ordi et al., 2010	Contexto litigante civil	Psicopatología	0,82	0,93
Edens et al., 2007	Contexto litigante penal	Psicopatología	0,85	0,40
Lewis et al., 2002	Contexto litigante penal	Psicopatología	1,00	0,61
>19				
Clegg et al., 2009	Contexto litigante civil	Psicopatología	0,90	0,69
>20				
Wisdom et al., 2010	Contexto litigante civil	Psicopatología o cognitivo	0,68	0,73
>22				
Wisdom et al., 2010	Contexto litigante civil	Psicopatología o cognitivo	0,55	0,91
>23				
Wisdom et al., 2010	Contexto litigante civil	Psicopatología o cognitivo	0,55	1,00
>24				
Wisdom et al., 2010	Contexto litigante civil	Psicopatología o cognitivo	0,50	1,00

Son muchos menos los estudios con grupos diagnósticos o autosuficientes en contextos litigantes que han presentado datos sobre la sensibilidad o especificidad de las distintas escalas del SIMS (tabla 3.80). En estos casos, apenas hay estudios que hayan considerado otros puntos de corte alternativos, a pesar de que, como se puede observar, estos puntos de corte pecan de una limitada especificidad, al igual que le ocurría al punto original para la puntuación total propuesto por los autores originales del SIMS.

Tabla 3.80
Sensibilidad y especificidad de la puntuación en las escalas del SIMS en los diferentes estudios de grupos diagnósticos conocidos o diseños autosuficientes en simulación de psicopatología o trastornos cognitivos en adultos

Escalas del SIMS	Clegg et al. (2009) Contexto litigante civil		Vitacco et al. (2007) Contexto litigante penal	
	Sensibilidad	Especificidad	Sensibilidad	Especificidad
Psicosis > 1	0,65	0,94	1,00	0,51
Deterioro neurológico > 2	0,80	0,40	0,95	0,66
Trastornos amnésicos > 2	1,00	0,54	0,90	0,67
Baja inteligencia > 2	0,70	0,63	0,90	0,57
Trastornos afectivos > 5	1,00	0,29	0,90	0,71

Finalmente, si se analizan en la literatura científica las correlaciones del SIMS con otras pruebas de validez de síntomas, se observan los resultados que aparecen en la tabla 3.81. En esta línea, pueden advertirse correlaciones moderadas del SIMS con las principales pruebas de referencia en el área como MMPI, PAI, SIRS o TOMM. Tal y como cabría esperar, son más altos los valores de correlación con las pruebas de autoinforme (MMPI, PAI o SIRS), que con las pruebas de rendimiento (TOMM, MVT, AMST), hecho frecuentemente observado en la literatura científica (Dandachi-FitzGerald et al., 2011; Haggerty, Frazier, Busch y Naugle, 2007; Lilienfeld,, Thames, y Watts,, 2013; Nelson et al., 2007; Whiteside et al., 2009). A este respecto y tal como se esperaba, la escala de Trastornos amnésicos es la que más correlación ha mostrado con estas pruebas de rendimiento, lo que refuerza la utilidad de esta escala para reflejar esta área de simulación específica (Leung, 2010).

Tabla 3.81
Correlaciones halladas en la literatura científica entre el SIMS y otras pruebas de evaluación

Test	Escala	Investigación	Tipo de muestra	Condición	n	Total	Ps	Dn	Am	Bi	Af
MMPI											
	F	Smith y Burger (1997)	Universitarios	Mezcla (control y simuladores)	238	0,84	0,84	0,78	0,53	0,53	0,63
		Cima et al., 2003	Pacientes forense criminal	Mezcla (clínico y simulador)	62	0,43					
		Heinze y Purisch (2001)	Presos	Simuladores	57	0,5		0,44			
		Carter et al., 2009	Pacientes forense civil	Mezcla (clínico y simulador)		0,8					
		Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	67	0,67		0,57	0,52		0,38
	F-K	Smith y Burger (1997)	Universitarios	Mezcla (control y simuladores)	238	0,81	0,83	0,76	0,49	0,47	0,65
		Heinze y Purisch (2001)	Presos	Simuladores	57	0,48		0,46			
	Fb	Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	67	0,57		0,39	0,38		0,43
	Fp	Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	67	0,53		0,38	0,54		0,05
	FBS	Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	61	0,44		0,4	0,29		0,36
	Md	Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	67	0,5		0,39	0,42		0,42
	MMDS	Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	67	0,49		0,4	0,4		0,41
	K	Cima et al., 2003	Pacientes forense criminal	Mezcla (clínico y simulador)	62	-0,2					
	L	Cima et al., 2003	Pacientes forense criminal	Mezcla (clínico y simulador)	62	0,03					
PAI											
	IMN	Edens et. al. (2007)	Presos	Mezcla (clínico y simulador)	116	0,8					
		West 2004	Solicitantes incapacidad	Mezcla (clínico y simulador)	20	0,94					
		Laffoon, 2009	Presos	Mezcla (clínico y simulador)	100	0,59					
		Tscheuschner, 2011	Presos	Mezcla (clínico y simulador)	101	0,67					
	SIM	Edens et. al. (2007)	Presos	Mezcla (clínico y simulador)	116	0,68					
		Laffoon, 2009	Presos	Mezcla (clínico y simulador)	100	0,29					
		Tscheuschner, 2011	Presos	Mezcla (clínico y simulador)	101	0,23					
	FDR	Edens et. al. (2007)	Presos	Mezcla (clínico y simulador)	116	0,45					
		Laffoon, 2009	Presos	Mezcla (clínico y simulador)	100	0,13					
		Tscheuschner, 2011	Presos	Mezcla (clínico y simulador)	101	0,12					

Tabla 3.81 (continuación)

Test	Escala	Investigación	Tipo de muestra	Condición	n	Total	Ps	Dn	Am	Bi	Af
SIRS	Total	Edens et. al. (2007)	Presos	Mezcla (clínico y simulador)	116	0,81					
		Laffoon, 2009	Presos	Mezcla (clínico y simulador)	100	0,71					
		Freeman et al., 2008	Veteranos guerra	Mezcla (clínico y simulador)	41	0,54					
		Rogers, Sewell y Gillard (2010)					0,64	0,71	0,54	0,54	0,56
	RS	Heinze y Purisch (2001)	Presos	Simuladores	57	0,74		0,59			
M-Fast		West, 2004	Adultos litigantes	Mezcla (clínico y simulador)	20	0,93					
		Naldony et al., 2010	Presos	Mezcla (clínico y simulador)	247	0,57					
TOMM	Ensayo 2	Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	--	-0,6		-0,5	-0,6		-0,1
		West, 2004	Adultos litigantes	Mezcla (clínico y simulador)	--	-0,9					
	Retención	Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	--	-0,6		-0,5	-0,6		-0,1
WMT	Total	Merten et al., 2009	Pacientes forense civil TEPT	Mezcla (clínico y simulador)	--	-0,5					
		Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	--	-0,6		-0,4	-0,7		-0,2
	DR	Merten et al., 2009	Pacientes forense civil TEPT	Mezcla (clínico y simulador)	--	-0,5					
		Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	--	-0,6		-0,4	-0,7		-0,2
	CNS	Merten et al., 2009	Pacientes forense civil TEPT	Mezcla (clínico y simulador)	--	-0,5					
		Leung y Carter (2011)	Pacientes forense civil	Mezcla (clínico y simulador)	--	-0,5		-0,3	-0,6		-0,3
ASMT		Dandachi-FitzGerald et al., 2010	Pacientes psiquiátricos	Mezcla (clínico y simulador)	183	-0,2					
MSVT	IR		Universitarios	Simuladores análogos	62	-0,3					
			Pacientes psiquiátricos	Respuesta honesta	21	-0,3					
	DR	Dandachi-FitzGerald y Merckelbach, 2012	Universitarios	Simuladores análogos	62	-0,5					
			Pacientes psiquiátricos	Respuesta honesta	21	-0,5					
	CNS		Universitarios	Simuladores análogos	62	-0,5					
			Pacientes psiquiátricos	Respuesta honesta	21	-0,3					
MENT		Merten et al., 2009	Pacientes forense civil TEPT	Mezcla (clínico y simulador)	77	0,36					

Nota. ASMT: Amsterdam Short Memory Test; WMT IR: Immediate Recognition; WMT DR (Delayed Recognition); WMT CNS: Consistency; MENT: Morel Emotional Numbing Test; WMT: Word Memory Test; MSVT: Medical Symptom Validity Test. -- Se desconoce el tamaño muestral exacto a partir del cual se calcularon esas correlaciones ya que no se precisa en la investigación publicada.

3.11. Conclusiones

Si bien el proceso de detección de simulación actualmente recomendado enfatiza los aspectos multimétodo y multisistema (González-Ordi et al., 2012; Rogers, 2008b), es indudable el papel destacado y predominante que han llegado a ocupar en él los instrumentos de autoinforme, acaparando una gran parte de la atención y la investigación realizada (Sweet y Guidotti Breting, 2013).

Dentro de los autoinformes, el uso de las pruebas de *screening* ha tenido un marcado auge por razones de eficiencia (Schinka y Borum, 1993, 1994) y, entre ellas, ha destacado, particularmente, el SIMS por su amplio barrido de las distintas áreas de simulación y su facilidad de uso e interpretación (Wisdom et al., 2010), habiendo sido adaptado al holandés (Merckelbach y Smith, 2003), al alemán (Cima, Hollnack, et al., 2003) y al afrikaans (Theron et al., 2001).

En este capítulo se ha tratado de sintetizar toda la investigación disponible actualmente sobre el SIMS a escala internacional, así como extraer los principales patrones de puntuaciones, perfiles y relaciones que se deducen del conjunto de investigaciones analizadas.

A partir de la revisión de la literatura científica disponible, cabe concluir que existe una amplia evidencia que respalda la capacidad del SIMS para discriminar entre simulación y clínica genuina en diferentes contextos y en diversos patrones sintomáticos. La evidencia científica muestra tamaños del efecto moderados o grandes y correlaciones significativas con las diversas pruebas de referencia en el área. Junto a la puntuación total, las puntuaciones en sus escalas han mostrado estar dirigidas específicamente a patrones sintomáticos concretos y resultar, por lo tanto, sensibles a la presencia de simulación en dicha área.

La revisión de la literatura también ha mostrado algunas limitaciones relevantes, como, por ejemplo, que un 38% de estas investigaciones tienen una utilidad muy reducida ya que son diseños de simulación análoga sin muestra clínica de referencia y que solo un 30% abordan la utilidad diagnóstica del SIMS en contextos reales con grupos diagnósticos conocidos o diseños autosuficientes en contexto civil o penal. Por otro lado, son escasas las investigaciones que informan del poder predictivo positivo o negativo (un 34%) o de la utilidad diagnóstica de diferentes puntos de corte alternativos para la puntuación total (apenas un 29%), a pesar de que algunos autores han recomendado enfáticamente la necesidad de aumentar el punto de corte para obtener una mayor especificidad necesaria en esta área de evaluación (Clegg et al., 2009; Edens et al., 1999; Wisdom et

al., 2010). La posibilidad de establecer puntos de corte diferentes según el contexto de evaluación (civil o penal), o la sintomatología alegada (psicopatología grave o leve, cognitiva, neurológica o somática), es una cuestión que apenas ha sido abordada, probablemente porque todavía son relativamente escasos los estudios que permitan llegar a este tipo de inferencias, siendo este aspecto una de las áreas más relevantes de estudio en un futuro.

La mayoría de la investigación se ha centrado principalmente en la detección de simulación de psicopatología (76%), siendo menor el número de investigaciones que han atendido a la simulación de problemas cognitivos (27%) o a la simulación de problemas neurológicos o somáticos (un 19%). Particularmente en el área del dolor, solo ha habido un estudio que lo haya abordado con un diseño de simulación análogo, empleando únicamente una muestra de universitarios, sin una muestra clínica de referencia. Esta ausencia de investigación en el área del dolor resulta sorprendente dada la relevancia de la prevalencia de simulación en este campo, probablemente una de las más notables (Greve et al., 2009; Mittenberg et al., 2002) y los continuos llamamientos a considerar esta área como la nueva frontera a traspasar en la investigación en simulación (Alwes, 2006; Rogers y Correa, 2008).

También, es escasa la atención prestada a las distintas escalas del SIMS, solo un 57% de las investigaciones referidas informan de los descriptivos de las escalas, omitiendo el resto cualquier información al respecto. Este dato resulta sorprendente dada la utilidad de las escalas para precisar el área de simulación destacada, tal y como han argumentado distintos investigadores (p. ej., Alwes, 2006).

En su conjunto, la investigación realizada hasta la fecha con el SIMS ha mostrado su utilidad y su potencial en la detección de simulación en muy diferentes áreas, si bien ha expuesto también algunas lagunas que deberían dar lugar a líneas de investigación futuras, tal y como se abordarán en la investigación empírica de esta tesis. Entre ellas, su adaptación a otras lenguas, particularmente a aquellas con mayor número de hablantes en el mundo, el empleo mayoritario de diseños con muestras clínicas de referencia ya sea mediante diseños de simulación análogos, de grupos conocidos o de diseños autosuficientes, la valoración de puntos de corte alternativos para la puntuación total y, finalmente, la mayor profundización en sus escalas y el papel que pueden jugar en la evaluación, así como los posibles puntos de corte alternativos que deberían establecerse.

Investigación empírica

El desarrollo del campo teórico, metodológico e investigador en la simulación ha sido ingente en los últimos años, siendo superior el número de investigaciones realizadas en los últimos quince años que en los cien anteriores (Rogers, 2008b), suponiendo una auténtica explosión de desarrollo y publicación de instrumentos con este objetivo (Otto y Heilbrun, 2002).

Este desarrollo ha llevado a la realización de trabajos de validación en población española de pruebas de simulación en el área neurocognitiva (v.g., Burton, Vilar-López y Puente, 2012; Sánchez-Crespo, Jiménez, Ampudia y Merino, 2012), de autoinformes (Arce et al., 2002; Arce, Fariña, Carballal y Novo, 2009; Ortiz-Tallo et al., 2011; Sánchez Crespo, Jiménez Gómez y Ampudia Rueda, 2008; Santamaría et al., 2010) o incluso de entrevistas (Zaldivar Basurto, García Montes, López Rios, Molina Moreno y Santiago, 2008a, 2008b).

No obstante, este desarrollo no ha estado presente en la adaptación y validación en población española de herramientas de autoinforme de *screening* de simulación. Esto es especialmente importante ya que, si bien las pruebas generales mencionadas anteriormente han mostrado satisfactorios resultados, no son menos ciertas sus limitaciones, entre ellas, el tiempo de aplicación que requieren y la formación especializada que exige su aplicación e interpretación (Schinka y Borum, 1993, 1994). Además, la mayoría de estas pruebas evalúan la presencia de simulación en un determinado dominio de funcionamiento, lo que puede resultar problemático ya que es difícil determinar en qué área va a presentarse la simulación (cognitiva, neurológica, somática o emocional), lo que hace que los profesionales deban siempre aplicar un amplio conjunto de pruebas extensas y complejas, para poder hacer un barrido completo de la posible simulación.

El SIMS (Smith y Burger, 1997; Widows y Smith, 2005) ha destacado dentro de estas pruebas de *screening* por la amplitud de áreas en las que rastrea la presencia de simulación, así como la investigación acumulada sobre su utilidad y capacidad discriminativa (Wisdom et al., 2010). Como prueba de *screening*, permite un acercamiento eficiente en términos de tiempo y esfuerzo, posibilitando un amplio rastreo de diversas áreas potenciales de simulación. Los resultados disponibles en la literatura científica han mostrado, adicionalmente, que ha alcanzado tamaños del efecto iguales a superiores a pruebas más largas y complejas (Edens et al., 2007; Laffoon, 2009; Vitacco et al., 2006).

El objetivo principal de esta tesis es, por lo tanto, responder a una cuestión de importancia pragmática como es determinar la utilidad del SIMS como prueba para el *screening* de simulación en la población española. Para ello, se han diseñado dos estudios empíricos complementarios que abordan su capacidad de discriminación en el área de psicopatología y de dolor crónico,

respectivamente, combinando el uso de diseños análogos, de medidas autosuficientes y de grupos conocidos, que permitan delimitar su utilidad. Estos estudios se presentan en los capítulos 5 y 6 respectivamente.

A diferencia de investigaciones previas realizadas en la literatura existente sobre el SIMS, que han tendido a focalizarse exclusivamente en la puntuación total y en el funcionamiento de un punto de corte, en estos dos estudios empíricos se analizará la utilidad de las distintas escalas del SIMS, indicando adicionalmente el rendimiento de un rango de posibles puntos de corte para cada una de ellas, así como para la puntuación total.

Ninguna otra investigación hasta la fecha ha abordado la utilidad de una prueba de *screening* de simulación en lengua española combinando diseños análogos, de grupos diagnósticos y de medidas autosuficientes con muestras amplias y para diversos trastornos; tampoco se ha analizado, en ningún estudio previo en la literatura científica internacional, la utilidad del SIMS en la detección de simulación del dolor en grupos diagnósticos, un área que representa el nuevo desafío en la investigación en simulación y donde más investigación ha sido requerida por distintos investigadores en sus prospecciones futuras (Berry y Nelson, 2010; Rogers y Correa, 2008); tampoco ha sido investigado hasta el momento la convergencia del SIMS en la detección de simulación de dolor con una de las más recientes y prometedoras medidas de autoinforme y sus escalas de exageración de síntomas, el MMPI-2-RF (Ben-Porath y Tellegen, 2008), otro aspecto que será analizado en este trabajo.

Adicionalmente, **un segundo objetivo de esta tesis** será el estudio de la percepción de profesionales médicos sobre la prevalencia de simulación de distintos trastornos, el perfil sociodemográfico que asocian a la simulación y los principales indicios que emplean para su detección. Este estudio empírico se ha realizado con el objetivo de orientar la investigación con el SIMS hacia las áreas con mayor presencia y prevalencia de la simulación.

En síntesis, en **esta tesis se pretenden abordar los siguientes objetivos:**

- **Estudiar la percepción de profesionales médicos sobre la prevalencia de simulación de distintos trastornos, el perfil sociodemográfico que asocian a simulación y los principales indicios que emplean para su detección.** Este estudio empírico se incluye en el **capítulo 4** de esta tesis bajo el título “Prevalencia de simulación por trastornos en incapacidad temporal: percepción de los profesionales de la salud”.
- **Determinar la utilidad del SIMS como prueba para el *screening* de simulación de psicopatología en población española.** Este estudio empírico se incluye en el **capítulo 5** de

esta tesis bajo el título “Utilidad diagnóstica del SIMS en la detección de simulación de psicopatología”.

- **Determinar la utilidad del SIMS como prueba para el *screening* de simulación de lumbalgia crónica en población española.** Este estudio empírico se incluye en el **capítulo 6** de esta tesis bajo el título “Utilidad diagnóstica del SIMS en la detección de simulación de lumbalgia”.

Los resultados de las investigaciones incluidas en esta tesis pretenden aportar información clave para determinar posibles tasas preliminares de simulación por trastornos en población española, así como posibles perfiles sociodemográficos asociados a los simuladores y estrategias empleadas por los profesionales para su detección. Asimismo, se pretende establecer cuál es la utilidad del SIMS en la detección de simulación en población española y qué confianza pueden tener los profesionales en su uso. En concreto, se analizará el uso del SIMS en los distintos contextos y para la detección de diferentes patologías, determinando posibles puntos de corte y las escalas más adecuadas a cada contexto de simulación en población española.

4. Prevalencia de simulación por trastornos en incapacidad temporal: percepción de los profesionales de la salud

4.1. Objetivos

Tal y como se expuso en el marco teórico, la estimación de tasas de prevalencia en simulación es un asunto particularmente complejo y de difícil solución ya que los simuladores buscan activamente no ser detectados, lo que dificulta o imposibilita de facto llegar a conocer su prevalencia real.

Para solventar este difícil punto de partida, dos han sido las estrategias usadas para estimar la prevalencia de simulación en distintos contextos: a) la consulta mediante encuestas estructuradas a jueces expertos, es decir, profesionales cualificados de la salud que habitualmente se encuentran con esta problemática en su desempeño profesional y b) la revisión de historias clínicas archivadas de individuos a los que se ha evaluado sistemáticamente la presencia de simulación. La ausencia de protocolos sistematizados de evaluación de simulación en la práctica profesional cotidiana de nuestro país dificulta y casi imposibilita la aplicación de esta segunda estrategia.

Como se señaló en los capítulos del marco teórico de esta tesis, diversos estudios han tratado de determinar la prevalencia de la simulación reflejando una notable variabilidad en sus resultados en función del campo de estudio o de la patología abordada. La evidencia parece apuntar a que las mayores tasas de simulación frecuentemente aparecen vinculadas al ámbito civil y médico, relacionado con la valoración de la discapacidad y los posibles beneficios laborales y socioeconómicos derivados de la misma, sin contar con que muchos de los casos procedentes de otros ámbitos pueden implicar posteriormente bajas laborales, procesos de rehabilitación prolongados y valoración de discapacidad con implicaciones económicas.

Desafortunadamente, son prácticamente nulos los estudios específicos realizados en muestras españolas sobre este punto, aspecto que motivó la realización de esta investigación. En ella se trata de conocer de manera preliminar la perspectiva de profesionales que tratan con estas valoraciones sobre qué patologías se perciben como con mayores tasas de simulación y qué características se asocian a los supuestos simuladores, además de los principales criterios de sospecha que utilizan para la detección de simulación.

En concreto, los objetivos de este estudio son obtener unas estimaciones preliminares acerca de la percepción de los profesionales sobre:

- qué patologías son percibidas como más frecuentemente simuladas;
- cuál es el posible perfil sociodemográfico (edad, sexo, ingresos...) asociado a simulación por los profesionales;
- cuáles son los mecanismos utilizados para la simulación;
- en qué contexto aparece más frecuentemente (contingencia común, laboral o accidente de trabajo);
- y cuáles son las estrategias que afirman utilizar los profesionales para la detección de simulación.

4.2. Método

4.2.1. Participantes

Se dispuso de una muestra incidental de 161 médicos procedentes de 26 provincias distintas (56% mujeres, 44% varones) con edades comprendidas entre los 30 y los 59 años (media = 43,62; Dt =6,52), con un promedio de 17 años de ejercicio profesional (Dt=6,3; rango: 3–34) y de 9,4 años de antigüedad en la organización donde trabajan (Dt=6,5; rango: 0,5–34).

De los 161 médicos participantes, 100 eran médicos de mutuas, los cuales valoran en la actualidad tanto patología de contingencia común como de accidente de trabajo y enfermedad profesional, 46 médicos pertenecían a los Equipos de Valoración de Incapacidades (EVI) del Instituto Nacional de la Seguridad Social (INSS), 5 médicos Forenses del Juzgado de lo Social y 10 médicos de Atención Primaria del Sistema Público de Salud.

4.2.2. Procedimiento y materiales

Se elaboró un cuestionario estructurado, siguiendo lo propuesto por Mittenberg et al. (2002), que constaba de nueve ítems (apéndice A) y que fue entregado a cada profesional para que lo respondiera de forma anónima e individual.

El cuestionario preguntaba a cada profesional por su percepción de los siguientes aspectos:

- patologías más simuladas (ordenación y prevalencia de cada una)
- perfil sociodemográfico más prevalente en simulación (sexo, edad, nivel de ingreso y profesión)

- contexto y mecanismo de simulación
- evidencias que le hacían sospechar de la presencia de simulación

Se incluyeron aquellas patologías más frecuentemente vinculadas a simulación en estudios previos así como aquellas sobre las que se quería indagar su prevalencia en muestras españolas. Posteriormente los datos fueron codificados, grabados y analizados con el programa SPSS 15.0.

4.3. Resultados

4.3.1. Percepción sobre las patologías más simuladas

La primera pregunta del cuestionario solicitaba a los profesionales que ordenaran de 1 a 10 las patologías según su percepción sobre la prevalencia de simulación (figura 4.1).

Indique, numerando de 1 a 10 las patologías en las que usted considera que se dan más casos de simulación (1=máxima prevalencia de simulación; 10 = mínima prevalencia de simulación):

- Fibromialgia /Síndrome de Fatiga Crónica
- TCE leve
- TCE moderado
- Esguince cervical
- Cervicalgia Aguda
- Cervicalgia Crónica
- Lumbalgia y/o ciática aguda
- Lumbalgia y/o ciática crónica
- Depresión
- Ansiedad
- Fiebre de Origen desconocido
- Síndrome de colon irritable
- Otros cuadros no mencionados y que usted desee puntuar de 1 a 10
-

Figura 4.1. Pregunta 1 del cuestionario

En la siguiente tabla se puede observar qué posición concedían los profesionales a cada enfermedad o trastorno y cómo se distribuían sus valoraciones (qué porcentaje de profesionales indicaban cada una de las categorías de 1 a 10 en cada enfermedad o trastorno).

Se puede advertir que la patología que se percibe como con mayor prevalencia de simulación es el esguince cervical (con una mediana de 2 y con el 40% de los médicos considerándola la patología con mayor prevalencia), seguido de fibromialgia, cervicalgia crónica, lumbalgia crónica, depresión y ansiedad; les siguen en prevalencia percibida de simulación la cervicalgia y lumbalgia en

su fase aguda y en la parte final de la tabla los traumatismos craneoencefálicos (TCE) y la fiebre de origen desconocido y el colon irritable. A partir del análisis de los porcentajes de profesionales que responden cada categoría de 1 a 10 puede advertirse la amplia variabilidad en cuanto a las valoraciones de prevalencia de los distintos profesionales.

Tabla 4.1

Ranking de prevalencia de simulación en diferentes patologías de acuerdo a la percepción de los profesionales (1= máxima prevalencia de simulación; 10 = mínima prevalencia de simulación)

Patología	Ranking de simulación (posición de 1 a 10 de mayor a menor prevalencia percibida)		Porcentaje de profesionales que responden cada categoría (posición de 1 a 10 de mayor a menor prevalencia de simulación percibida)									
	Mediana	Rango intercuartílico	1	2	3	4	5	6	7	8	9	10
Esguince cervical	2	1 – 4	40	19	9	8	7	7	7	1	1	1
Fibromialgia	4	1 – 8	26	15	7	9	5	9	3	14	3	8
Cervicalgia crónica	4	3 – 6	6	9	23	15	15	10	13	3	1	4
Lumbalgia crónica	4	3 – 6	3	13	13	23	16	11	7	6	2	5
Depresión	5	3 – 7	7	9	17	11	9	15	8	11	6	6
Ansiedad	5	3 – 7	5	11	10	11	17	11	9	5	5	14
Cervicalgia aguda	7	4 – 10	3	9	7	7	13	6	9	11	7	27
Lumbalgia aguda	7	4 – 10	5	7	7	8	8	7	11	10	5	32
TCE leve	9	6 – 10	1	3	4	3	8	7	8	6	22	37
TCE moderado	10	8 – 10	2	1	2	1	8	7	9	7	63	--
Fiebre desconocida	10	10 – 10	--	--	--	1	1	--	--	3	8	87
Colon irritable	10	10 – 10	--	1	1	--	1	2	5	5	9	76

Nota. TCE: Traumatismo craneoencefálico.

Cabe señalar que se dejó un apartado abierto para que los profesionales mencionaran otras patologías que consideraban relevantes en cuanto a su prevalencia de simulación. En la tabla siguiente se indican qué otras patologías se mencionaron. Como se puede observar, solo hubo tres que fueron mencionadas por más de un 4% (omalgia, epicondilitis y gonalgia) todas ellas con una estimación relativamente baja en cuanto a su prevalencia de simulación.

Tabla 4.2

Otros cuadros citados por los profesionales en la percepción de 1 a 10 de prevalencia de simulación (1= máxima prevalencia de simulación; 10 = mínima prevalencia de simulación)

	Frecuencia		Ranking estimado por los profesionales que la mencionaban (1-10) Mediana
	Número de veces nombrada	Porcentaje de profesionales que la mencionaban	
Omalgia	11	6,8	8
Epicondilitis	10	6,2	7
Gonalgia	7	4,3	8
Mareos	2	1,2	5
Tendinitis	2	1,2	10
Vértigos	2	1,2	-
Cefalea crónica	1	0,6	10
Cefalea ocular	1	0,6	9
Cefaleas	1	0,6	7
Coxalgia	1	0,6	10
Dolor en trauma costal	1	0,6	9
Dorsalgia	1	0,6	6
Efectos secundarios de la medicación	1	0,6	10
Esguince muñeca	1	0,6	7
Esguince tobillo	1	0,6	10
Esguinces de extremidades	1	0,6	8
Lesiones por agresiones	1	0,6	10
Pérdida visión	1	0,6	8
Síncopes	1	0,6	8
Tendinitis muñeca	1	0,6	6
Tenosinovitis EESS	1	0,6	6

Nota. TCE: Traumatismo craneoencefálico.

La segunda pregunta que se hacía a los profesionales indagaba sobre su percepción acerca del porcentaje de simulación que consideraban que existía en cada una de las patologías (figura 4.2).

En la tabla 4.3 puede observarse cuál es el porcentaje promedio estimado por los profesionales. Puede volver a advertirse que las patologías que se estiman con mayor porcentaje de simulación son esguince cervical, fibromialgia, cervicalgia crónica, lumbalgia crónica, depresión y ansiedad, todas ellas con porcentajes cercanos o superiores al 50%. Se estimaron tasas sensiblemente menores en las patologías dolorosas agudas (cervicalgia y lumbalgia), en los

traumatismos craneoencefálicos (leve y moderado) y en la fiebre de origen desconocido y colon irritable.

En las mismas patologías de la pregunta anterior estime, a su criterio, el porcentaje que hay de exageración de síntomas/simulación:

- Fibromialgia/Síndrome de Fatiga Crónica
- TCE leve
- TCE moderado
- Esguince cervical
- Cervicalgia Aguda
- Cervicalgia Crónica
- Lumbalgia o ciática aguda
- Lumbalgia o ciática crónica
- Depresión
- Ansiedad
- Fiebre de Origen desconocido
- Síndrome de colon irritable
- Otros cuadros no mencionados y que usted desee valorar
-

Figura 4.2. Pregunta 2 del cuestionario

El análisis del rango intercuartílico (la diferencia entre el primer y el tercer cuartil de la distribución de respuestas de los profesionales) y del rango (mínimo porcentaje estimado y máximo porcentaje estimado) permite ver la amplia variabilidad y diversidad en las valoraciones de los profesionales que parece evidenciar una importante falta de consenso.

Tabla 4.3
Porcentaje de prevalencia de simulación percibido por los profesionales en diferentes patologías

Patología	Porcentaje de simulación estimado en cada patología		
	Mediana	Rango intercuartílico	Rango (min. – max.)
Esguince cervical	60	50 – 80	5 - 100
Fibromialgia	50	30 – 80	1 - 100
Cervicalgia crónica	50	30 – 60	5 - 100
Lumbalgia crónica	45	30 – 60	5 - 100
Depresión	50	30 – 60	5 - 90
Ansiedad	50	30 – 60	0 - 90
Cervicalgia aguda	30	20 – 50	1 - 100
Lumbalgia aguda	30	20 – 50	1 - 90
TCE Leve	20	10 – 40	1 - 90
TCE Moderado	10	5 – 30	0 - 80
Fiebre origen desconocido	15	5 – 20	0 - 60
Colon irritable	20	10 – 28	0 - 60

4.3.2. Percepción sobre el perfil sociodemográfico prevalente en simulación

Se les pidió a los profesionales que valoraran una serie de características sociodemográficas (sexo, edad, nivel de ingresos, profesión...) de acuerdo a su vinculación con la presencia de simulación. Así, se les pidió que indicaran de los pacientes que simulaban síntomas, qué porcentaje eran varones y cuál mujeres (figura 4.3).

Entre los pacientes que usted cree que exageran/simulan síntomas (escriba el porcentaje que usted estime)

-----% son mujeres

-----% son varones

Figura 4.3. Pregunta 8 del cuestionario

Como se puede advertir en la siguiente figura, el 21% de los profesionales consideraban que, entre los pacientes que simulaban sus síntomas, era idéntico el porcentaje de varones y mujeres, el 15% consideraban que eran más frecuentes los varones y el 63% que eran más frecuentes las mujeres.

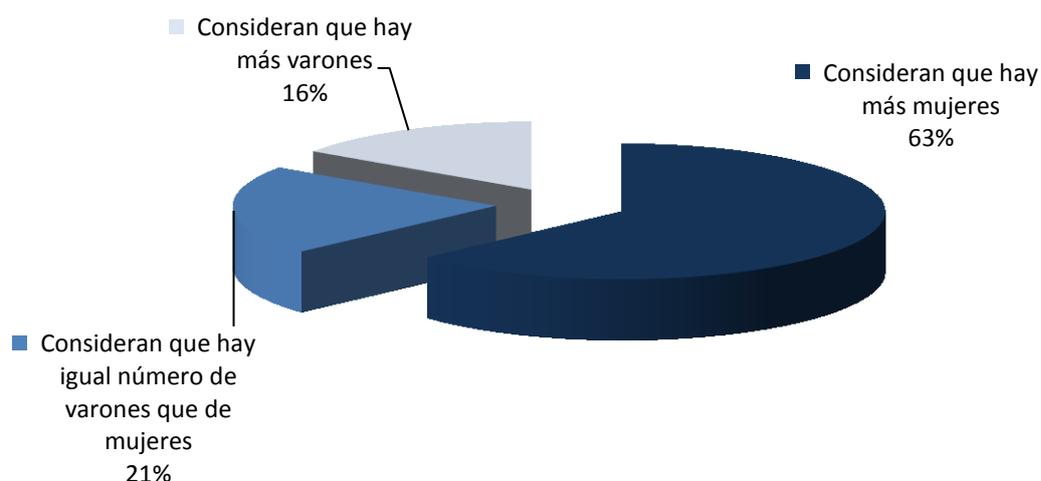


Figura 4.4. Percepción de los profesionales sobre la presencia de mujeres y varones en los pacientes que simulan sus síntomas

Cuando se indaga exactamente sobre la ratio mujeres-varones que informan los profesionales (tabla 4.4), se puede ver que la mayoría de los profesionales se inclinan por considerar que del conjunto de pacientes que simulan el 60% son mujeres y el 40% son varones (el 40,6% de los profesionales responden esta distribución de porcentajes). La amplia mayoría de los profesionales (75,5%) respondían una distribución de porcentajes que oscilaba entre el 40% varones y 60% mujeres y el 60% varones y 40% mujeres.

Tabla 4.4**Percepción de los profesionales sobre la presencia de mujeres y varones en simulación**

De los pacientes que simulan, ¿qué porcentaje son mujeres y qué porcentaje son varones?(% mujeres - % varones) Opciones de respuesta dadas en la encuesta	Porcentaje de profesionales que respondían cada posible opción de respuesta %
90 – 10	0,6
85 – 15	0,6
80 – 20	4,5
75 – 25	1,9
70 – 30	12,3
65 – 35	0,6
60 – 40	40,6
55 – 45	1,9
50 – 50	21,3
45 – 55	0,6
40 – 60	11,0
30 – 70	1,9
25 – 75	1,9

También se les solicitó a los profesionales que ordenarán los grupos de edad y niveles de ingresos según la prevalencia de simulación en cada uno de ellos, de acuerdo a su experiencia (figura 4.5). Así, tuvieron que ordenar los cinco grupos de edad asignando a cada uno de ellos un número entre 1 y 5 (siendo 1 el que más prevalencia de simulación tenía y 5 el que menos) y los seis niveles de ingresos asignando a cada uno de ellos un número entre 1 y 6 (siendo 1 el que más prevalencia de simulación tenía y 6 el que menos).

En qué tramos de edad piensa usted que se dan más casos de exageración de síntomas simulación:

numere de 1 (más casos) a 5 (menos casos):

- 16 a 25 años
- 26 a 35 años
- 36 a 45 años
- 46 a 55 años
- 55 al final de la vida laboral

Qué salario medio mensual bruto piensa usted que tiene con más frecuencia el paciente que exagera/simula síntomas (numere de 1 a 6 de más frecuente a menos)

- menos de 1.000 euros/mes
- de 1.000 a 1.500 "
- de 1.500 a 2.000 "
- de 2.000 a 2.500 "
- de 2.500 a 3.000 "
- más de 3.000 euros /mes

Figura 4.5. Preguntas 3 y 5 del cuestionario

Puede observarse en la siguiente tabla como los profesionales consideran que es más probable la presencia de simulación en el rango de mediana edad que en los extremos. El grupo de 46 a 55 años es el que más prevalencia de simulación presenta de acuerdo a la percepción de los profesionales.

Tabla 4.5

Ranking de prevalencia de simulación en los grupos de edad de acuerdo a la experiencia y percepción de los profesionales (1= máxima prevalencia de simulación; 5 = mínima prevalencia de simulación)

Grupo	Ranking (posición de 1 a 5)		Porcentaje de profesionales que responden cada categoría de 1 a 5				
	Mediana	Rango intercuartílico	1	2	3	4	5
46-55 años	2	2 – 3	23	33	20	18	6
36-45 años	3	1 – 3	26	23	28	13	10
26-35 años	3	2 – 4	21	13	17	43	6
> 55 años	3	2 – 5	20	17	23	14	26
16-25 años	5	2 – 5	12	13	13	12	50

Nota. Se destaca en negrita la respuesta (de 1 a 5) más marcada por los profesionales en cada grupo de edad.

En relación a los niveles de ingresos, hay una tendencia marcada por parte de los profesionales a considerar que la prevalencia de simulación es más marcada conforme el nivel de ingresos es menor, siendo el grupo con más prevalencia el que tiene ingresos inferiores a 1.000 euros al mes.

Tabla 4.6

Ranking de prevalencia de simulación según los niveles de ingresos de acuerdo a la experiencia y percepción de los profesionales (1= máxima prevalencia de simulación; 6 = mínima prevalencia de simulación)

Grupo	Ranking (posición de 1 a 6)		Porcentaje de profesionales que responden cada categoría de 1 a 6					
	Mediana	Rango intercuartílico	1	2	3	4	5	6
Menos de 1.000 euros/mes	1	1 – 3	59	15	16	5	3	2
De 1.000 a 1.500 euros/mes	2	1 – 2	26	66	3	3	2	-
De 1.500 a 2.000 euros/mes	3	3 – 3	8	11	74	4	1	1
De 2.000 a 2.500 euros/mes	4	4 – 4	3	5	5	83	2	1
De 2.500 a 3.000 euros/mes	5	5 – 5	2	2	1	3	91	1
Más de 3.000 euros/mes	6	6 – 6	1	-	-	4	1	94

Nota. Se destaca en negrita la respuesta (de 1 a 6) más marcada por los profesionales en cada grupo de salarios.

También se les solicitó que indicaran cinco profesiones en las que, desde su experiencia profesional, consideraran que se producían más casos de exageración de síntomas/simulación. En

este caso no se le daba un listado de profesiones sino que era un campo abierto donde anotaban las profesiones que consideraban oportunas (figura 4.6).

Escriba las 5 profesiones en las que, desde su experiencia y en el medio en el que usted trabaja, se producen más casos de exageración de síntomas /simulación (la primera en la que usted ha observado más casos y así hasta 5)

- 1-.....
- 2-.....
- 3-.....
- 4-.....
- 5-.....

Figura 4.6. Pregunta 4 del cuestionario

En la siguiente tabla se indican las profesiones más mencionadas, el porcentaje de los profesionales que las mencionaban y en qué posición (de 1º a 5º) solía aparecer en la enumeración (siendo la 1º la que más veces se vincula a simulación).

Al consultar la tabla, puede advertirse como en la parte superior aparecen principalmente profesiones del sector de la limpieza, de la construcción, la industria y los operarios (cadena de montaje/producción, mozo almacén carga y descarga, mantenimiento, servicios, reponedora, albañil, ...), de administración, de atención al público (cajera, dependiente, azafata, camarero, auxiliar hostelería, teleoperador...) y de cuidados sanitarios (auxiliar de clínica, de enfermería o geriatría, ayuda a domicilio...). Este conjunto de profesiones aglutinan el 80% de las menciones y, por lo tanto, representan el eje central de ocupaciones que los profesionales perciben como más asociadas a la presencia de simulación. El 20% restante se reparte en ocupaciones menos señaladas como son policías, profesores, vigilantes, jardineros, cámaras de televisión...

Tabla 4.7**Profesiones más vinculadas a la presencia de exageración de síntomas de acuerdo a la percepción de los profesionales**

Profesión	Nº de veces citada	% de los profesionales que la citan	Nº veces citada en cada posición				
			1	2	3	4	5
Limpieza	130	82,3	72	25	18	10	5
Construcción (albañil, obrero, encofrador, soldador, carpintería, electricista...)	110	69,6	23	35	27	13	12
Administrativo	73	46,2	10	17	13	19	14
Cajera, reponedora, dependiente, azafata	47	29,7	2	7	6	10	8
Auxiliar de clínica/enfermería/geriatria/sanitario	41	25,9	6	11	11	8	5
Cadena de montaje/producción	36	22,8	7	9	9	5	6
Ayuda a domicilio	33	20,9	11	8	7	4	3
Teleoperador	26	16,5	4	3	5	11	3
Mozo almacén y carga-descarga	25	15,8	3	0	6	11	5
Funcionario	23	14,6	2	8	7	3	3
Hostelería (camarero, auxiliar hostelería...)	21	13,3	0	2	5	7	7
Mantenimiento, operario, montador, servicios, conserje	16	10,1	--	3	3	3	7
Agrario-ganadería	14	8,9	2	6	1	5	0
Autónomo	14	8,9	3	2	4	3	2
Policía	13	8,2	4	3	4	0	2
Profesor	11	7,0	1	1	0	6	3
Vigilante	11	7,0	0	4	3	2	2
Comercial	10	6,3	0	1	3	1	5
Conductor	8	5,1	0	2	0	0	6
Camarera piso	6	3,8	0	2	3	1	0
Barrendero	6	3,8	0	1	1	3	1

Nota. n=158, tres profesionales no incluyeron su respuesta a esta pregunta. En aras de la reducción de espacio, no se han incluido en la tabla aquellas profesiones mencionadas por menos de 6 profesionales de la salud.

4.3.3. Contexto y mecanismo de simulación

Otra área en la que se solicitó su opinión a los profesionales fue el tipo de contingencia que se asociaba con mayor frecuencia a simulación y el porcentaje con que se empleaban distintos mecanismos de simulación (figura 4.7).

En relación al tipo de contingencia, tal y como se indica en la tabla 4.8, el más frecuente era la contingencia común, seguido del accidente de trabajo y finalmente la enfermedad profesional.

Tabla 4.8**Percepción de los profesionales sobre el tipo de contingencia donde se da más simulación**

Tipo de contingencia	n	%
Accidente de trabajo	47	29,2
Contingencia común	110	68,3
Enfermedad profesional	3	1,9

En qué tipo de contingencia cree usted que se produce más simulación: (marque con una x)

- Accidente de trabajo
- Enfermedad profesional
- Contingencia Común

Entre los pacientes que usted piensa que simulan/exageran síntomas: escriba el porcentaje que estima usted que lo hace en las siguientes situaciones:

- % simulan el mecanismo de producción
- % perpetúan síntomas una vez que estos han mejorado
- % simulan la patología completa (no existe y la simulan)
- % simulan o exageran síntomas por que adoptan el rol de enfermo
- % simulan síntomas por que tiene así un beneficio externo (tribunal, incapacidad, invalidez, seguro coche....)
- % simulan por que existe un conflicto laboral

Figura 4.7. Preguntas 6 y 7 del cuestionario

Cuando se les preguntó a los profesionales sobre el tipo de simulación que consideraban que empleaban los pacientes simuladores, tal y como se puede observar en la tabla 4.9, indicaron que los mecanismos más empleados eran la perpetuación de síntomas una vez habían mejorado (en término medio los profesionales consideraban que esto ocurría en el 55% de los casos) y la simulación de síntomas cuando tenían un beneficio externo (en el 50% de los casos). En menor medida aparecían la simulación por un conflicto laboral (30%), la simulación del mecanismo de producción (20%), el rol de enfermo (20%) o la simulación completa de la patología (10%).

Tabla 4.9*Percepción de los profesionales sobre qué porcentaje de simuladores utiliza cada una de las siguientes estrategias de simulación*

Estrategias de simulación	Estadísticos descriptivos		Porcentaje de profesionales que responden en cada rango de porcentajes										
	Mediana	Rango intercuartílico	0% - 5%	6% - 10%	11% - 20%	21% - 30%	31% - 40%	41% - 50%	51% - 60%	61% - 70%	71% - 80%	81% - 90%	91% - 100%
Perpetúan síntomas una vez que estos han mejorado	55	30 – 80	4	3	10	18	7	8	9	11	13	14	4
Simulan síntomas por que tiene así un beneficio externo (tribunal, incapacidad, invalidez, seguro coche....)	50	30 – 80	3	4	9	18	9	11	10	3	19	9	5
Simulan por que existe un conflicto laboral	30	20 – 60	7	10	20	20	7	9	8	6	10	2	1
Simulan el mecanismo de producción	20	10 – 50	8	21	23	11	9	11	5	4	5	4	1
Simulan o exageran síntomas por que adoptan el rol de enfermo	20	10 – 50	15	27	11	9	6	15	3	5	6	2	1
Simulan la patología completa (no existe y la simulan)	10	5 – 24	29	30	16	10	5	7	--	1	2	1	1

4.3.4. Datos que hacen sospechar simulación a los profesionales

Finalmente se solicitó a los profesionales que indicaran cuatro datos que le hicieran sospechar la existencia de simulación, siendo un campo abierto donde anotaban las evidencias que consideraban oportunas (figura 4.8).

Escriba, por orden de importancia, los 4 datos que a usted le hacen sospechar la existencia de exageración de síntomas/simulación:

- 1-----
- 2-----
- 3-----
- 4-----

Figura 4.8. Pregunta 9 del cuestionario

En la siguiente tabla se indican las evidencias más mencionadas, el porcentaje de los profesionales que las mencionaban y en qué posición (de 1º a 4º) solía aparecer en la enumeración de los profesionales (siendo la 1ª a la que más importancia concedían como evidencia de simulación).

Al consultar la tabla, puede advertirse como los principales datos que hacen sospechar a los profesionales la existencia de simulación son la ausencia de datos objetivos que justifiquen los síntomas referidos por el paciente (lo menciona un 69%), la exageración de síntomas (43%), la presencia de beneficios colaterales (un 38%), la ausencia de respuesta al tratamiento (un 28%), una situación laboral conflictiva (un 28%), la falta de colaboración o adherencia al tratamiento (19%), la descripción de los síntomas o mecanismos de producción imprecisa, contradictoria o con dudas y omisiones (un 16%) y una actitud hostil o agresiva (un 16%). En su conjunto, estos datos agrupan el 70% de las menciones de los profesionales (391 de los 554 datos mencionados en la encuesta), pudiendo considerarse el núcleo central de evidencias en que se basan para la sospecha de simulación. El 30% restante se reparte en datos menos señalados. En la figura 4.9 puede observarse gráficamente el número de veces que se menciona cada dato y la pendiente decreciente que aparece.

Tabla 4.10

Datos clínicos que hacen sospechar a los profesionales la existencia de simulación de acuerdo a su experiencia

Datos de sospecha clínica	Nº de veces citada	% de la muestra que lo cita	Nº veces citada en cada posición			
			1º	2º	3º	4º
Ausencia de evidencias objetivas o signos objetivos: discordancia síntomas y signos, pruebas normales (exploración normal y muchos síntomas referidos, ausencia de datos objetivos, pruebas diagnósticas negativas, no concuerdan signos y síntomas, discordancia clínica y exploración y pruebas complementarias...)	105	68,6	45	26	23	11
Dramatización o magnificación de síntomas: descripción de los síntomas o reacción a la exploración desproporcionada a la patología (sobreactuación sintomatológica, quejas excesivas y exageradas, "no le has tocado y ya le ha dolido", dolor al roce o palpación de la piel, llanto incontrolado, "chillidos" con patología leve...)	66	43,1	21	28	13	4
Presencia de beneficios colaterales a la baja: accidente y compensaciones, presencia de conflictos personales, económicos o familiares, época del año (accidente de tráfico, posibilidad de rentabilizar baja con compañía de seguros, pendiente evaluación de secuelas, problemas económicos o familiares, edad prejubilación, fecha fin de contrato, los niños ya no tienen colegio, proximidad de fiestas o vacaciones...)	58	37,9	11	12	10	25
Ausencia de respuesta al tratamiento (no cede el dolor con el tratamiento, no mejora con el tratamiento, sigue empeorando a pesar del tratamiento, mala respuesta a múltiples tratamientos, nunca está bien ni refiere mejoría...)	43	28,1	8	15	12	8
Situación laboral conflictiva o poco motivante, tipo de trabajo o profesión (problemas o conflictos laborales, críticas a su empresa o trabajo, refiere situación laboral difícil, baja estima por su trabajo, escasa remuneración y poco gratificante el trabajo...)	42	27,5	4	11	16	11
Ausencia de colaboración o de adherencia al tratamiento (no se deja explorar, no sigue el tratamiento, es reticente al tratamiento, no quiere ir a una interconsulta...)	29	19,0	4	3	9	13
Descripciones de los síntomas o mecanismos de producción imprecisas, contradictorias, con dudas y omisiones o mal localizados (cambian los síntomas, no saben concretar mecanismo de producción, responden diferente sobre la misma pregunta, signos contradictorios al explorar, incongruencia de respuestas...)	25	16,3	7	9	5	4
Actitud hostil, agresividad (susplicacia, desconfianza, inconformismo, enfado al no cumplir las expectativas, descalificaciones al profesional, conflicto con la seguridad social o la mutua...)	25	16,3	2	6	11	7
Incoherencia de la patología con las alteraciones o dolor que informa, sintomatología atípica, clínica no habitual	22	14,4	11	5	3	3
Alargamiento y perpetuación de los síntomas (duración exagerada de la lesión, prolongación fuera de lo razonable, persisten síntomas y mejora la exploración, relatan los mismos síntomas que el primer día...)	17	11,1	5	5	2	5

Tabla 4.10 (continuación)

Datos de sospecha clínica	Nº de veces citada	% de la muestra que lo cita	Nº veces citada en cada posición			
			1º	2º	3º	4º
Comenta con excesivo énfasis e insistencia que no quiere bajas y que lo único que quiere es curarse y volver a trabajar ("no he venido antes porque no quiero bajas", deseo extremo de curación, el deseo continuo manifiesto de irse a trabajar, "lo primero es curarme"...)	14	9,2	3	2	4	5
Múltiples síntomas, dolor "móvil", siempre aparece un nuevo síntoma (cuando mejora una patología aparece otra nueva, múltiples quejas que van cambiando de localización, pluripatología, dolores múltiples...)	13	8,5	1	2	5	5
Muestran más intereses por aspectos económicos, legales y administrativos que médicos (preguntan cuánto van a cobrar, solicitan ellos mismos la baja, te hablan lo primero de la baja, contratan a un abogado lo primero, tratan de dirigirte hacia el diagnóstico que quieren...)	12	7,8	1	4	5	2
Mecanismo causal (de producción) no coincide con la lesión (incongruencia entre el mecanismo de producción y la patología, falta de relación causa efecto...)	11	7,2	6	4	1	0
Historial (historial de bajas repetidas, hiperfrecuentación a servicio sanitario...)	10	6,5	0	3	5	2
Actitud extraña o poco frecuente del paciente hacia su lesión y su tratamiento	9	5,9	5	1	0	3
Características sociodemográficas (edad, sexo, nacionalidad, nivel de ingresos...)	8	5,2	1	0	4	3
Mejorías o empeoramientos vinculados a factores ajenos a la clínica o el tratamiento (al pedirle pruebas complementarias mejora, al obtener su objetivo o indemnización mejora, al acercarse al alta empeora, al haber días festivos mejora...).	7	4,6	0	2	2	3
Discrepancia con maniobras de distracción, no ortodoxas o disuasorias	7	4,6	2	3	0	2
Descripción del síntoma o dolor excesivamente detallada, florida o barroca.	6	3,9	3	2	0	1
Forma de caminar, posturas y movilidad forzada y no natural del paciente cuando se siente observado	6	3,9	2	1	3	0
Lenguaje no verbal	6	3,9	5	1	0	0

Nota. n=153, ocho profesionales no incluyeron su respuesta a esta pregunta. En aras de la reducción de espacio, no se han incluido en la tabla aquellos datos de sospecha clínica mencionados por menos de 6 profesionales de la salud.

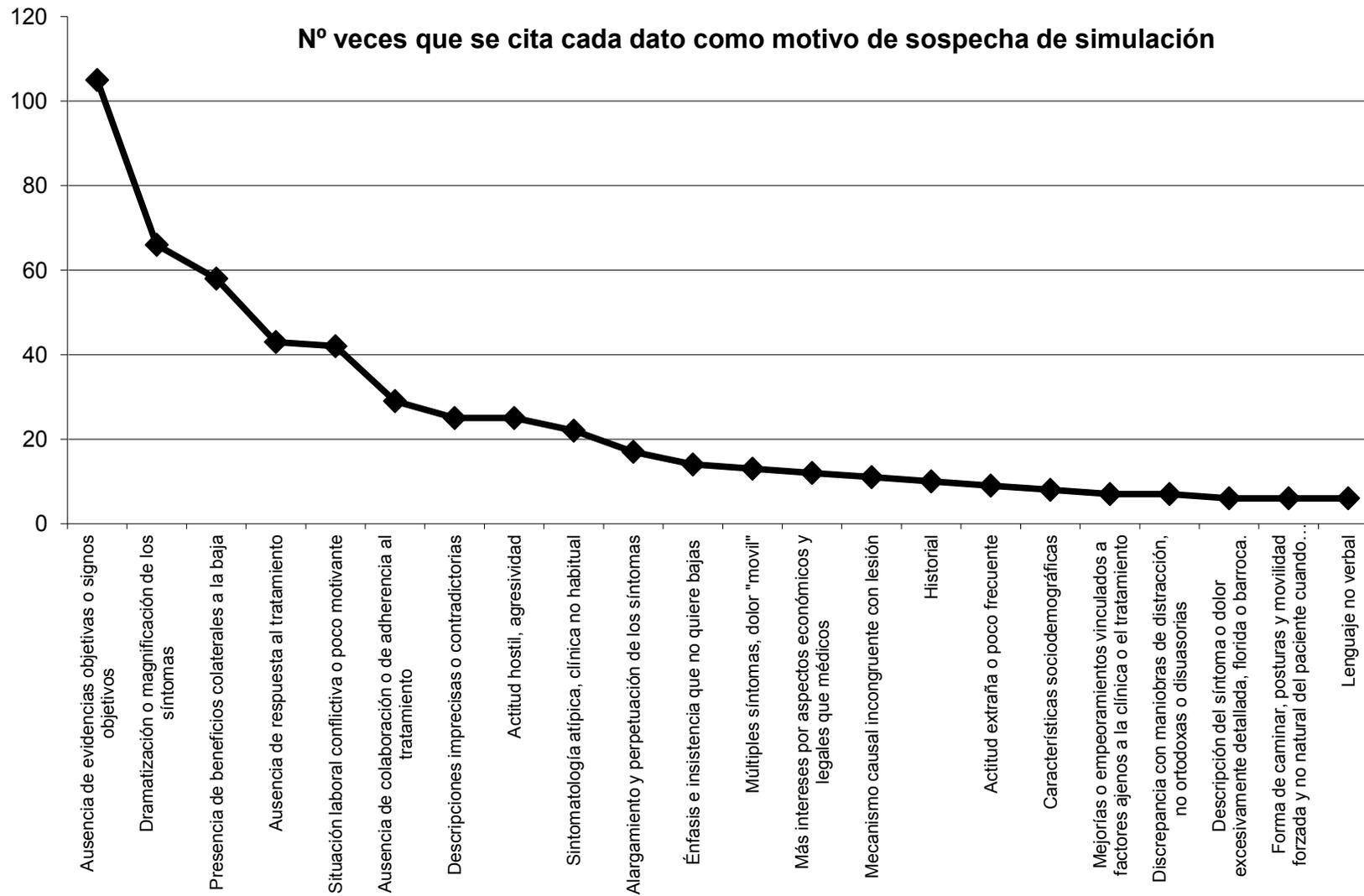


Figura 4.9. Número de menciones por parte de los profesionales a cada dato de sospecha de simulación

4.4. Discusión

Los objetivos de este estudio eran obtener unas estimaciones preliminares sobre la percepción de los profesionales de las patologías percibidas con mayores tasas de simulación, el posible perfil sociodemográfico (edad, sexo, ingresos...) asociado a simulación y los datos que afirmaban utilizar como indicadores de sospecha de simulación. Hasta donde llega nuestro conocimiento, ningún estudio hasta la fecha ha abordado esta cuestión en España, a pesar de su relevancia económica, social y científica.

En relación con las **patologías con mayores tasas de simulación** los resultados muestran una mayor percepción de simulación de las patologías más asociadas a dolor crónico y trastornos de ansiedad y afectivos (esguince cervical, fibromialgia, cervicalgia crónica, lumbalgia crónica, depresión y ansiedad), todas ellas con porcentajes estimados de simulación con medianas cercanas o superiores al 50%. Es interesante destacar que el conjunto de patologías señaladas parecen caracterizarse por su carácter crónico y perdurable en el tiempo, algo relevante desde el punto de visto del simulador, su marcada subjetividad con una limitada evidencia objetiva de las patologías y su mayor prevalencia clínica, lo que hace que sea probable que tengan experiencia de las mismas, de sus síntomas y de su curso (por su propia experiencia o de familiares o amigos cercanos) o que pueda prolongarlas en el tiempo una vez que las padezca.

Las patologías percibidas con menores tasas de simulación y valores estimados inferiores al 35% fueron las patologías dolorosas agudas (cervicalgia y lumbalgia aguda), los traumatismos craneoencefálicos (leve y moderado) y la fiebre de origen desconocido y colon irritable. Estas patologías pudieran caracterizarse por un carácter menos crónico y más puntual, menos interesante por lo tanto a priori para un simulador, o con menor prevalencia clínica y de las que hay menos experiencia de sus síntomas, presentación o curso.

Adicionalmente cabe señalar que no parece que haya ninguna otra patología que haya sido señalada por los profesionales en el campo de respuesta abierta que ocupe los primeros lugares o que haya sido mencionada de forma relevante o significativa. Las más mencionadas que no aparecían en el listado proporcionado han sido omalgia, epicondilitis y gonalgia, con posiciones medias en el ranking cercanas a la posición 8ª (en una escala de 1 a 10, siendo 1 máxima prevalencia y 10 mínima prevalencia).

Estos resultados concuerdan con los obtenidos por Mittenberg et al. (2002) al destacar la relevancia de la simulación en los trastornos asociados a dolor y a la ansiedad y depresión. En general, los datos indican que en el contexto médico-legal los patrones de simulación están más

relacionados con trastornos y enfermedades físicas o trastornos afectivos o emocionales con la posibilidad de obtener incentivos económicos (McDermott y Feldman, 2007).

No obstante, los resultados obtenidos en este estudio discrepan en cuanto a la relevancia de la simulación del traumatismo craneoencefálico leve (TCE leve). Esta patología es considerada en el estudio de Mittenberg et al. (2002) como entre las más simuladas, mientras que de acuerdo a los resultados de nuestro estudio se encontraría entre las menos simuladas. Esta divergencia puede ser explicada tanto por diferencias en las muestras de profesionales consideradas como por distintas contingencias diagnósticas, económicas y laborales asociadas al TCE leve en España y en Estados Unidos. En relación a las muestras consideradas, el estudio de Mittenberg et al. fue respondido por 131 neuropsicólogos que estimaron las tasas de prevalencia de acuerdo a su archivo de casos de consultas privadas, siendo los neuropsicólogos típicamente la profesión más asociada a la evaluación de la simulación en Estados Unidos y Europa; en la muestra considerada en esta investigación fueron predominantemente médicos de mutuas y de valoración de discapacidades del sistema público de salud cuya experiencia reside en la valoración y supervisión de todo tipo de bajas laborales.

Con respecto a las tasas de simulación informadas en esta investigación, los porcentajes percibidos por los profesionales en las patologías más prevalentes de simulación han oscilado entre el 45 y el 60%. Estos valores se encuentran en el rango superior informado por la literatura científica, que habitualmente oscila entre el 20% y el 50% (Gervais et al., 2001; Greve et al., 2009; Kay y Morris-Jones, 1998; Leavitt y Sweet, 1986; Mittenberg et al., 2002; van der Werf, Prins, Jongen, van der Meer y Bleijenbergh, 2000), si bien concuerdan con diversos estudios que afirman tasas comprendidas entre el 40 y el 60% o incluso superiores (Chafetz, 2008; Hammond, 2006; Jorge y Jorge, 2011; Merten et al., 2006; Schmand et al., 1998; Stevens et al., 2008), particularmente en trabajos que evalúan la prevalencia de simulación en solicitantes de discapacidad en sistemas públicos de salud.

Es importante reseñar de los resultados obtenidos la gran variabilidad existente entre los profesionales en cuanto a la ordenación y estimación de porcentajes de simulación de las patologías. Esto puede ser debido a la heterogeneidad de la muestra y a la diferente frecuencia con que se enfrentan a la simulación, al componente subjetivo del “diagnóstico” de simulación, a la diferente experiencia de los profesionales en este campo o la heterogénea formación y conocimiento de esta área entre los profesionales.

Parece que esta heterogeneidad es habitual en estos estudios. Así, por ejemplo Rogers et al. (1998) informaban de una marcada variabilidad en las estimaciones de la prevalencia o en el estudio de Slick, Tan, Strauss y Hultsch (2004) con una muestra de 24 expertos en contexto neuropsicológico

(seleccionados por su participación en revistas científicas con artículos en esta área y considerados expertos específicos en simulación) encontraba que un 16,7% consideraban que la prevalencia de simulación era inferior al 5%, un 16,7% entre el 5 y el 10%, un 33,3 que entre el 10 y el 20%, un 20,8% que era entre el 20 y el 30% y un 12,5% que era superior al 30%.

En cuanto al **perfil sociodemográfico** percibido por los profesionales, estos consideran que la simulación es más prevalente conforme menores son los ingresos, en medianas edades (particularmente 46-55 años), en profesiones menos estimulantes y vocacionales y en mujeres ligeramente más que en varones. Estos resultados concuerdan con algunos de los encontrados en la literatura científica, que asocian la mayor prevalencia de bajas laborales y sospecha de simulación en mujeres frente a varones, conforme mayor es la edad, menores son los ingresos y los trabajos son más pesados y menos satisfactorios (Barraquer, 1992; Gjesdal, Ringdal, Haug y Mæland, 2004; Karlsson, Carstensen, Gjesdal y Alexanderson, 2008; Quezada–Ortega, Razo–Mondragón, Marín-Cotoñieto, Salinas-Tovar y López-Rojas, 2006; Steenstra, Verbeek, Heymans y Bongers, 2005). Así, por ejemplo, el estudio de Aurrekoetxea Agirre, Sanzo Ollakarizketa, Zubero Oleagoitia y Alamillo Gascón (2009) mostraba un incremento significativo de la reiteración en las bajas a partir de los 45 años o el estudio de Benavides et al. (2007) observó una duración media en las bajas más altas para personas mayores de 40 años, mujeres y para los que trabajan en hostelería o el estudio de Catalina-Romero et al. (2013) relacionaba la mayor duración de la baja por trastorno de ansiedad con una edad superior a 35 años, un menor nivel de estudios y la presencia de desempleo durante la baja.

En particular, en relación con la mayor prevalencia de las mujeres, el estudio de Rogers et al. (1998) basado también en una encuesta a profesionales sobre el prototipo de simulador, encontraba que en contextos no forenses eran muy similares las tasas de varones y mujeres (49,9%). Ahora bien, los profesionales percibían que las mujeres que se embarcaban en una conducta de simulación solían informar más típicamente de síntomas médicos o somáticos y menos de síntomas cognitivos que los varones. Resulta incierto determinar en qué medida esta percepción de los profesionales es verosímil o en qué medida es debida a un sesgo posiblemente relacionado con los conceptos históricos de histeria o trastorno somatomorfo asociado al sexo femenino (Rogers y Reinhardt, 1998). Es interesante citar también a este respecto el estudio de Brage, Nygard y Tellnes (1998) que halló notables diferencias en la prevalencia y duración de las bajas relacionadas con enfermedades musculoesqueléticas en mujeres frente a varones, si bien dichas diferencias se veían notablemente reducidas cuando se controlaba la edad y el nivel de ingresos. Este hecho puede sugerir que las diferencias por sexo pueden ser explicadas, cuando menos parcialmente, por diferencias subyacentes relacionadas con el nivel de ingresos, las profesiones, la proyección profesional o las

demandas sociales y familiares asociadas habitualmente al género femenino. No obstante, este es un punto claramente incierto a delimitar en posteriores investigaciones.

En relación al tipo de contingencia y al mecanismo de producción, los resultados de la investigación han señalado la mayor prevalencia de la simulación en el contexto de contingencias comunes (con menos control y seguimiento) y empleando como mecanismos de producción fundamentales la perpetuación de los síntomas una vez que habían mejorado y la simulación de los síntomas cuando tenían un beneficio externo. Conviene destacar que, de acuerdo a la opinión de los profesionales evaluados en este estudio, rara vez es una simulación completa de la patología, la simulación del mecanismo de producción o el rol de enfermo. En la terminología clásica de Lipman, (1962) podría decirse que predomina la simulación parcial (exageración o perseveración) frente a la simulación pura o inventada. Esto es un aspecto con ciertas implicaciones a nivel metodológico ya que cuestiona la utilidad del simulador análogo, que debe recrear desde cero una patología, como modelo de investigación en simulación. No parece que la simulación pura sea habitual, sino más bien suele presentarse una simulación parcial realizada por pacientes clínicos con cierto nivel genuino de sintomatología pero que deciden mantener, exacerbar o perdurar en tiempo dichos trastornos.

En relación a las **estrategias de detección** mencionadas por los profesionales, en general concuerdan con las mencionadas en la literatura científica y en otras encuestas realizadas internacionalmente (Dandachi-FitzGerald et al., 2013; Mittenberg et al., 2002; Rogers, 1984; Sharland y Gfeller, 2007). Por ejemplo, el estudio de Rogers (1984) con clínicos mencionaba la inconsistencia con lo observado y la extrema severidad de los síntomas como dos de los criterios fundamentales empleados por los profesionales, o en la referencia clásica de Resnick (2002) describía a los simuladores como tendentes a evitar las evaluaciones o exámenes médicos o psicológicos, así como a ser hoscos, huraños, suspicaces o resentidos.

Resulta interesante ver un posible paralelismo de las estrategias mencionada por los profesionales con la definición y los criterios de sospecha de simulación del DSM-5. Así, el primero de los criterios de sospecha del DSM-5 (presentación en un contexto médico-legal), podría englobar las categorías de “Presencia de beneficios colaterales a la baja: accidente y compensaciones, presencia de conflictos personales, económicos o familiares...” y situación laboral conflictiva, que son citados respectivamente por el 38% y el 28% de los profesionales. El segundo, discrepancia con los datos objetivos, podría considerarse equivalente a la categoría de “Ausencia de evidencias objetivas o signos objetivos: discordancia síntomas y signos, pruebas normales”, que es mencionado por la amplia mayoría de los profesionales, el 68,6% y al de “exageración de los síntomas” nombrado por

un 43%. El tercer criterio del DSM-5, falta de cooperación en la valoración diagnóstica e incumplimiento del tratamiento, es equivalente a la categoría “Ausencia de colaboración o de adherencia al tratamiento” y es mencionado por el 19%. El cuarto criterio, presencia de trastorno antisocial, podría relacionarse con la categoría de “actitud hostil y agresividad”, mencionado por el 16% de los profesionales.

En su conjunto, los criterios del DSM-5 podrían explicar hipotéticamente alrededor del 59% de las estrategias sugeridas por los profesionales, quedando otro 41% ajeno a estos criterios, entre las que se podrían mencionar la ausencia de respuesta al tratamiento, la descripción imprecisa de síntomas o mecanismo de producción, el alargamiento o perpetuación de los síntomas o el énfasis en su deseo de volver a trabajar.

Cabe llamar la atención sobre el hecho de que ninguno de los profesionales evaluados mencione en ningún caso el uso de pruebas, protocolos o procedimientos objetivos o sistematizados como vías de sospecha de simulación o exageración. Se confía la detección de la simulación a métodos más o menos subjetivos como los vinculados al DSM-5 que diversos autores han criticado por el alto número de falsos positivos que pueden ir asociados (p.ej., Berry y Nelson, 2010; Gerson, 2002; Rogers, 2008a). Esta misma tendencia a confiar primordialmente en evidencias subjetivas para el establecimiento de la simulación ha sido encontrada también en distintas encuestas realizadas a profesionales a nivel internacional (Dandachi-FitzGerald et al., 2011; Mittenberg et al., 2002; Sharland y Gfeller, 2007).

Estos resultados, así como los expuestos anteriormente en relación a las patologías y su prevalencia y el perfil sociodemográfico asociado a simulación, encuentran condicionada su generalización a las características particulares de los profesionales (médicos), a la incidentalidad de la muestra y a su relativamente reducido tamaño. Si bien se dispuso de profesionales de muy diversos centros y áreas médicas, con amplia experiencia profesional y de 26 de las 50 provincias españolas, se considera necesaria la replicación de este estudio con muestras seleccionadas aleatoriamente y estratificadas por área médica y centro de trabajo. Así mismo cabe resaltar el hecho de que las estimaciones de prevalencia proceden de las percepciones de los profesionales y no de un registro sistemático de archivos históricos o de aplicación de protocolos objetivos de detección de simulación. Son conocidos los múltiples sesgos que pueden condicionar dicha percepción, tales como la tendencia de los clínicos a sobreestimar la prevalencia de condiciones inusuales o salientes (Wedding y Faust, 1989), los efectos de las probabilidades apriorísticas o los puntos de anclaje (Kahneman y Tversky, 1984) o la tendencia a estimar mayores tasas de prevalencia generales que las específicamente encontradas en la propia práctica profesional (Dandachi-FitzGerald et al., 2013).

Este posible conjunto de sesgos nos advierte de la limitada objetividad de estas prevalencias y de las probables divergencias que puedan hallarse cuando se realicen estudios objetivos con protocolos o historiales. Estas percepciones sobre la prevalencia de la simulación han de considerarse más bien como un acercamiento al punto de vista del profesional y a la relevancia que conceden a una problemática con un fuerte impacto en los recursos económicos de los sistemas sanitarios.

Serán necesarias investigaciones posteriores que repliquen este estudio con muestras más amplias y seleccionadas aleatoriamente, así como estimaciones convergentes de prevalencias procedentes de la revisión de historiales y archivos o de la aplicación de procedimientos sistemáticos y objetivos de detección de simulación. En futuros estudios será oportuno incluir en el listado de patologías algunas de las mencionadas de forma destacada en los campos abiertos del cuestionario (p. ej., omalgia, gonalgia y epicondilitis), así como otras que puedan considerarse relevantes a nivel psicopatológico y que no fueron incluidas aquí. También será deseable que se amplíe el ámbito de la investigación al campo forense. Sin duda los resultados de este estudio están fuertemente restringidos al campo médico-legal y cabe esperar resultados muy diferentes en otros contextos como los peritajes judiciales o el contexto militar. En estos casos, se deberá incluir a otros profesionales que participan habitualmente en la valoración de incapacidades, lesiones o trastornos asociados a simulación. Sería particularmente recomendable extender esta investigación al campo de la neuropsicología y a la estimación por parte de dichos profesionales de tasas de prevalencia en su contexto profesional y en sus áreas de competencia.

4.5. Conclusiones

A pesar de la relevancia de la simulación a nivel económico y profesional en el ámbito médico, apenas existe investigación en España que indague sobre la percepción de los profesionales sobre cuál es su prevalencia, a qué perfil sociodemográfico suelen vincularse y qué evidencias emplean para su detección. Conocer esta información se consideró un primer paso necesario para poder abordar las áreas donde se percibe mayor necesidad de instrumentos para la detección de simulación y complementar las estrategias ya empleadas por los profesionales.

En su conjunto los resultados llaman la atención sobre la magnitud de la presencia de la simulación percibida por los profesionales, cercana en su mediana al 50%, particularmente en trastornos muy frecuentes y prevalentes en la población como son los trastornos por dolor crónico (esguince cervical, lumbalgia y cervicalgia) y los afectivos y emocionales (ansiedad y depresión).

El perfil sociodemográfico de simulación percibido por los profesionales parece encajar dentro del modelo adaptativo propuesto por diversos autores (González-Ordi et al., 2012; Rogers, 2008b), que considera la respuesta de simulación como una elección adaptativa que les reporta el mejor balance de costes-beneficios en un contexto adverso. Así, la mayores tasas de simulación han sido percibidas asociadas a personas de mediana edad (habitualmente con mayores cargas familiares asociadas y más beneficiadas con disponer de una baja laboral), con trabajos poco estimulantes y bajos o moderados ingresos, que comienzan sufriendo originalmente algún tipo de lesión o trastorno y deciden prologarlo en el tiempo para mantener los beneficios obtenidos por dicha baja a nivel económico o laboral. En síntesis podría resumirse en que, de acuerdo a la percepción de los profesionales médicos:

- simulan más aquellos sujetos con peores condiciones económicas y laborales (menor nivel de ingresos, profesiones de menor cualificación y menos vocacionales o satisfactorias, medianas edades), más motivados a considerar alternativas al trabajo para conseguir ingresos o unas mejores condiciones personales;
- simulan aquellas patologías más prevalentes en la población y, por lo tanto, más probable que hayan experimentado y que estén dentro de su universo de repertorio conductual, aquellas de carácter crónico (preferentemente a las leves o agudas) y más subjetivas de verificar: esguince cervical, fibromialgia, lumbalgia crónica, cervicalgia crónica, ansiedad y depresión.
- simulan como una alternativa que les surge de obtener beneficios externos a partir de una patología real que o bien mantienen en el tiempo artificialmente o exageran (simulación parcial, perpetuación de los síntomas y con beneficios externos).

Ante esta acuciante necesidad profesional, llama la atención la carencia de estrategias objetivas puestas en marcha por los profesionales, sin que ninguno de ellos mencione en ningún caso el uso de pruebas, protocolos o procedimientos objetivos o sistematizados como vías de sospecha de simulación o exageración. Las estrategias de detección se articulan mayoritariamente alrededor de la definición y criterios de sospecha del DSM, criterios que han sido duramente criticados por diversos autores por su subjetividad y por el alto número de falsos positivos que pueden ir asociados (v.g., Berry y Nelson, 2010; Gerson, 2002; Rogers, 2008a).

En este contexto se hace importante la investigación en el desarrollo de protocolos y tests específicos de simulación que permitan dar una respuesta científica a un asunto relevante y preocupante. La inclusión de tests y protocolos específicos de evaluación deberían ayudar tanto a facilitar la detección del paciente simulador como a asegurar la determinación de que el paciente

honesto realmente lo es y no es tratado como un simulador (falso positivo). Tanto la confirmación (presencia) como refutación (ausencia) de la existencia de conductas fraudulentas de exageración de síntomas y discapacidad son objetivos igualmente importantes a la hora de implementar un protocolo de evaluación.

5. Utilidad diagnóstica del SIMS en la detección de simulación de psicopatología

5.1. Objetivos e hipótesis de trabajo

El objetivo de esta investigación es analizar la utilidad diagnóstica del SIMS en la detección de simulación de psicopatología en muestras españolas empleando para ello dos estudios, uno con diseño análogo con muestras clínicas de referencia para su comparación y otro con diseño autosuficiente en contexto médico-legal de bajas laborales por psicopatología.

En el **primer estudio** se dispondrá de tres grupos (control, análogo y clínico) y tres condiciones de simulación: simulación de trastorno psicótico para evitar la cárcel o reducir la condena, simulación de secuelas tras un accidente de tráfico para conseguir indemnización económica y simulación de depresión para conseguir baja laboral. Se pretende determinar la capacidad del SIMS para detectar a los participantes instruidos a simular en cada una de estas tres condiciones específicas de psicopatología y en su conjunto.

Específicamente los objetivos perseguidos en este estudio son:

- 1) Analizar la utilidad diagnóstica de la puntuación total del SIMS para detectar la simulación de distintos tipos de psicopatologías simuladas experimentalmente.
- 2) Analizar la utilidad diagnóstica específica de cada una de las escalas del SIMS diseñadas para la detección de la simulación en las tres áreas sintomatológicas determinadas.
- 3) Analizar la eficacia de los puntos de corte propuestos en literatura científica y estudiar su adecuación a la población española con el fin de modificar o sugerir otros que fueran oportunos.

En el **segundo estudio** se pretende poner a prueba la eficacia del SIMS en contextos médico-legales de personas con baja laboral por problemas psicopatológicos clasificados como con presencia o ausencia de simulación por un criterio externo. También pretende analizar su validez convergente con una prueba de referencia en el campo de detección de simulación. Este estudio en contexto real

pretende complementar al estudio de simuladores análogos para determinar la utilidad diagnóstica de la prueba.

En concreto se plantean tres *objetivos* para este segundo estudio:

- 1) Analizar la utilidad diagnóstica de la puntuación total del SIMS y de cada una de sus escalas en la detección de simulación de psicopatología en contextos médico-legales reales.
- 2) Analizar la eficacia de los puntos de corte de la puntuación total y las escalas.
- 3) Estudiar la validez convergente del SIMS con las escalas del MMPI-2 y del MMPI-2-RF.

Estudio 1

5.2. Método

5.2.1. Participantes

Se dispusieron de tres muestras: controles, pacientes clínicos y simuladores análogos. La primera muestra (grupo control) consistió en 174 participantes sanos (sin patología previa diagnosticada), 115 mujeres y 59 varones con una edad media de 27,9 años ($Dt = 12,8$; rango 18 - 62 años), a los que se les pidió que respondieran honestamente el SIMS. De ellos, 113 eran estudiantes universitarios y 61 adultos en situación laboral activa. No se hallaron diferencias significativas entre las puntuaciones en las distintas escalas de SIMS entre el grupo de adultos en situación laboral activa y el grupo de universitarios (MANOVA, λ de Wilk = 0,959, $F(5, 168) = 1,422$, $p = 0,219$, $\eta^2 = 0,04$).

La segunda muestra (grupo de simulación análoga) estuvo compuesta por 159 participantes sanos (sin patología previa diagnosticada), 109 mujeres y 50 varones con una edad media de 22,6 ($Dt = 6,3$; rango 18 – 59 años), a los que se les pidió responder al SIMS simulando un tipo de trastornos o problema diferente con unos determinados incentivos hipotéticos acordes al trastorno concreto. De ellos 137 eran estudiantes universitarios y 22 adultos en situación laboral activa. No se hallaron diferencias significativas entre las puntuaciones en las distintas escalas de SIMS entre ambos grupos, adultos en situación laboral activa y estudiantes universitarios (MANOVA, λ de Wilk = 0,046, $F(5, 153) = 1,485$, $p = 0,198$, $\eta^2 = 0,05$).

La tercera muestra (grupo clínico) estuvo compuesta por 145 participantes clínicos (casos con algún trastorno psicológico o neuropsicológico diagnosticado), 70 mujeres y 75 varones con una edad

media de 43,0 (Dt = 12,2; rango 18 – 71 años), procedentes de centros de salud mental, hospitales, asociaciones de afectados y centros de rehabilitación psicosocial (CRPS) a los que se les pidió que respondieran al SIMS honestamente y que habían sido diagnosticados independientemente por profesionales de dichos centros. Específicamente fueron seleccionados casos cuyos diagnósticos se ajustaban a alguna de las categorías consideradas en los grupos análogos (trastornos psicóticos, deterioro neurológico o trastornos afectivos). Así, un 31% de los casos (n= 45) eran pacientes con trastornos psicóticos, un 23% (n=33) padecían deterioro neurológico y un 46% (n=67) tenían trastornos afectivos.

5.2.2. Material

Los participantes fueron evaluados con la adaptación española del SIMS (González-Ordi y Santamaría, 2009). En la versión española, los ítems del SIMS fueron traducidos y adaptados con el fin de evitar ambigüedades y localismos americanos (uso de dólares, mención al presidente de los Estados Unidos...).

5.2.3. Procedimiento

Los participantes de las muestras 1 (grupo control) y 3 (grupo clínico) fueron instruidos a responder honestamente a los ítems del SIMS, bajo las instrucciones estándar de la prueba (González-Ordi y Santamaría, 2009).

Los participantes de la muestra 2 (grupo simuladores análogos) fueron instruidos a responder en tres condiciones de simulación diferentes asignadas aleatoriamente a cada participante. En la primera condición, se les solicitaba a los participantes que simularan un trastorno psicótico (n=53) ya que habían cometido un delito violento y querían reducir o atenuar su condena y tipo de prisión; en la segunda condición, se les solicitaba que simularan problemas neurológicos consecuencia de un accidente (n=53) con el objetivo de obtener una indemnización económica; en la tercera condición (n=53) se les pedía que simularan un trastorno afectivo con el objetivo de obtener una baja laboral. En cada una de las condiciones, se les presentó por escrito a los participantes un pequeño texto descriptivo de las quejas que tenían que simular y el contexto en que tenían lugar (p. ej., accidente de tráfico y búsqueda de indemnización) y se instruía a la persona para que se presentara de una manera creíble ante el clínico al informar de sus síntomas. En el apéndice B pueden encontrarse las instrucciones específicas dadas a cada uno de los grupos de simulación análoga.

5.3. Resultados

5.3.1. Puntuaciones en las escalas por grupo

Se realizó un MANOVA de un factor con tres grupos para analizar si existían diferencias globales en las escalas del SIMS de acuerdo a la condición establecida (control, simulador análogo, clínico). El modelo global fue estadísticamente significativo (λ de Wilk = 0,337, $F(10, 942) = 67,95$, $p < 0,001$, $\eta^2 = 0,42$). A continuación se realizaron ANOVAs con corrección de Bonferroni y test post hoc (diferencia honestamente significativa de Tukey) de cada una de las escalas individuales del SIMS (Ps, Dn, Am, Bi y Af) y de la puntuación total para valorar la significación de las diferencias entre los tres grupos en cada escala. La magnitud de las diferencias entre los grupos fue analizada con la d de Cohen como medida del tamaño del efecto. De acuerdo a los criterios establecidos por Rogers (2008d) específicamente para la investigación en simulación se consideró los siguientes intervalos para etiquetar la magnitud de la d de Cohen: 0,75 para moderado, 1,25 para grande y 1,50 para muy grande. Los resultados se muestran en la tabla 5.1 y en la figura 5.1. Todas las escalas del SIMS obtuvieron puntuaciones significativamente más altas en el grupo de simuladores análogos frente al grupo clínico de pacientes.

Las estimaciones del tamaño del efecto fueron siempre mayores en la comparación del grupo de control frente a grupo de simuladores análogos que en la comparación del grupo clínico frente a grupo de simuladores análogos, tal y como cabría esperar. La puntuación total en el SIMS tuvo el mayor tamaño del efecto, seguido de la escala Trastornos afectivos (Af). Las escalas Psicosis (Ps), Deterioro neurológico (Dn) y Trastornos amnésicos (Am) tuvieron tamaños del efecto similares, mientras que la escala Baja inteligencia (Bi) fue la que menores tamaños del efecto obtuvo.

Tabla 5.1
Medias, desviaciones típicas, estadístico F y tamaños del efecto para los grupos control, clínico y simuladores análogos

	Controles (n=174)	Clínicos (n=145)	Simuladores análogos (n=159)	F	d_1	d_2
Total SIMS	7,38 ^a (3,13)	18,52 ^b (9,72)	32,5 ^c (11,7)	337,937***	3,00	1,30
Psicosis (Ps)	0,43 ^a (0,72)	1,84 ^b (2,16)	5,19 ^c (4,5)	117,76***	1,52	0,94
Deterioro neurológico (Dn)	0,71 ^a (1,01)	4,12 ^b (3,13)	6,97 ^c (3,1)	251,34***	2,78	0,92
Trastornos amnésicos (Am)	0,74 ^a (1,1)	3,54 ^b (3,24)	7,01 ^c (4,3)	167,22***	2,04	0,91
Baja inteligencia (Bi)	1,75 ^a (1,31)	2,61 ^b (1,57)	3,86 ^c (2,45)	54,69***	1,09	0,60
Trastornos afectivos (Af)	3,75 ^a (1,69)	6,41 ^b (2,63)	9,47 ^c (2,18)	287,65***	2,96	1,28

Nota. Las medias marcadas con diferente letra son significativamente diferentes a un nivel de $p < 0,05$
 d_1 tamaño del efecto para la comparación del grupo control frente al grupo de simuladores análogos,

d_2 tamaño del efecto para la comparación del grupo clínico frente al grupo de simuladores análogos

*** $p < 0,001$

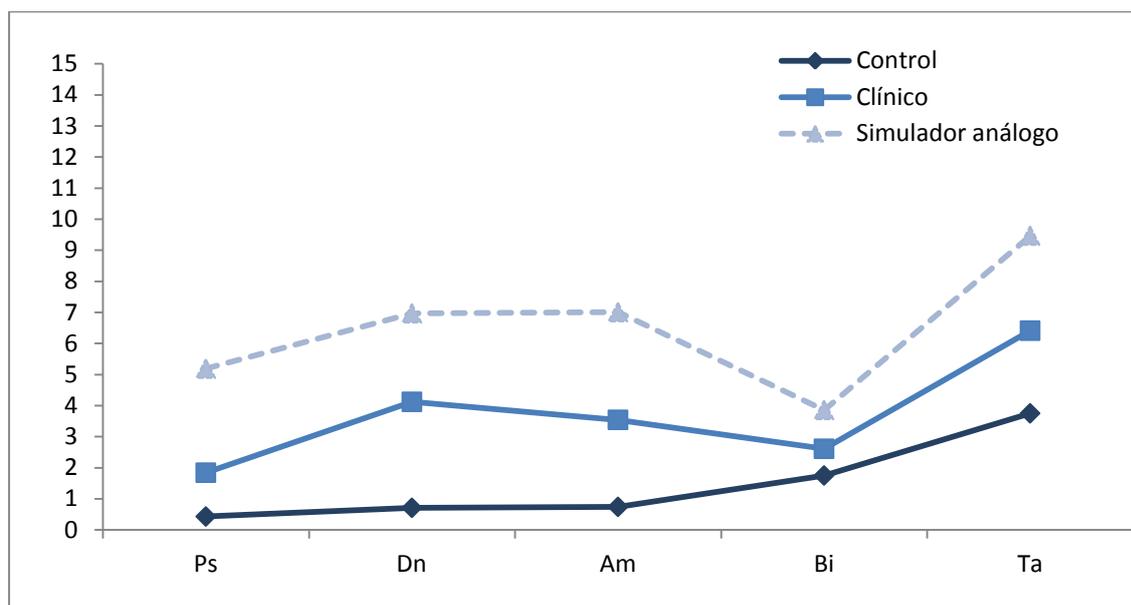


Figura 5.1. Puntuaciones medias para el grupo control, clínico y de simulador análogos en las escalas del SIMS

A continuación, se analizó si las diferencias en las escalas del SIMS se mantenían en cada una de las tres condiciones del grupo de simuladores análogos, comparándolo con sus homólogos del grupo clínico (aquellos casos clínicos asignados a esa categoría). Para ello se realizó un MANOVA para cada condición (trastornos psicóticos, deterioro neurológico y trastornos afectivos) considerando a las escalas del SIMS como variables dependientes y al grupo (clínico, simulador análogo) como factor. El modelo global fue estadísticamente significativo para la condición de trastorno psicótico (Wilk's $\lambda=0,302$ $F(5,92)=42,61$, $p < 0,001$, $\eta^2 = 0,70$), para la condición de deterioro neurológico (Wilk's $\lambda=0,578$, $F(5,80)=11,69$, $p < 0,001$, $\eta^2 = 0,42$) y para la condición de trastorno afectivo (Wilk's $\lambda=0,687$, $F(5,114)=10,37$, $p < 0,001$, $\eta^2 = 0,31$).

Se complementó el análisis anterior, con una serie de ANOVAs con corrección de Bonferroni y pruebas post hoc (Tukey HDS) para determinar entre qué grupos se establecían diferencias significativas para cada una de las escalas en cada una de las condiciones. Para estudiar el tamaño del efecto se calculó la d de Cohen entre cada par de grupos en cada condición. Los resultados pueden observarse en la tabla 5.2 y en las figuras 5.2 a 5.4.

En conjunto, el grupo de simuladores análogos obtuvo puntuaciones significativamente más altas que el grupo clínico de pacientes. Los mayores valores de la d de Cohen se observaron en la condición de trastorno psicótico, con clara diferencia respecto a las condiciones de deterioro

neuroológico y trastornos afectivos. En las tres condiciones, la puntuación Total se encontraba de forma consistente entre las dos escalas que mayor d de Cohen obtenían. En el caso de Psicosis y Trastornos afectivos eran la escala total y la propia escala de la condición (Ps y Af respectivamente) las dos que mayor poder discriminativo tenían; en la condición de Deterioro neurológico eran Total y Trastornos afectivos (Af), seguido en tercer lugar por la escala Deterioro neurológico (Dn).

Tabla 5.2

Medias, desviaciones típicas, estadístico F y tamaños del efecto para los grupos clínico y simulador análogo por condición

Condición / categoría diagnóstica	Escala del SIMS	Clínicos	Simuladores Análogos	F	d
Trastorno psicótico		(n=45)	(n=53)		
	Total SIMS	16,90 (9,56)	40,60 (10,20)	139,06***	2,42
	Psicosis (Ps)	2,29 (2,25)	9,21 (3,50)	129,74***	2,34
	Deterioro neurológico (Dn)	3,69 (3,11)	7,64 (2,97)	41,21***	1,31
	Trastornos amnésicos (Am)	2,80 (2,47)	9,94 (3,95)	110,38***	2,15
	Baja inteligencia (Bi)	2,73 (1,75)	4,66 (2,91)	15,10***	0,80
Deterioro neurológico		(n=33)	(n=53)		
	Total SIMS	16,60 (8,39)	26,10 (10,40)	19,474***	0,99
	Psicosis (Ps)	1,03 (1,85)	2,72 (3,62)	6,14*	0,56
	Deterioro neurológico (Dn)	3,82 (2,63)	6,02 (2,95)	12,28**	0,79
	Trastornos amnésicos (Am)	3,88 (3,43)	4,91 (3,73)	1,63	0,29
	Baja inteligencia (Bi)	2,06 (1,34)	3,60 (2,36)	11,74**	0,77
Trastorno afectivo		(n=67)	(n=53)		
	Total SIMS	20,50 (10,20)	30,80 (9,56)	31,62***	1,05
	Psicosis (Ps)	1,94 (2,15)	3,66 (3,31)	11,80**	0,64
	Deterioro neurológico (Dn)	4,55 (3,34)	7,26 (3,19)	20,30***	0,83
	Trastornos amnésicos (Am)	3,88 (3,54)	6,17 (3,57)	12,29**	0,65
	Baja inteligencia (Bi)	2,79 (1,51)	3,30 (1,79)	2,86	0,31
	Trastornos afectivos (Af)	7,34 (2,72)	10,4 (1,76)	48,89***	1,32

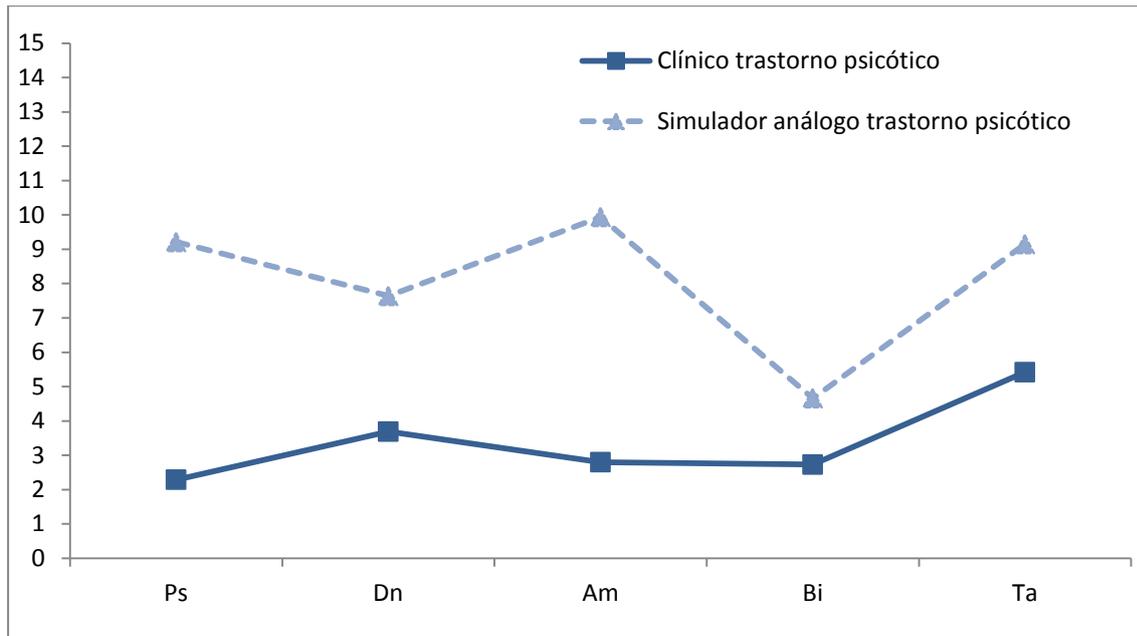


Figura 5.2. Medias para la condición de trastorno psicótico: grupo clínico y grupo de simuladores análogos.

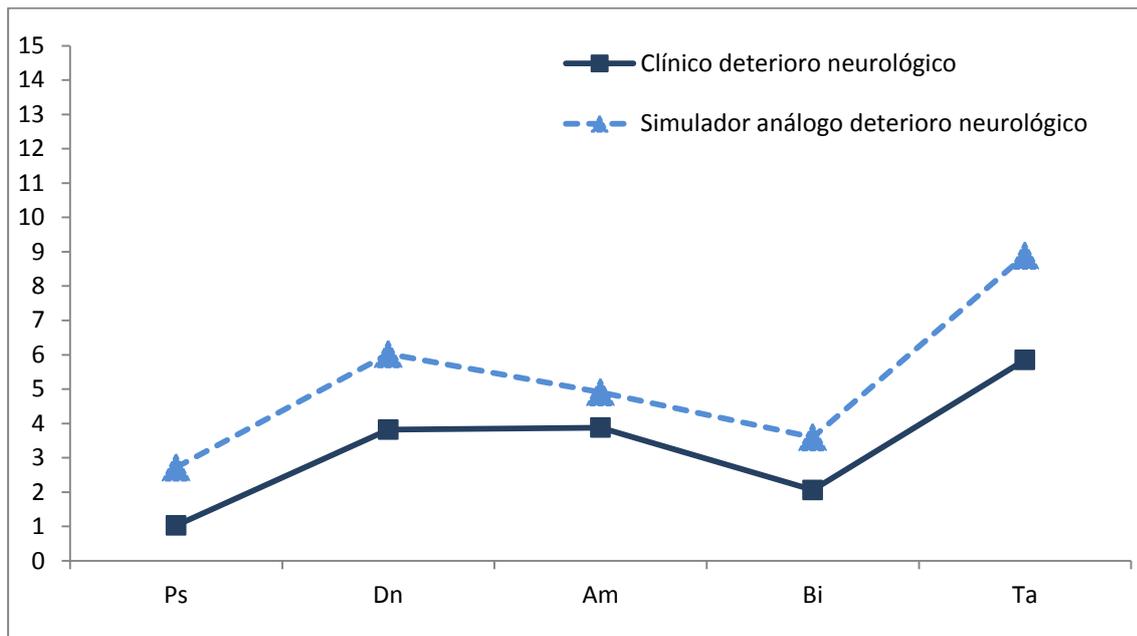


Figura 5.3. Medias para la condición de deterioro neurológico: grupo clínico y grupo de simuladores análogos

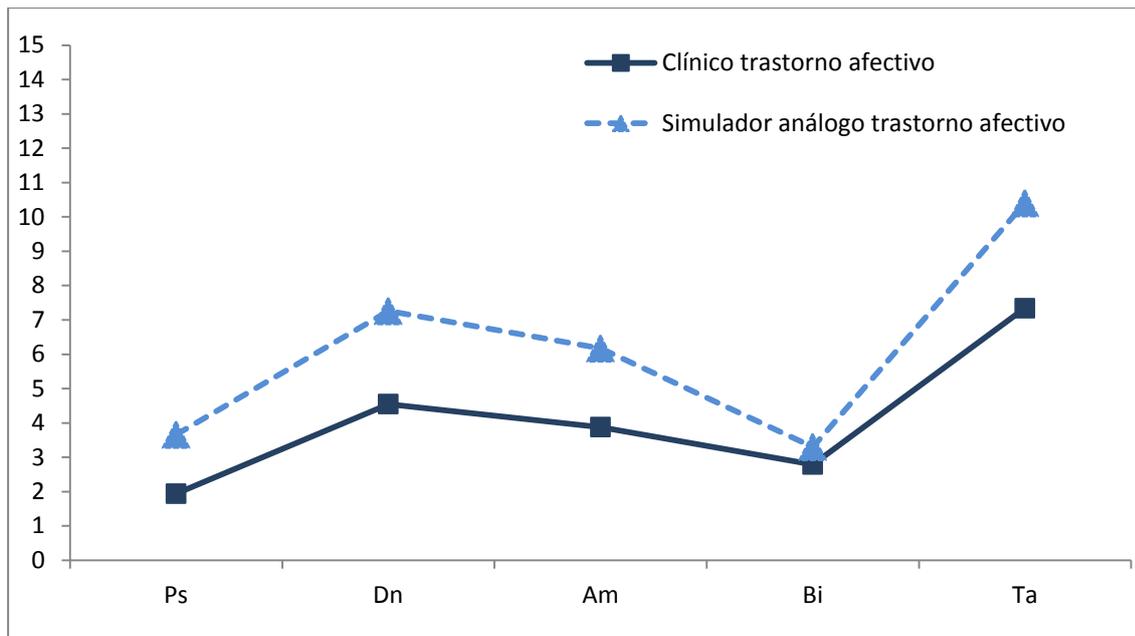


Figura 5.4. Medias para la condición de trastornos afectivos: grupo clínico y grupo de simuladores análogos

5.3.2. Utilidad diagnóstica

Se estudió la utilidad diagnóstica de los distintos puntos de corte para la puntuación total del SIMS en la discriminación entre el grupo clínico y el grupo de simulación análoga (excluyendo al grupo control de estos estudios). Se consideró un amplio rango de posibles puntos de corte sugeridos por investigaciones previas.

Se emplearon curvas COR para examinar la precisión clasificatoria de las escalas del SIMS a la hora de diferenciar el grupo de simuladores análogos del grupo clínico de pacientes. Las curvas COR se calculan como función de la sensibilidad y 1-especificidad y su rendimiento global puede ser evaluado con el área bajo la curva (ABC). Adicionalmente, se calcularon el poder predictivo positivo y el poder predictivo negativo (PPP y PPN) estimados de acuerdo a diferentes tasas base de prevalencia (0,15, 0,30 y 0,50) tal y como es recomendado en la literatura (Sellbom et al., 2010). El área bajo la curva de la puntuación total del SIMS fue de 0,82 (ET=0,024). En la tabla 5.3 se muestran los resultados tanto para el grupo global, sin diferenciar la condición; en la tabla 5.4 se muestran los resultados para cada una de las condiciones consideradas en el estudio.

En las tablas 5.5 a 5.7 se analizaron específicamente los puntos de corte por escala del SIMS en cada una de las condiciones. Se consideró en cada condición aquellas escalas que mayor tamaño del efecto habían mostrado. Así, en la condición de trastorno psicótico se estudiaron las escalas Psicosis y Trastornos amnésicos (con áreas bajo la curva respectivamente de 0,93 y 0,92); en la condición de deterioro neurológico las de Trastornos afectivos, Deterioro neurológico y Baja inteligencia (con áreas bajo la curva respectivamente de 0,82, 0,72 y 0,69) y en la condición de trastornos afectivos, la escala del mismo nombre (área bajo la curva igual a 0,81).

Tabla 5.3

Estadísticos de utilidad diagnóstica de la puntuación total en el SIMS en la discriminación entre pacientes clínicos y simuladores análogos

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
> 14	0,95	0,40	0,69	0,22	0,98	0,40	0,95	0,61	0,89
> 16	0,91	0,48	0,71	0,24	0,97	0,43	0,93	0,64	0,85
> 19	0,84	0,59	0,72	0,26	0,95	0,47	0,90	0,67	0,79
> 23	0,76	0,68	0,72	0,30	0,94	0,51	0,87	0,71	0,74
> 24	0,75	0,71	0,73	0,31	0,94	0,53	0,87	0,72	0,74
> 26	0,68	0,81	0,74	0,38	0,93	0,60	0,85	0,78	0,72
> 32	0,48	0,90	0,68	0,47	0,91	0,68	0,80	0,83	0,63
> 40	0,28	0,99	0,62	0,78	0,89	0,90	0,76	0,95	0,58

Nota. Sen. = sensibilidad, Esp. = especificidad, %GI = proporción global de clasificaciones correctas, TP = Tasa de prevalencia, PPP = poder predictivo positivo, PPN = poder predictivo negativo.

Tabla 5.4

Estadísticos de utilidad diagnóstica de la puntuación total en el SIMS en la discriminación entre pacientes clínicos y simuladores análogos por condición

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Grupo de simuladores análogos de trastorno psicótico frente a pacientes con diagnósticos dentro de la categoría de trastorno psicótico									
> 14	0,98	0,51	0,77	0,26	0,99	0,46	0,98	0,67	0,96
> 16	0,98	0,56	0,79	0,28	0,99	0,49	0,99	0,69	0,97
> 19	0,96	0,62	0,81	0,31	0,99	0,52	0,97	0,72	0,94
> 23	0,96	0,69	0,84	0,35	0,99	0,57	0,98	0,76	0,95
> 24	0,96	0,76	0,87	0,41	0,99	0,63	0,98	0,80	0,95
> 26	0,96	0,87	0,92	0,56	0,99	0,76	0,98	0,88	0,96
> 32	0,77	0,93	0,85	0,67	0,96	0,83	0,91	0,92	0,80
> 40	0,55	1,00	0,76	1,00	0,93	1,00	0,84	1,00	0,69
Grupo de simuladores análogos de deterioro neurológico tras accidente de tráfico frente a pacientes con diagnósticos dentro de la categoría de deterioro neurológico									
> 14	0,91	0,42	0,72	0,22	0,96	0,40	0,91	0,61	0,82
> 16	0,81	0,64	0,74	0,28	0,95	0,49	0,89	0,69	0,77
> 19	0,68	0,73	0,70	0,31	0,93	0,52	0,84	0,71	0,69
> 23	0,55	0,79	0,64	0,31	0,91	0,53	0,80	0,72	0,64
> 24	0,55	0,82	0,65	0,35	0,91	0,56	0,81	0,75	0,64
> 26	0,42	0,88	0,59	0,38	0,89	0,59	0,78	0,77	0,60
> 32	0,23	0,91	0,49	0,31	0,87	0,52	0,73	0,71	0,54
> 40	0,11	1,00	0,45	1,00	0,86	1,00	0,72	1,00	0,53
Grupo de simuladores análogos de depresión para obtener baja laboral frente a pacientes con diagnósticos dentro de la categoría de trastornos afectivos									
> 14	0,96	0,31	0,60	0,20	0,98	0,38	0,95	0,58	0,89
> 16	0,94	0,36	0,62	0,21	0,97	0,39	0,94	0,60	0,86
> 19	0,87	0,52	0,68	0,24	0,96	0,44	0,90	0,65	0,80
> 23	0,77	0,63	0,69	0,27	0,94	0,47	0,87	0,67	0,73
> 24	0,74	0,63	0,68	0,26	0,93	0,46	0,85	0,66	0,70
> 26	0,66	0,73	0,70	0,30	0,92	0,51	0,83	0,71	0,68
> 32	0,40	0,88	0,67	0,37	0,89	0,59	0,77	0,77	0,59
> 40	0,15	0,97	0,61	0,47	0,87	0,68	0,73	0,83	0,53

Nota. Sen. = sensibilidad, Esp. = especificidad, %GI = porcentaje global de clasificaciones correctas, TP = Tasa de prevalencia, PPP = poder predictivo positivo, PPN = poder predictivo negativo.

Tabla 5.5

Estadísticos de utilidad diagnóstica de las escalas del SIMS en la discriminación entre pacientes clínicos y simuladores análogos en la condición de trastorno psicótico

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Escala Psicosis (Ps)									
Ps > 0	0,98	0,18	0,61	0,17	0,98	0,34	0,96	0,54	0,90
Ps > 1	0,94	0,51	0,74	0,25	0,98	0,45	0,95	0,66	0,90
Ps > 2	0,92	0,71	0,83	0,36	0,98	0,58	0,96	0,76	0,90
Ps > 3	0,91	0,76	0,84	0,40	0,98	0,61	0,95	0,79	0,89
Ps > 4	0,91	0,80	0,86	0,44	0,98	0,66	0,95	0,82	0,89
Ps > 5	0,89	0,87	0,88	0,54	0,98	0,74	0,95	0,87	0,88
Ps > 6	0,83	0,91	0,87	0,62	0,97	0,80	0,93	0,90	0,84
Ps > 7	0,77	0,98	0,87	0,86	0,96	0,94	0,91	0,97	0,81
Ps > 8	0,60	1,00	0,79	1,00	0,93	1,00	0,85	1,00	0,72
Escala Trastornos amnésicos (Am)									
Am > 0	1,00	0,22	0,64	0,18	1,00	0,36	1,00	0,56	1,00
Am > 1	0,98	0,38	0,70	0,22	0,99	0,40	0,98	0,61	0,95
Am > 2	0,94	0,51	0,74	0,25	0,98	0,45	0,95	0,66	0,90
Am > 3	0,89	0,64	0,78	0,31	0,97	0,52	0,93	0,71	0,85
Am > 4	0,89	0,76	0,83	0,39	0,97	0,61	0,94	0,78	0,87
Am > 5	0,85	0,87	0,86	0,53	0,97	0,73	0,93	0,86	0,85
Am > 6	0,81	0,89	0,85	0,56	0,96	0,76	0,92	0,88	0,82
Am > 7	0,74	0,98	0,85	0,85	0,95	0,93	0,90	0,97	0,79
Am > 8	0,70	0,98	0,83	0,85	0,95	0,93	0,88	0,97	0,76
Am > 9	0,58	0,98	0,77	0,82	0,93	0,92	0,85	0,96	0,70
Am > 10	0,45	1,00	0,70	1,00	0,91	1,00	0,81	1,00	0,65

Nota. Sen. = sensibilidad, Esp. = especificidad, %GI = porcentaje global de clasificaciones correctas, TP = Tasa de prevalencia, PPP = poder predictivo positivo, PPN = poder predictivo negativo.

Tabla 5.6

Estadísticos de utilidad diagnóstica de las escalas del SIMS en la discriminación entre pacientes clínicos y simuladores análogos en la condición de deterioro neurológico

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Escala Trastornos afectivos (Af)									
Ta > 3	1,00	0,21	0,70	0,18	1,00	0,35	1,00	0,56	1,00
Ta > 4	0,96	0,33	0,72	0,20	0,98	0,38	0,95	0,59	0,90
Ta > 5	0,91	0,42	0,72	0,22	0,96	0,40	0,91	0,61	0,82
Ta > 6	0,81	0,55	0,71	0,24	0,94	0,43	0,87	0,64	0,74
Ta > 7	0,70	0,76	0,72	0,34	0,93	0,55	0,85	0,74	0,72
Ta > 8	0,62	0,94	0,74	0,64	0,93	0,81	0,85	0,91	0,71
Ta > 9	0,42	0,94	0,62	0,55	0,90	0,75	0,79	0,87	0,62
Ta > 10	0,26	1,00	0,55	1,00	0,89	1,00	0,76	1,00	0,58
Escala Deterioro neurológico (Dn)									
Dn > 0	0,98	0,06	0,63	0,16	0,95	0,31	0,88	0,51	0,76
Dn > 1	0,96	0,21	0,67	0,18	0,97	0,34	0,93	0,55	0,85
Dn > 2	0,91	0,30	0,67	0,19	0,95	0,36	0,88	0,57	0,76
Dn > 3	0,81	0,52	0,70	0,23	0,94	0,42	0,86	0,63	0,73
Dn > 4	0,66	0,73	0,69	0,30	0,92	0,51	0,83	0,71	0,68
Dn > 5	0,49	0,79	0,60	0,29	0,90	0,50	0,78	0,70	0,61
Dn > 6	0,42	0,82	0,57	0,29	0,89	0,49	0,77	0,70	0,58
Dn > 7	0,34	0,88	0,55	0,33	0,88	0,55	0,76	0,74	0,57
Dn > 8	0,17	0,91	0,45	0,25	0,86	0,44	0,72	0,65	0,52
Dn > 9	0,11	0,97	0,44	0,40	0,86	0,62	0,72	0,79	0,52
Dn > 10	0,09	1,00	0,44	1,00	0,86	1,00	0,72	1,00	0,52
Escala Baja inteligencia (Bi)									
Bi > 0	0,94	0,12	0,63	0,16	0,92	0,32	0,83	0,52	0,68
Bi > 1	0,83	0,39	0,66	0,19	0,93	0,37	0,84	0,58	0,70
Bi > 2	0,64	0,61	0,63	0,22	0,91	0,41	0,80	0,62	0,63
Bi > 3	0,43	0,85	0,59	0,34	0,89	0,55	0,78	0,74	0,60
Bi > 4	0,28	0,97	0,55	0,62	0,88	0,80	0,76	0,90	0,57
Bi > 5	0,19	1,00	0,50	1,00	0,87	1,00	0,74	1,00	0,55

Nota. Sen. = sensibilidad, Esp. = especificidad, %GI = porcentaje global de clasificaciones correctas, TP = Tasa de prevalencia, PPP = poder predictivo positivo, PPN = poder predictivo negativo.

Tabla 5.7

Estadísticos de utilidad diagnóstica de las escalas del SIMS en la discriminación entre pacientes clínicos y simuladores análogos en la condición de trastorno afectivo

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Escala Trastornos afectivos (Af)									
Ta > 2	1,00	0,04	0,47	0,16	1,00	0,31	1,00	0,51	1,00
Ta > 3	1,00	0,07	0,48	0,16	1,00	0,32	1,00	0,52	1,00
Ta > 4	1,00	0,16	0,53	0,17	1,00	0,34	1,00	0,54	1,00
Ta > 5	1,00	0,31	0,62	0,20	1,00	0,38	1,00	0,59	1,00
Ta > 6	0,96	0,40	0,65	0,22	0,98	0,41	0,96	0,62	0,91
Ta > 7	0,94	0,49	0,69	0,25	0,98	0,44	0,95	0,65	0,90
Ta > 8	0,85	0,60	0,71	0,27	0,96	0,47	0,90	0,68	0,80
Ta > 9	0,72	0,73	0,73	0,32	0,94	0,53	0,86	0,73	0,72
Ta > 10	0,47	0,88	0,70	0,41	0,90	0,63	0,80	0,80	0,63
Ta > 11	0,34	0,97	0,69	0,67	0,89	0,83	0,77	0,92	0,59
Ta > 12	0,08	0,99	0,58	0,47	0,86	0,68	0,71	0,83	0,52

Nota. Sen. = sensibilidad, Esp. = especificidad, %GI = porcentaje global de clasificaciones correctas, TP = Tasa de prevalencia, PPP = poder predictivo positivo, PPN = poder predictivo negativo.

5.4. Discusión

La capacidad del SIMS para realizar un *screening* amplio de las distintas áreas donde puede presentarse simulación es una de las aportaciones más relevantes de la prueba (Wisdom et al., 2010). Esto puede resultar un valor añadido al contexto profesional español donde, si bien existen pruebas parcialmente validadas para la evaluación de simulación de áreas específicas, no existe ninguna herramienta breve de despistaje de simulación que derive posteriormente a protocolos más completos y especializados.

En este primer estudio análogo se pretendió explorar precisamente esa capacidad para detectar simulación inducida experimentalmente en un conjunto amplio y diverso de áreas, en concreto en tres grandes conjuntos: trastornos psicóticos, deterioro neurológico y trastornos afectivos. Se pretendía analizar la capacidad discriminativa de la puntuación total del SIMS y sus escalas tanto en el conjunto global de la muestra, sin distinguir el área específica simulada, como en cada uno de los grupos. Esto permitiría obtener unos primeros datos preliminares sobre su eficacia en un grupo indeterminado y variado de simulación y posteriormente específicamente en grupos más acotados.

Los resultados han mostrado una significativa y satisfactoria capacidad discriminativa de la puntuación total del SIMS entre pacientes clínicos y simuladores análogos tanto en el conjunto global de la muestra como en cada una de las condiciones específicas. Para el conjunto global de la muestra

ha mostrado un tamaño del efecto grande, mostrando una adecuada capacidad para discernir la sospecha de simulación en un grupo indeterminado y variado de simulación y clínicos (d cohen para la puntuación total = 1,30). Cuando se ha pasado a analizar cada una de las condiciones, la puntuación total del SIMS ha mostrado un tamaño del efecto muy grande en la condición de trastorno psicótico, grande en trastorno afectivo y moderado en deterioro neurológico.

Este gradiente descendiente de tamaño del efecto conforme menor es el componente psicopatológico implicado (psicopatología grave, psicopatología moderada, baja psicopatología) concuerda con los obtenidos por otros autores que han señalado que conforme más grave es la psicopatología alegada, mayor es el patrón de simulación y mayores tamaños del efecto se encuentran (Greene, 2008). Así mismo diversos autores han encontrado patrones de exageraciones más amplios en contextos forenses criminales que en contextos civiles (Wygant et al., 2007, 2010), lo que concordaría con la mayor magnitud de la condición de trastornos psicóticos, en la que se colocaba a los simuladores análogos en un contexto criminal, que las condiciones de trastorno afectivo y deterioro neurológico, en las que se les situaba en un contexto civil.

Específicamente, las magnitudes de tamaño del efecto halladas en este estudio son convergentes con las encontradas en otros estudios de simulación análogas realizados con el SIMS, que hallaban valores de la d de Cohen de entre 2,19 y 2,42 en el caso de trastornos psicóticos (Cima, Hollnack, et al., 2003; Merckelbach y Smith, 2003), entre 0,63 y 1,96 en el caso de Deterioro neurológico (Cima, Hollnack, et al., 2003; Merckelbach y Smith, 2003) y entre 1,00 y 1,50 en el caso de Trastornos afectivos (Cima, Hollnack, et al., 2003; Dearth, 2007; Efendov, 2007).

Cuando se analizan los resultados de cada escala en cada condición, se advierte que la escala Psicosis y la escala Trastornos afectivos son las que mayor tamaño del efecto obtienen respectivamente en sus condiciones, ambas con tamaños del efecto grandes. Esto avala su objetivo específico de detectar simulación vinculada a dicha área sintomatológica y cumple con su cometido, una vez detectada simulación a partir de la puntuación Total, de orientar en aquellas áreas que puedan ser las principales simuladas. Es reseñable también que en la condición de Trastorno psicótico la segunda escala con mayor tamaño del efecto es Trastornos amnésicos, de forma coherente con la instrucción de simular que han perdido la memoria y cualquier recuerdo relevante.

En el caso de la condición de Deterioro neurológico, se observa un perfil diferente ya que la escala que más se eleva es Trastornos afectivos, seguido de Deterioro neurológico y Baja inteligencia. Esto probablemente es debido a las instrucciones específicas dadas a este grupo, menos centradas en el deterioro neurológico *per se* y más en el hecho de haber padecido un accidente y querer

obtener indemnización por los daños psicológicos, mentales y físicos sufridos. Una posible hipótesis es que los resultados obtenidos reflejen el perfil de síntomas que los simuladores análogos del estudio consideraban vinculados a un accidente, con un marcado componente emocional seguido de efectos neurológicos y cognitivos. Para desarrollar una condición relativa a deterioro neurológico probablemente hubiera sido más deseable dar instrucciones más circunscritas a estos problemas y menos genéricas.

De todas las escalas, Trastornos afectivos es la que mayores tamaños del efecto muestra de forma consistente en todas las condiciones. Una posible hipótesis explicativa es que frecuentemente los simuladores consideran que en cualquier patología existe un componente asociado de problemas emocionales y malestar que buscan simular también, provocando la elevación en esta escala.

En su conjunto, los resultados obtenidos a nivel de las escalas del SIMS respaldan su papel de detección de simulación específica al área evaluada y su relevancia en la interpretación de las puntuaciones obtenidas en el perfil de la prueba en su adaptación española.

En relación a los puntos de corte de la puntuación Total, los valores recomendados en el manual original americano (>14, Widows y Smith, 2005) han mostrado una excelente sensibilidad pero una muy baja especificidad, algo ya señalado de forma sistemática en la literatura científica que ha investigado este aspecto del SIMS (Alwes, 2006; Cima, Hollnack, et al., 2003; Clegg et al., 2009; Dandachi-FitzGerald et al., 2011; Edens et al., 1999, 2007; Lewis et al., 2002; Merckelbach y Smith, 2003; Peters, Jellicic, et al., 2013; Rogers, Hinds, et al., 1996; Vitacco et al., 2006; Wisdom et al., 2010).

De acuerdo a estos resultados, podrían resultar más recomendables puntos de corte más altos que permitan una mayor especificidad, aunque se mantenga su objetivo de prueba de cribaje con tasas elevadas de sensibilidad. Los puntos de corte recomendados en otras investigaciones como 16 (Rogers, Hinds, et al., 1996; Merckelbach y Smith, 2003; Lewis et al., 2002; Cima, Hollnack, et al., 2003), 19 (Clegg et al., 2009) o 23 (Wisdom et al., 2010), obtienen valores más satisfactorios y parecieran alternativas más recomendables en población española de acuerdo a este estudio.

En concreto, los resultados de este estudio de detección de simulación de psicopatología con la adaptación española del SIMS descartarían preliminarmente el uso del punto de 14. Alternativamente sugerirían tentativamente considerar las puntuaciones por debajo de 16 como indicativas de ausencia de exageración de síntomas psicopatológicos (≤ 16 "Paciente clínico, sin

sospecha de exageración de síntomas de psicopatología”) y puntuaciones por encima de 16 como indicadores de sospecha de exageración de síntomas de psicopatología en diferentes niveles: >16 posible sospecha de exageración de síntomas; >23 probable sospecha de exageración de síntomas; >40 clara sospecha de exageración de síntomas, empleando los niveles de gradientes de certeza sugeridos por Berry y Nelson (2010).

Si bien los resultados hallados en este estudio también han mostrado oscilaciones en la sensibilidad y especificidad de los puntos de corte según la sintomatología alegada, no se ha considerado oportuno proponer puntos de corte alternativos de acuerdo al contexto de evaluación y los síntomas presentados. La ausencia de investigación previa en esta dirección y la incertidumbre al determinar en la práctica qué sintomatología alegará una persona en su evaluación, hacen aconsejable esta decisión. No obstante, se considera este aspecto una línea de investigación futura prioritaria como ya señalaba Alwes (2006) en su estudio con el SIMS.

El análisis de los resultados de utilidad diagnóstica de las escalas, arroja resultados similares a los comentados previamente para la puntuación total. Los puntos de corte sugeridos originalmente por los autores originales (>1 en Ps, >2 en Dn, Am y Bi y >5 en Ta) muestran una alta sensibilidad pero una muy baja especificidad. Este hecho ya ha sido observado también en las escasas investigaciones con muestras clínicas que han indagado en este aspecto tanto en contextos civiles como penales (Clegg et al., 2009; Vitacco et al., 2006). Los resultados obtenidos en este primer estudio, sugieren preliminarmente un aumento sensible de los puntos de corte para su uso en población española. Serán necesarias investigaciones futuras que determinen exactamente los valores más recomendables, si bien tentativamente se podrían sugerir el punto de corte de >2 para la escala Psicosis, de >3 para las escalas Deterioro neurológico, Trastornos Amnésicos y Baja inteligencia y >7 para la escala Trastornos afectivos. Estos puntos de corte permitirían una ligera mejora de la especificidad, aun manteniendo su condición de puntos de cribaje que requieren la posterior confirmación por una batería más completa y exhaustiva. En algunos contextos podría ser recomendable incluso considerar puntos de corte más elevados si se desea aumentar la especificidad.

Dentro de las limitaciones de este estudio cabría señalar varias. En primer lugar la ausencia de un tiempo amplio de preparación previa para los participantes análogos así como incentivos positivos y negativos como recompensa o castigo al éxito en la simulación. Si bien las investigaciones realizadas hasta la fecha han mostrado un menor efecto de estos aspectos con relación al factor clave el remarcar la necesidad de ser creíbles (Elhai et al., 2007; Viglione et al., 2001), aspecto que sí se incluye en las instrucciones suministradas a los participantes en este estudio, no es menos cierto

que el haber introducido ambos aspectos hubiera permitido hacer más vívidas y reales las condiciones de simulación análoga. También hubiera sido deseable considerar diferentes tipos de instrucciones (información básica, información detallada de los síntomas, información de los síntomas y de la existencia de escalas de detección y su funcionamiento...) y haber indagado en el efecto del tipo de preparación en las puntuaciones obtenidas. Será necesario delimitar en futuros estudios con la versión española del SIMS el efecto de estas variables, más aún cuando recientes estudios apuntan a la interacción entre el tipo de entrenamiento y la sintomatología simulada en el efecto en las puntuaciones obtenidas y la detección (Veltri y Williams, 2013). Igualmente será deseable incluir en futuras investigaciones análogas con el SIMS nuevos contextos con diferentes descripciones de los mismos analizando su efecto en las puntuaciones obtenidas. Los contextos y patologías incluidas en este primer estudio son relativamente reducidos y una muestra del posible conjunto más amplio de contextos posibles y dignos de ser investigados. También requerirá de investigaciones futuras la consideración de variables de personalidad y capacidad cognitiva de las personas evaluadas en su capacidad de simular síntomas, como una variable más a considerar dentro del análisis de los resultados obtenidos.

No obstante las limitaciones señaladas, los resultados obtenidos en este estudio muestran en su conjunto una satisfactoria capacidad de la puntuación total del SIMS para discriminar a los simuladores análogos de psicopatología de forma tanto indiferenciada (en el grupo global, sin distinguir condición) como en cada condición. Adicionalmente, las escalas del SIMS han mostrado una adecuada elevación concurrente con la simulación pretendida, lo que respalda su carácter particular asociado a determinadas áreas de sintomatología. Estos resultados apuntan a que la elevación en la puntuación total será indicativa de sospecha de exageración de síntomas mientras que el perfil en las escalas podrá orientar sobre las áreas donde se centra dicha simulación. En síntesis, los resultados hallados aportan evidencia preliminar sobre la bondad del uso de la adaptación española del SIMS en el cribaje de simulación de psicopatología.

Estudio 2

5.5. Método

5.5.1. Participantes

La muestra estuvo compuesta por 207 pacientes, 131 mujeres y 76 varones con una edad media de 38,0 (Dt=10,5; rango 18 - 63 años), que estaban de baja laboral por problemas psicológicos y eran evaluados por psicólogos forenses en su valoración inicial de la patología subyacente.

5.5.2. Materiales

Todos los participantes completaron la versión española del SIMS (González-Ordi y Santamaría, 2009) y del MMPI-2 (Butcher et al., 1989).

El MMPI-2 (Butcher et al., 1989) es uno de los cuestionarios más ampliamente utilizados para evaluar alteraciones psicopatológicas (Archer et al., 2006; Lally, 2003; Muñiz Fernández y Fernández Hermida, 2010) y con mayor evidencias de sus escalas de exageración de síntomas para la detección de simulación (Greene, 2008; Sánchez Crespo, 2002; Santamaría y González-Ordi, 2012).

A partir de las respuestas al MMPI-2 también se pudieron calcular las puntuaciones en el MMPI-2-RF (Ben-Porath y Tellegen, 2008), dado que los 338 ítems del MMPI-2-RF están contenidos dentro del MMPI-2. La investigación previa sobre la prueba ha establecido la equivalencia de las puntuaciones obtenidas en el MMPI-2-RF aplicando tanto el MMPI-2 como sólo los ítems del MMPI-2-RF (Tellegen y Ben-Porath, 2008; Van Der Heijden, Egger y Derksen, 2010). El MMPI-2-RF ha sido desarrollado con el objetivo de proporcionar una evaluación alternativa al MMPI-2 que fuera más breve y psicométricamente más eficiente. Dispone de cinco escalas de exageración de síntomas (F-r, Fp-r, Fs, FBS-r y RBS) que han mostrado su capacidad discriminativa de simulación en distintos contextos (v.g., Rogers, Gillard, et al., 2011; Schroeder et al., 2011; Sellbom et al., 2010, 2012; Wygant et al., 2009; Youngjohn et al., 2011).

5.5.3. Procedimiento

Todos los participantes dieron su consentimiento informado previo a la evaluación y respondieron a las pruebas bajo instrucciones estándar, esto es, respuesta honesta, de acuerdo a las instrucciones indicadas en sus respectivos manuales de aplicación. La evaluación fue realizada dentro del proceso de valoración inicial de los problemas psicopatológicos en contexto forense civil.

5.5.4. Diseño

Para el análisis de los datos se empleó un diseño autosuficiente como vía de aproximación al diseño de grupos diagnósticos (Rogers y Bender, 2003; Rogers, 2008e). Este diseño emplea puntos de corte estrictos de pruebas de referencia de detección de simulación para identificar grupos de probables fingidores frente a grupos de probables pacientes clínicos genuinos. Posteriormente se estudian las puntuaciones en la prueba bajo estudio en los diferentes grupos para determinar su capacidad discriminativa. Tal y como observaron Rogers et al.(2009), este diseño es especialmente útil en la evaluación de la validez externa de una medida mediante el estudio de los tamaños del efecto que establece entre los grupos clasificados.

Como prueba de referencia en simulación se ha empleado el MMPI-2 que ha sido la que mayor investigación y atención ha recibido por parte de los profesionales (Greene, 2008; Lally, 2003; Rogers, 2008b; Santamaría y González-Ordi, 2012), lo que ha motivado que fuera considerada como el criterio independiente empleado en esta investigación para establecer los grupos. Esta prueba proporciona tanto escalas más generales (F, Fb) como especializadas (Fp, Ds...) que han sido analizadas y estudiadas de forma comprensiva en un extenso meta-análisis de 65 estudios en el campo de la simulación mostrando que obtenían tamaños del efecto entre grande y muy grandes (Rogers et al., 2003).

Aunque las escalas estándar (F, Fb) han mostrado tamaños del efecto grandes, Rogers et al. (2003) informó de problemas en su aplicación debido a la gran variabilidad en los puntos de corte sugeridos y el riesgo potencial de falsos positivos con ciertos trastornos. Por este motivo, se seleccionó para este estudio Fp como la escala empleada para el establecimiento de puntos de corte estrictos ya que es menos sensible a la presencia de psicopatología genuina y en la investigación se han encontrado puntos de corte más estables (Arbisi y Ben-Porath, 1998; Greene, 2008; Hellman, 2010).

Las puntuaciones de cada participante en la escala Fp del MMPI-2 fueron empleadas para determinar su asignación al grupo de probable simulación, al grupo de probable paciente clínico genuino o al grupo indeterminado, casos en que no se puede establecer claramente su pertenencia a ninguno de los dos grupos establecidos. Los puntos de corte empleados para cada grupo fueron establecidos de acuerdo a estudios previos e investigaciones sobre la eficacia diagnóstica de Fp tanto a nivel internacional como en estudios con muestras españolas (Greene, 2008; Jiménez Gómez y Sánchez Crespo, 2002; Rogers et al., 2009, 2003), clasificando dentro del grupo de probable simulación a aquellos sujetos con puntuaciones directas en Fp iguales o mayores a 7; en el grupo

indeterminado aquellos con puntuaciones comprendidas entre 4 y 6 y en el grupo de probable respuesta honesta aquellos con puntuaciones menores de 4. Previa a cualquier clasificación, se estudió la posible presencia de patrones de respuesta aleatoria en alguno de los protocolos del MMPI-2. Para ello se estableció la exclusión de aquellos protocolos con VRIN >14 (Berry et al., 1991), sin que ningún protocolo fuera excluido por este motivo. Una vez implementadas las reglas de clasificación, el resultado fue la clasificación de 9 participantes en el grupo de probable simulación, 163 en el grupo de probable respuesta honesta y 35 en el grupo indeterminado.

5.6. Resultados

5.6.1. Puntuaciones en las escalas por grupo

De forma homogénea con el resto de análisis de esta tesis, se llevo a cabo un MANOVA de un factor con dos grupos para analizar si existían diferencias globales en las escalas del SIMS de acuerdo a la clasificación establecida (probable paciente clínico genuino, probable simulador). El modelo global fue estadísticamente significativo (Wilk's $\lambda = 0,697$, $F(5,166) = 14,41$, $p < 0,001$, $\eta^2 = 0,30$). A continuación se realizaron ANOVAs con corrección de Bonferroni y test post hoc (diferencia honestamente significativa de Tukey) de cada una de las escalas individuales del SIMS (Ps, Dn, Am, Bi y Af) y de la puntuación total para valorar la significación de las diferencias entre los tres grupos en cada escala. La magnitud de las diferencias entre los grupos fue analizada con la d de Cohen como medida del tamaño del efecto. Los resultados se muestran en la tabla 5.8 y en la figura 5.5.

Todas las escalas del SIMS obtuvieron puntuaciones significativamente más altas en el grupo de probable simulación frente al grupo de probable paciente clínico genuino, con tamaños del efecto muy grandes en el caso de la puntuación total y en las escalas Deterioro neurológico, Trastornos amnésicos y Psicosis y tamaños del efecto moderados en Trastornos afectivos y Baja inteligencia.

Tabla 5.8
Medias, desviaciones típicas, estadístico F y tamaño del efecto para los grupos de probable paciente clínico genuino y probable simulador

	Probables pacientes clínicos genuinos (n=163)	Probables simuladores (n=9)	F	d
Total SIMS	13,82 (6,28)	31,00 (13,53)	54,47***	2,54
Psicosis (Ps)	0,72 (1,15)	4,00 (4,82)	39,04***	2,15
Deterioro neurológico (Dn)	2,38 (2,17)	8,00 (3,81)	52,13***	2,49
Trastornos amnésicos (Am)	1,82 (2,09)	6,67 (4,03)	40,62***	2,20
Baja inteligencia (Bi)	2,15 (1,4)	3,33 (2,24)	5,72*	0,82
Trastornos afectivos (Af)	6,74 (2,55)	9,00 (1,66)	6,87*	0,90

*** $p < 0,001$ * $p < 0,05$

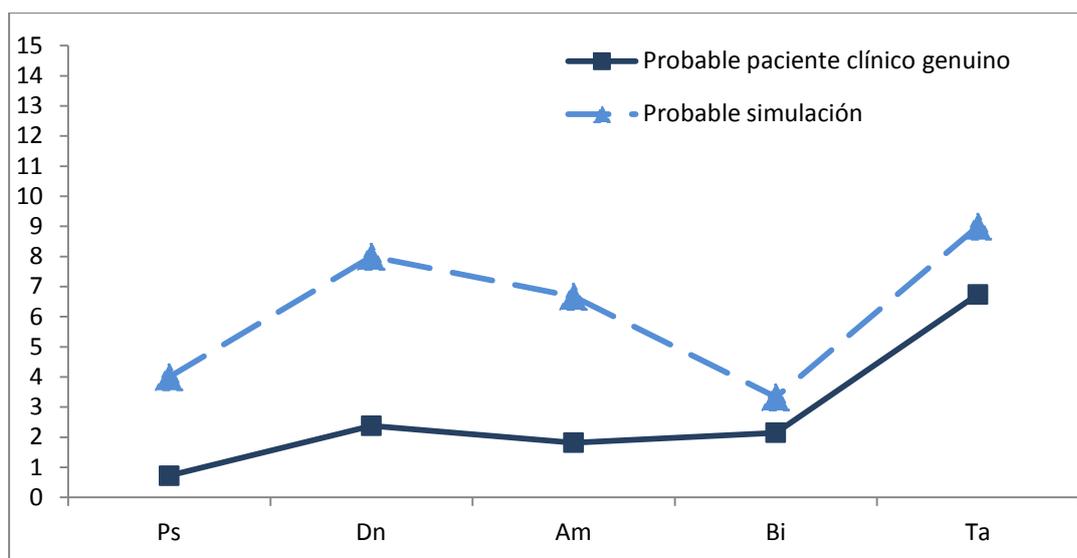


Figura 5.5. Medias para los grupos de probable respuesta honesta y probable simulación

Dado el reducido tamaño del grupo de probable simulación ($n=9$), se decidió realizar de forma complementaria un análisis no paramétrico de los datos para verificar los resultados del análisis paramétrico comentado previamente, utilizando para ello el estadístico U de Mann-Whitney, cuyos resultados pueden observarse en la tabla 5.9. Tal y como puede advertirse, se observan diferencias significativas en todas las escalas, de forma convergente a lo señalado previamente.

Tabla 5.9

Mediana, rango intercuartílico (IQR), rango promedio y estadístico U de Mann-Whitney del grupo sin sospecha de simulación y con sospecha de simulación en las escalas del SIMS

	Probables pacientes clínicos genuinos ($n=163$)		Probables simuladores ($n=9$)		U	Z	p
	Mdn. (IQR)	Rango promedio	Mdn. (IQR)	Rango promedio			
Total SIMS	13 (9-17)	82,5	27 (23-35)	158,3	87,0	4,45	<0,001
Psicosis (Ps)	0 (0-1)	84,1	2 (1-6)	129,4	347,5	2,95	0,002
Deterioro neurológico (Dn)	2 (1-3)	83,0	8 (6-11)	150,4	158,5	4,01	<0,001
Trastornos amnésicos (Am)	1 (0-3)	83,0	5 (4-10)	149,8	163,5	4,02	<0,001
Baja inteligencia (Bi)	2 (1-3)	85,0	3 (2-4)	112,8	496,5	1,67	0,047
Trastornos afectivos (Af)	7 (5-9)	84,1	10 (8-11)	130,3	339,5	2,73	0,002

5.6.2. Utilidad diagnóstica

Se estudió la utilidad diagnóstica de los distintos puntos de corte para la puntuación total del SIMS en la discriminación entre el grupo de probables pacientes clínicos genuinos y el grupo de probables simuladores (excluyendo al grupo indeterminado). Al igual que en el estudio 1, se consideró un amplio rango de posibles puntos de corte sugeridos por investigaciones previas.

Se emplearon curvas COR para examinar la precisión clasificatoria de las escalas del SIMS a la hora de diferenciar el grupo de simuladores análogos del grupo clínico de pacientes. Adicionalmente, se calcularon el poder predictivo positivo y el poder predictivo negativo (PPP y PPN) estimados de acuerdo a diferentes tasas base de prevalencia recomendadas en la literatura. En la tabla 5.10 se muestran los resultados obtenidos para la puntuación total. El área bajo la curva fue de 0,94 (ET=0,028). En las tablas 5.11 a 5.15 se muestran los resultados para cada una de las escalas del SIMS, cuyas áreas bajo la curva fueron 0,76 (ET =0,10) en la escala Psicosis, 0,89 (ET=0,07) en la escala Deterioro neurológico, 0,89 (ET = 0,04) en la escala Trastornos amnésicos, 0,66 (ET = 0,07) en la escala Baja inteligencia y 0,77 (ET=0,07) en la escala Trastornos afectivos.

Tabla 5.10

Estadísticos de utilidad diagnóstica de la puntuación total en el SIMS en la discriminación entre pacientes con y sin sospecha de simulación

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
> 14	1,00	0,58	0,60	0,29	1,00	0,50	1,00	0,70	1,00
> 16	1,00	0,73	0,74	0,40	1,00	0,61	1,00	0,79	1,00
> 19	0,89	0,83	0,83	0,48	0,98	0,69	0,95	0,84	0,88
> 23	0,67	0,94	0,93	0,68	0,94	0,84	0,87	0,92	0,74
> 24	0,56	0,94	0,92	0,64	0,92	0,81	0,83	0,91	0,68
> 26	0,56	0,96	0,94	0,70	0,92	0,85	0,83	0,93	0,68
> 32	0,44	0,99	0,96	0,86	0,91	0,94	0,81	0,97	0,64
> 40	0,11	1,00	0,95	1,00	0,86	1,00	0,72	1,00	0,53

Nota. Sen = sensibilidad, Esp = especificidad, %GI = porcentaje global de clasificaciones correctas, PPP = poder predictivo positivo, PPN = poder predictivo negativo. TP = Tasa de prevalencia.

Tabla 5.11

Estadísticos de utilidad diagnóstica de la puntuación en la escala Psicosis (Ps) en la discriminación entre pacientes con y sin sospecha de simulación

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Ps > 0	0,78	0,56	0,58	0,24	0,94	0,43	0,86	0,64	0,72
Ps > 1	0,56	0,87	0,85	0,43	0,92	0,65	0,82	0,81	0,66
Ps > 2	0,44	0,91	0,89	0,48	0,90	0,69	0,79	0,84	0,62
Ps > 3	0,44	0,95	0,92	0,62	0,91	0,80	0,80	0,90	0,63
Ps > 4	0,44	0,98	0,95	0,81	0,91	0,91	0,80	0,96	0,64
Ps > 5	0,22	0,99	0,95	0,86	0,88	0,94	0,75	0,97	0,56
Ps > 6	0,22	1,00	0,96	1,00	0,88	1,00	0,75	1,00	0,56
Ps > 0	0,78	0,56	0,58	0,24	0,94	0,43	0,86	0,64	0,72

Nota. Sen = sensibilidad, Esp = especificidad, %GI = porcentaje global de clasificaciones correctas, PPP = poder predictivo positivo, PPN = poder predictivo negativo. TP = Tasa de prevalencia.

Tabla 5.12

Estadísticos de utilidad diagnóstica de la puntuación en la escala Deterioro neurológico (Dn) en la discriminación entre pacientes con y sin sospecha de simulación

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Dn > 0	1,00	0,20	0,24	0,18	1,00	0,35	1,00	0,55	1,00
Dn > 1	0,89	0,44	0,46	0,22	0,96	0,40	0,90	0,61	0,80
Dn > 2	0,89	0,61	0,63	0,29	0,97	0,50	0,93	0,70	0,85
Dn > 3	0,89	0,75	0,76	0,39	0,97	0,61	0,94	0,78	0,87
Dn > 4	0,89	0,83	0,83	0,48	0,98	0,69	0,95	0,84	0,88
Dn > 5	0,78	0,91	0,90	0,60	0,96	0,78	0,91	0,89	0,80
Dn > 6	0,67	0,93	0,92	0,64	0,94	0,81	0,87	0,91	0,74
Dn > 7	0,56	0,96	0,94	0,73	0,92	0,87	0,83	0,94	0,68
Dn > 8	0,44	0,99	0,96	0,86	0,91	0,94	0,81	0,97	0,64
Dn > 9	0,44	1,00	0,97	1,00	0,91	1,00	0,81	1,00	0,64

Nota. Sen = sensibilidad, Esp = especificidad, %GI = porcentaje global de clasificaciones correctas, PPP = poder predictivo positivo, PPN = poder predictivo negativo. TP = Tasa de prevalencia.

Tabla 5.13

Estadísticos de utilidad diagnóstica de la puntuación en la escala Trastornos amnésicos (Am) en la discriminación entre pacientes con y sin sospecha de simulación

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Am > 1	1,00	0,56	0,59	0,29	1,00	0,50	1,00	0,70	1,00
Am > 2	0,89	0,71	0,72	0,35	0,97	0,57	0,94	0,76	0,86
Am > 3	0,78	0,82	0,82	0,44	0,95	0,65	0,90	0,81	0,79
Am > 4	0,56	0,89	0,87	0,47	0,92	0,68	0,82	0,83	0,67
Am > 5	0,44	0,93	0,90	0,52	0,90	0,72	0,80	0,86	0,63
Am > 6	0,44	0,96	0,94	0,68	0,91	0,84	0,80	0,92	0,63
Am > 7	0,44	0,98	0,95	0,81	0,91	0,91	0,80	0,96	0,64
Am > 8	0,33	0,99	0,95	0,83	0,89	0,92	0,78	0,96	0,60

Nota. Sen = sensibilidad, Esp = especificidad, %GI = porcentaje global de clasificaciones correctas, PPP = poder predictivo positivo, PPN = poder predictivo negativo. TP = Tasa de prevalencia.

Tabla 5.14

Estadísticos de utilidad diagnóstica de la puntuación en la escala Baja inteligencia (Bi) en la discriminación entre pacientes con y sin sospecha de simulación

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Bi > 1	1,00	0,36	0,39	0,22	1,00	0,40	1,00	0,61	1,00
Bi > 2	0,56	0,59	0,59	0,19	0,88	0,37	0,76	0,57	0,57
Bi > 3	0,22	0,85	0,82	0,21	0,86	0,39	0,72	0,60	0,52
Bi > 4	0,11	0,95	0,91	0,29	0,86	0,49	0,71	0,69	0,52
Bi > 5	0,11	0,99	0,94	0,62	0,86	0,80	0,72	0,90	0,53

Nota. Sen = sensibilidad, Esp = especificidad, %GI = porcentaje global de clasificaciones correctas, PPP = poder predictivo positivo, PPN = poder predictivo negativo. TP = Tasa de prevalencia.

Tabla 5.15

Estadísticos de utilidad diagnóstica de la puntuación en la escala Trastornos afectivos (Af) en la discriminación entre pacientes con y sin sospecha de simulación

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Ta > 5	1,00	0,32	0,35	0,21	1,00	0,39	1,00	0,59	1,00
Ta > 6	0,89	0,45	0,48	0,22	0,96	0,41	0,91	0,62	0,80
Ta > 7	0,78	0,58	0,59	0,24	0,94	0,44	0,86	0,65	0,72
Ta > 8	0,67	0,72	0,72	0,30	0,92	0,51	0,84	0,71	0,68
Ta > 9	0,56	0,87	0,85	0,43	0,92	0,65	0,82	0,81	0,66
Ta > 10	0,11	0,93	0,89	0,23	0,86	0,41	0,71	0,62	0,51
Ta > 11	0,00	0,99	0,94	0,00	0,85	0,00	0,70	0,00	0,50

Nota. Sen = sensibilidad, Esp = especificidad, %GI = porcentaje global de clasificaciones correctas, PPP = poder predictivo positivo, PPN = poder predictivo negativo. TP = Tasa de prevalencia.

5.6.3. Relación con otras pruebas

Se examinó la correlación del SIMS con las escalas de exageración y minimización de síntomas del MMPI-2 en el conjunto total de la muestra, sin distinguir entre con y sin sospecha de simulación. Tal y como cabría predecir, se halló un claro patrón de correlaciones entre la puntuación total del SIMS con las distintas escalas de exageración de síntomas del MMPI-2 ($M r = 0,63$, calculada con la z de Fisher; véase la tabla 5.16) y con las escalas de minimización, particularmente con la escala K. Las escalas del SIMS muestran también un claro patrón de correlación con las escalas del MMPI-2, más marcado en aquellos casos convergentes en cuanto al área de simulación al que están destinadas.

Tabla 5.16

Correlaciones entre las escalas del SIMS y las escalas de exageración y minimización de síntomas del MMPI-2 en pacientes evaluados en contexto forense civil.

Escalas del SIMS	MMPI-2 Escalas de validez							
	Exageración de síntomas				Minimización de síntomas			
	F	Fb	Fp	FBS	Ds-R	F - K	L	K
Total SIMS	0,69**	0,74**	0,49**	0,56**	0,62**	0,70**	-0,05	-0,49**
Psicosis (Ps)	0,51**	0,55**	0,45**	0,23**	0,48**	0,53**	-0,13	-0,40**
Deterioro neurológico (Dn)	0,63**	0,62**	0,45**	0,53**	0,54**	0,61**	-0,01	-0,41**
Trastornos amnésicos (Am)	0,56**	0,60**	0,39**	0,43**	0,48**	0,57**	-0,05	-0,40**
Baja inteligencia (Bi)	0,30**	0,31**	0,22**	0,21**	0,24**	0,30**	-0,02	-0,21**
Trastornos afectivos (Af)	0,41**	0,49**	0,23**	0,47**	0,44**	0,43**	0,01	-0,32**

Nota. N = 207 F = Incoherencia; Fb = F posterior; Fp = Psicopatología infrecuente; Ds-R=Escala de disimulación – revisada; FBS = Escala de fingirse enfermo; F-K =Índice F-K; MMPI-2 = Inventario Multifásico de Personalidad de Minnesota-2.

** $p < 0,01$

A continuación, se realizó el mismo análisis pero con las escalas de validez del MMPI-2-RF. Al igual que ocurría con el MMPI-2, la correlación entre la escala total del SIMS y las escalas de exageración de síntomas del MMPI-2-RF mostraron valores notables ($M r = 0,66$; véase tabla 5.17), obteniendo su correlación más alta con la escala más general de exageración de síntomas del MMPI-2-RF, la escala F-r.

A nivel de las escalas del SIMS, puede observarse la convergencia entre ciertas escalas específicas del SIMS y del MMPI-2-RF destinadas a la detección de áreas similares de simulación. Así, la escala Psicosis del SIMS (Ps) muestra su máxima correlación con la escala Fp-r del MMPI-2-RF; la escala Trastornos amnésicos del SIMS (Am) obtiene el mayor valor de correlación con la escala RBS del MMPI-2-RF creada específicamente para la detección de simulación de problemas de memoria o la escala Deterioro neurológico del SIMS (Dn) alcanza su valor más alto de correlación con la escalas de Quejas somáticas infrecuentes (Fs), la escala de problemas de memoria (RBS) y la escala general de exageración de síntomas (F-r). Al igual que ocurría en el estudio del MMPI-2, se observa un leve patrón de correlaciones negativas de las escalas del SIMS con la escalas de minimización de síntomas del MMPI-2-RF, particularmente con K-r.

Tabla 5.17
Correlaciones entre las escalas del SIMS y las escalas de exageración y minimización de síntomas del MMPI-2-RF en pacientes en contexto forense civil.

Escalas del SIMS	MMPI-2-RF Escalas de validez						
	Exageración de síntomas				Minimización de síntomas		
	F-r	Fp-r	Fs	FBS-r	RBS	L-r	K-r
Total SIMS	0,74**	0,60**	0,67**	0,56**	0,74**	-0,08	-0,43**
Psicosis (Ps)	0,50**	0,55**	0,43**	0,33**	0,40**	-0,08	-0,34**
Deterioro neurológico (Dn)	0,69**	0,54**	0,64**	0,52**	0,66**	-0,10	-0,31**
Trastornos amnésicos (Am)	0,61**	0,43**	0,49**	0,43**	0,63**	-0,06	-0,32**
Baja inteligencia (Bi)	0,31**	0,33**	0,30**	0,29**	0,35**	-0,05	-0,13**
Trastornos afectivos (Af)	0,47**	0,34**	0,43**	0,40**	0,50**	0,00	-0,38**

Nota. N = 207. F-r = Infrecuencia; Fp-r = Psicopatología infrecuente; Fs = Quejas somáticas infrecuentes; FBS-r = Validez de los síntomas; RBS = Escala de sesgo de respuestas; L-r = Virtudes inusuales; K-r = Validez del ajuste; MMPI-2-RF = Inventario Multifásico de Personalidad de Minnesota-2 Reestructurado.

** $p < 0,01$.

5.7. Discusión

Los resultados mostrados en este estudio han reflejado una alta capacidad discriminativa de la versión española del SIMS entre pacientes clasificados con probable respuesta sincera o probable simulación a partir de sus puntuaciones en la escala Fp del MMPI-2, la prueba de referencia en el campo de la detección de simulación (Rogers, 2008b; Santamaría y González-Ordi, 2012).

Los tamaños del efecto han sido muy grandes en el caso de la puntuación total y en las escalas Deterioro neurológico, Trastornos amnésicos y Psicosis y moderados en las escalas Trastornos afectivos y Baja inteligencia. Estos resultados concuerdan con los hallados en la literatura científica que han mostrado las mayores elevaciones en la escala Deterioro neurológico en diseños análogos ante la instrucción de simular baja laboral para obtener compensaciones (Rogers et al., 2005) o en estudios en contexto forense civil en solicitud de indemnizaciones o compensaciones donde aparecían también las mayores elevaciones en las escalas Deterioro neurológico, Psicosis y Trastornos amnésicos (Alwes, 2006) o en Psicosis y Trastornos amnésicos (Clegg et al., 2009). No obstante, serán necesarios estudios futuros que indaguen en el perfil específico de elevación de las escalas del SIMS según el contexto de evaluación (civil o penal) y el área psicopatológica alegada.

Los análisis de la utilidad diagnóstica de los puntos de corte de la puntuación total y de las escalas apoyan la necesidad de aumentar los puntos de corte originales sugeridos por Smith y Burger (1997) y Widows y Smith (2005) con el objetivo de aumentar marcadamente su especificidad disminuyendo parcialmente su sensibilidad. Los puntos de corte sugeridos en el estudio 1 han mostrado preliminarmente un satisfactorio comportamiento, si bien, como se señalará posteriormente, el limitado tamaño de la muestra de probable simulación en este estudio y la exclusión del “grupo indeterminado” en los análisis de acuerdo al diseño de medidas autosuficientes, hace muy inestable las estimaciones de utilidad diagnóstica obtenidas que deben considerarse con suma precaución y de forma meramente orientativa

A estos resultados ha de añadirse que el patrón global de correlaciones entre las escalas del SIMS y del MMPI-2 y del MMPI-2-RF respalda la validez del SIMS cuando es aplicado a casos civiles forenses en población española. Existe una clara convergencia entre ambas pruebas no solo a nivel global sino a nivel específico de sus escalas y las áreas de simulación que abarcan. Los valores obtenidos con el MMPI-2 son análogos a los hallados en investigaciones previas con el SIMS, con valores promedio entre la puntuación total del SIMS y la escala F cercanos a 0,64 y oscilando entre 0,43 y 0,84 (Carter et al., 2009; Cima, Hollnack, et al., 2003; Heinze y Purisch, 2001; Leung, 2010; Widows y Smith, 2005). El patrón de correlación de las escalas del SIMS con la escala K de

minimización de síntomas también es una pauta habitual en los estudios realizados (Cima, Hollnack, et al., 2003; Rogers et al., 2009), obteniendo sistemáticamente correlaciones inferiores y no significativas con la escala L, tal y como también aparece en este estudio.

En el caso del MMPI-2-RF, se obtienen patrones satisfactorios de correlaciones y entre aquellas escalas del SIMS y del MMPI-2-RF que aluden a una misma área de exageración de síntomas.

Varias limitaciones relevantes de este estudio han de ser mencionadas. En primer lugar, el diseño autosuficiente empleado en este estudio se considera una aproximación al diseño de grupos conocidos. Será deseable poder emplear diseños de grupos conocidos en próximos estudios que garanticen estimaciones más precisas y veraces de la utilidad diagnóstica de los distintos puntos de corte. Si bien el SIRS (Rogers et al., 1992) ha sido una de las pruebas frecuentemente empleadas en la investigación con el SIMS para el establecimiento de los grupos diagnósticos conocidos, el hecho de no disponer de una adaptación validada en España en la actualidad, hizo que se optara en este estudio por el MMPI-2 como la prueba validada en España con mayores evidencias en la detección de simulación y exageración de síntomas (González-Ordi y Iruarrizaga Díez, 2005; Jiménez y Sánchez, 2004; Jiménez Gómez y Sánchez Crespo, 2002; Sánchez Crespo et al., 2008; Sánchez Crespo, 2002; Santamaría y González-Ordi, 2012).

La elección de la escala Fp como escala criterio para la asignación de participantes al grupo de probable respuesta honesta o probable simulación también afecta a los resultados del estudio. Si bien esta escala ya ha sido utilizada previamente en la configuración de grupos en diseños autosuficientes en contexto forense civil (Rogers et al., 2009) y ha sido elegida por haberse mostrado como una de las escalas del MMPI-2 más estable en sus puntos de corte y con el menor número de falsos positivos (Rogers et al., 2003), un aspecto crítico en la configuración de un diseño autosuficiente, no es menos cierto que precisamente estas características probablemente han hecho que dentro del grupo de probable simulación solo se hayan clasificado aquellos simuladores con una presentación de su simulación más global, marcada y clara, pudiendo dejar en el grupo indeterminado a simuladores más sutiles o sofisticados.

Así mismo la existencia del grupo indeterminado, algo recomendable a nivel metodológico para garantizar la correcta asignación de los casos (Rogers, 2008e), hace que las estimaciones de sensibilidad y especificidad obtenidas probablemente sean sobre-estimaciones de los valores reales ya que se han eliminado aquellos casos precisamente más dudosos, aumentando probablemente de forma artificial los valores de sensibilidad y especificidad obtenidos por la propia configuración y

características del diseño. Esto explicaría el hecho de haber obtenido valores mucho más altos tanto de sensibilidad como de especificidad comparativamente con el estudio 1.

Finalmente, ha de señalarse, como se apuntaba anteriormente, que el reducido tamaño del grupo de probable respuesta de simulación ($n=9$) afecta marcadamente a la estabilidad de las estimaciones de tamaño del efecto y utilidad diagnóstica, haciéndolas más inestables e imprecisas. Serán necesarios estudios futuros, preferentemente con grupos diagnósticos conocidos, con tamaños muestrales más amplios que puedan replicar los análisis realizados y confirmar o refutar estos resultados.

5.8. Conclusiones

En esta investigación se ha examinado la validez y utilidad diagnóstica de la adaptación española del SIMS para detectar simulación de psicopatología en diversas muestras de pacientes clínicos y simuladores. Para ello se han llevado a cabo dos estudios, el primero con un diseño de grupos análogos en el que se han considerado diferentes áreas psicopatológicas y contextos de simulación (trastornos psicóticos, deterioro neurológico, trastornos afectivos) y se ha analizado la capacidad discriminativa del SIMS en cada una de las áreas tanto para la puntuación total como para cada una de sus escalas. El segundo estudio pretendía indagar en la capacidad del SIMS para detectar simulación en contextos reales de valoración de bajas laborales, una de las principales áreas de utilización futura de la prueba.

Los resultados obtenidos con grupos análogos han respaldado la capacidad de la prueba de detección simulación de distintas sintomatologías psicopatológicas (simulación de trastornos psicóticos, deterioro neurológico y trastornos afectivos) tanto de la puntuación total como de las escalas específicas a cada caso, con tamaños del efecto entre moderados y muy grandes, de acuerdo a la clasificación más exigente de Rogers (2008d) para la investigación específica en simulación.

La puntuación total ha sido la que mayor tamaño del efecto ha obtenido en todas las condiciones y en el grupo indiferenciado, lo que apoya su carácter de *screening* global de diferentes sintomatologías simuladas y sustenta su potencial como prueba de despistaje de amplio espectro. Por otra parte, las escalas de la prueba han mostrado su mayor utilidad diagnóstica para la detección de aquella área sintomatológica para la que fueron diseñadas. Finalmente, el análisis de los puntos de corte, ha sugerido el aumento de sus valores para el uso de la prueba en población española con puntos de corte más elevados que los establecidos originalmente por los autores, de forma convergente con la mayoría de la investigación científica realizada sobre el SIMS (Alwes, 2006; Cima,

Hollnack, et al., 2003; Clegg et al., 2009; Dandachi-FitzGerald et al., 2011; Edens et al., 1999, 2007; Lewis et al., 2002; Merckelbach y Smith, 2003; Peters, Jelicic, et al., 2013; Rogers, Hinds, et al., 1996; Vitacco et al., 2006; Wisdom et al., 2010).

Los resultados obtenidos con casos reales en contexto civil forense refuerzan la adecuación de la prueba para la detección de simulación de psicopatología en muestras españolas, mostrando tamaños del efecto muy grandes entre el grupo de probables pacientes clínicos genuinos y el grupo de probables simuladores establecidos a partir de sus puntuaciones en la escala Fp del MMPI-2. El estudio de los puntos de corte también coincide en recomendar un aumento con relación a los originales propuestos por los autores Smith y Burger (1997) y aporta una evidencia preliminar de apoyo a los sugeridos a partir del estudio 1 de grupos análogos para su aplicación en muestras españolas. Adicionalmente, el claro patrón de correlaciones hallado entre las puntuaciones del SIMS y las diversas escalas de exageración de síntomas del MMPI-2 y del MMPI-2-RF refuerza la validez de la prueba en la detección de simulación de psicopatología en muestras españolas.

Son diversas las limitaciones de los estudios aquí presentados, como ya se ha señalado en los apartados previos de discusión y son variadas las líneas de investigación abiertas a partir de los resultados encontrados que permitirán refutar o confirmar algunas de las inferencias deducidas de los datos presentados. Particularmente serán necesarios estudios de grupos diagnósticos en contextos civiles y penales que permitan obtener evidencias adicionales de la capacidad de detección de simulación de psicopatología de la versión española del SIMS. También serán deseables estudios análogos que aborden la capacidad discriminativa de simulación de trastornos cognitivos, un aspecto no abordado en el estudio análogo presentado en esta tesis. No obstante estas limitaciones y la necesidad de estudios futuros, el conjunto de los resultados encontrados respalda la capacidad discriminativa del SIMS entre pacientes clínicos y simuladores en muestras españolas tanto en diseños controlados de sujetos análogos con muestras clínicas de referencia como en contextos reales. Los resultados aportan evidencias de la utilidad de la versión española del SIMS.

Esta investigación también aporta contribuciones significativas a la literatura científica existente sobre el SIMS, particularmente en el análisis de las escalas y su funcionamiento en condiciones específicas de simulación de ciertos trastornos y áreas psicopatológicas así como la sugerencia de puntos de corte alternativos a los sugeridos originalmente en las distintas escalas del SIMS (Ps, Dn, Am, Bi y Af).

Finalmente, esta investigación relaciona las puntuaciones del SIMS con las escalas de exageración y minimización de síntomas del MMPI-2-RF. El MMPI-2-RF es una prueba de desarrollo

relativamente reciente pero que rápidamente está convirtiéndose en un estándar para la evaluación de exageración de síntomas a nivel internacional (Gervais et al., 2010; Gervais, Wygant, Sellbom y Ben-Porath, 2011; Gervais, Lees-Haley, et al., 2007; Harp et al., 2011; Hoelzle, Nelson y Arbisi, 2012; Marion et al., 2011; Rogers, Gillard, et al., 2011; Santamaría et al., 2010; Santamaría y González-Ordi, 2012; Schroeder et al., 2011; Sellbom y Bagby, 2010; Sellbom et al., 2010, 2012; Wygant et al., 2009, 2010, 2011; Youngjohn et al., 2011). La convergencia entre las escalas del SIMS y del MMPI-2-RF en las áreas sintomatológicas esperadas y el claro patrón de correlaciones hallados en las muestras españolas estudiadas resulta una fuente de respaldo adicional y relevante a la versión española del SIMS y su validez en la detección de síntomas simulados.

6. Utilidad diagnóstica del SIMS en la detección de simulación de lumbalgia

6.1. Objetivos e hipótesis de trabajo

Se calcula que ocho de cada diez personas sufrirán dolor en la zona baja de la espalda, o lumbalgia, al menos una vez en su vida (Muñoz-Gómez, 2003), siendo su etiología inespecífica en el 80% de los casos (Moix, Cano y Grupo español de trabajo del programa COST B13 de la comisión europea, 2006) y en el 90% de los casos con compensación económica (León Mateos, Jover Jover, Loza Santamaría y Pérez Nieto, 2007).

Entre los asalariados españoles se ha calculado una media de 22 días anuales de incapacidad temporal por dolor de espalda (González Viejo y Condón Huerta, 2000; Kovacs y Gil del Real, 2000). Se ha estimado que hay abuso respecto a esta prestación en un tercio de los casos, bien por ser innecesario el reposo, bien por prolongarse en exceso (Quijano, 1999; Ruiz Téllez, 1999).

La capacidad de los médicos para detectar con su juicio clínico la simulación en estos casos parece ser notablemente limitada (Jung y Reidenberg, 2007) por lo que las herramientas de evaluación objetiva cobran una importancia clave, particularmente las pruebas de *screening* por su posible uso en primera instancia y en la toma de decisiones sobre si derivar al paciente a una evaluación más exhaustiva. Esto es especialmente relevante dada la frecuente comorbilidad de los problemas de dolor crónico con síntomas emocionales, cognitivos y neurológicos (Brage, Sandanger y Nygard, 2007; Casado Morales, Moix y Vidal Fernández, 2008; Casado Morales y Urbano, 2001; Steenstra et al., 2005) y la necesidad de hacer un cribaje de las diferentes áreas donde puede estar presente una exageración de la sintomatología asociada.

Si bien en los últimos años ha habido avances notables en el área de detección de simulación en trastornos mentales y trastornos cognitivos, estos avances han sido sensiblemente menores en el ámbito de la detección de simulación de quejas médicas, un área generalmente relegada en la investigación en simulación (Rogers et al., 1998) y considerada la nueva frontera en la investigación en detección de la simulación (Berry y Nelson, 2010; Rogers y Correa, 2008).

Prueba de ello es el escaso número de estudios que han analizado la capacidad del SIMS en la detección de simulación en dolor crónico. La revisión realizada de la literatura científica solo informa

de un único estudio centrado en la simulación de esguince cervical y realizado con universitarios en un diseño de simulación análogo sin muestra clínica de referencia (Merten et al., 2008).

Esta ausencia de investigación con el SIMS en simulación de dolor crónico es llamativa dado el potencial de la prueba para cribaje de simulación en dolor crónico. Dado que los pacientes con cronificación de problemas somáticos, como la lumbalgia, suelen cursar con síntomas cognitivos y emocionales, además de físicos, se estima que el SIMS podrá ser una herramienta interesante que permita una amplia evaluación de la exageración de síntomas en las distintas áreas, dada sus escalas específicas referidas a problemas cognitivos (escalas de Trastornos amnésicos), neurológicos (escala de Deterioro neurológico) y afectivos (escala de trastornos afectivos), así como su puntuación total como compendio de las distintas áreas pretendidamente afectadas.

El objetivo de esta investigación es analizar la capacidad y utilidad de las escalas del SIMS para identificar la presencia de quejas somáticas no creíbles relacionadas con la lumbalgia crónica, empleando para ello por primera vez en la literatura científica un diseño de grupos diagnósticos conocidos. Para ello se compararán los resultados en el SIMS de participantes sin patología (controles) y de pacientes clínicos con lumbalgia crónica con pacientes clínicos con sospecha de simulación y simuladores análogos instruidos para simular lumbalgia de forma contingente a una reclamación de compensación económica. Adicionalmente se estudiará la convergencia del SIMS con otra prueba de referencia en el campo para la detección de simulación de quejas somáticas, el MMPI-2-RF (Ben-Porath and Tellegen, 2008). Ningún otro estudio hasta la fecha ha relacionado las puntuaciones entre ambas pruebas en el contexto de quejas somáticas.

Por lo tanto, se plantean tres *objetivos* para este estudio:

- 1) Analizar la utilidad diagnóstica de la puntuación total del SIMS y de cada una de sus escalas en la detección de simulación de dolor crónico de lumbalgia en contextos médico-legales reales.
- 2) Analizar la eficacia de los puntos de corte de la puntuación total y las escalas.
- 3) Estudiar la validez convergente del SIMS con las escalas del MMPI-2-RF, particularmente aquellas que más eficacia han mostrado en la detección de simulación de quejas somáticas.

6.2. Método

6.2.1. Participantes

Se dispusieron de cuatro muestras: controles, pacientes clínicos con lumbalgia sin sospecha de simulación, pacientes clínicos con lumbalgia y con sospecha de simulación y simuladores análogos instruidos para simular lumbalgia de forma contingente a una reclamación de compensación económica.

La muestra 1 (control) estuvo compuesta por 30 participantes sanos (sin patología previa diagnosticada), 13 mujeres y 17 varones con una edad media de 48,8 años (dt=7,4; rango: 33–61).

La muestra 2 (pacientes clínicos con lumbalgia crónica sin sospecha de simulación) estuvo formada por 45 participantes que eran pacientes con lumbalgia crónica diagnosticada en el Sistema Público de Salud (SPS), 26 eran mujeres y 19 varones, con una edad media de 42,2 años (dt=8,8; rango 24-65). Estos participantes fueron evaluados en el contexto de un programa de educación postural para el tratamiento del dolor crónico y no se encontraban en situación de incapacidad temporal ni tenían un beneficio externo manifiesto.

La muestra 3 (pacientes clínicos con lumbalgia y con sospecha de simulación) consistió en 48 participantes, 33 mujeres y 15 varones, con una edad media de 43,9 años (dt=11,1; rango 21-64 años), identificados por expertos del hospital FREMAP Madrid como con sospecha de simulación de lumbalgia crónica debido al cumplimiento de los criterios establecidos en los protocolos de detección de simulación de lumbalgia en Capilla Ramírez y González-Ordi (2009) y González-Ordi, Capilla Ramírez y Matalobos Veiga (2008):

- no se encuentra evidencia física en la exploración física y en las pruebas de imagen y función (Rx, RMN y EMG) que justifiquen el dolor y discapacidad referidos por el paciente;
- la frecuencia, intensidad o duración de los síntomas excede con mucho lo usual en la patología diagnosticada y hay escasos o nulos avances terapéuticos;
- baja adherencia a las prescripciones médicas;
- cumplen al menos tres de los cinco signos de Waddell como criterio discriminante de sospecha clínica;
- existe sospecha clínica de simulación por al menos un profesional de la salud implicado en su tratamiento (médico o fisioterapeuta);

- existencia de beneficio externo, es decir, percepción de una retribución económica estando en situación de incapacidad temporal (IT) o pendientes de valoración por el Equipo de Valoración de Incapacidades (EVI) con el fin de obtener una incapacidad permanente (IP).

La muestra 4 (simuladores análogos instruidos para simular lumbalgia de forma contingente a una reclamación de compensación económica) estuvo integrada por 20 participantes sanos (sin patología previa diagnóstica), 11 mujeres y 9 varones, con una edad media de 42,7 años (dt=10,4; rango 24-62 años), que fueron instruidos para simular problemas físicos de salud relacionados con dolor en la parte baja de la espalda con el objetivo de obtener compensaciones económicas o laborales y que eran evaluados en un contexto forense civil.

6.2.2. Materiales

A todos los participantes se les aplicó una batería de pruebas para evaluar su estado físico y psicológico, entre las que se incluyeron el SIMS (González-Ordi y Santamaría, 2009; Widows y Smith, 2005) y el MMPI-2-RF (Ben-Porath y Tellegen, 2008; Santamaría, 2009).

El MMPI-2-RF es un inventario de evaluación de la personalidad psicopatológica que consta de 338 ítems dicotómicos (verdadero/falso) y que fue diseñado para representar eficientemente todos aquellos aspectos clínicos significativos que subyacían al conjunto de ítems del MMPI-2. Incluye seis conjuntos de escalas: validez, dimensiones globales, clínicas reestructuradas, problemas específicos, intereses y los cinco de personalidad psicopatológica. Para este estudio, se centró la atención en las escalas de validez, particularmente en las escalas de exageración de síntomas. El MMPI-2-RF dispone de cinco escalas de exageración de síntomas (F-r, Fp-r, Fs, FBS-r y RBS) que han mostrado su capacidad discriminativa de simulación en distintos contextos y particularmente en la detección de simulación en contextos médico-legales (Gervais et al., 2010, 2011; Rogers, Gillard, et al., 2011; Santamaría et al., 2010; Sellbom et al., 2012; Wygant et al., 2009, 2010, 2011)

6.2.3. Procedimiento

A los participantes de las muestras 1 (control), 2 (pacientes clínicos de lumbalgia sin sospecha de simulación) y 3 (pacientes clínicos de lumbalgia con sospecha de simulación) se les dieron indicaciones para que respondieran a las pruebas de forma sincera, tratando de reflejar de la forma más precisa posible los síntomas y características clínicas que presentaban para que pudieran ser

tenidas en cuenta en el proceso de posterior intervención y tratamiento. Se siguieron para ello las instrucciones estándar incluidas en los manuales de aplicación respectivos del SIMS y del MMPI-2-RF (Ben-Porath y Tellegen, 2008; Widows y Smith, 2005).

A los participantes de la muestra 4 (simuladores análogos) se les instruyeron para que simularan problemas físicos relacionados con dolores en la parte baja de su espalda en un contexto de solicitud de compensaciones económicas y laborales. Se les especificó que debían hacerlo de una manera creíble y convincente sin que pudiera detectarse que estaban mintiendo. En el apéndice C pueden encontrarse las instrucciones específicas dadas.

Todos los pacientes dieron su consentimiento informado previo a su evaluación. La clasificación de los participantes en cada uno de los grupos fue independiente y previa a sus respuestas al SIMS y al MMPI-2-RF.

6.3. Resultados

6.3.1. Puntuaciones en las escalas por grupo

Se realizó un MANOVA de un factor con cuatro grupos para analizar si existían diferencias globales en las escalas del SIMS de acuerdo a la condición establecida (control, clínico sin sospecha, clínico con sospecha, simulador análogo). El modelo global fue estadísticamente significativo (Wilk's $\lambda = 0,377$, $F_{(15, 367,55)} = 10,4$, $p < 0,001$, $\eta^2 = 0,28$). A continuación se realizaron ANOVAs con corrección de Bonferroni y test post hoc (diferencia honestamente significativa de Tukey) de cada una de las escalas individuales del SIMS (Ps, Dn, Am, Bi y Af) y de la puntuación total para valorar la significación de las diferencias entre los tres grupos en cada escala.

La magnitud de las diferencias entre los grupos fue analizada con la d de Cohen como medida del tamaño del efecto. De acuerdo a los criterios establecidos por Rogers (2008d) específicamente para la investigación en estilos de respuestas, se consideró los siguientes intervalos: 0,75 para moderado, 1,25 para grande y 1,50 para muy grande. Los resultados se muestran en la tabla 6.1 y en la figura 6.1.

Tabla 6.1

Medias, desviaciones típicas, estadístico F y tamaños del efecto para los grupos control, clínico sin sospecha de simulación, clínico con sospecha de simulación y simuladores análogos

Escalas del SIMS	Control (n=30)	Pacientes sin sospecha de simulación (n=45)	Pacientes con sospecha de simulación (n=48)	Simuladores análogos (n=20)	F	d ₁	d ₂
Total SIMS	7,31 ^a (3,04)	7,09 ^a (2,98)	15,08 ^b (7,54)	19,16 ^c (6,19)	35,493***	1,39	2,94
Psicosis (Ps)	0,59 ^{a,b} (0,68)	0,27 ^a (0,69)	0,46 ^{a,b} (0,74)	0,95 ^b (1,13)	3,61*	0,27	0,82
Deterioro neurológico (Dn)	0,72 ^a (0,92)	0,80 ^a (1,12)	4,06 ^b (3,21)	5,68 ^c (2,67)	35,14***	1,36	2,89
Trastornos amnésicos (Am)	0,69 ^{a,b} (1,2)	0,18 ^a (0,58)	1,65 ^b (2,32)	1,37 ^b (2,01)	6,65***	0,87	1,02
Baja inteligencia (Bi)	1,52 ^a (1,21)	2,11 ^{a,b} (1,47)	2,5 ^b (1,22)	2,53 ^b (1,26)	3,92*	0,29	0,30
Trastornos afectivos (Af)	3,79 ^a (1,82)	3,73 ^a (1,47)	6,42 ^b (2,5)	8,63 ^c (2,5)	34,39***	1,31	2,72

Nota. Las medias marcadas con diferente letra son significativamente diferentes a un nivel de $p < 0,05$

d₁ tamaño del efecto para la comparación del grupo de pacientes clínicos sin sospecha de simulación frente al grupo de pacientes clínicos con sospecha de simulación.

d₂ tamaño del efecto para la comparación del grupo de pacientes clínicos sin sospecha de simulación frente al grupo de simuladores análogos.

* $p < 0,05$ *** $p < 0,001$

Se puede advertir como existen diferencias significativas entre los cuatro grupos en todas las escalas. El análisis de las pruebas post hoc y la asignación a grupos homogéneos permite observar que el grupo control y el grupo clínico sin sospecha de simulación no muestran diferencias significativas entre sí, mientras que sí lo hacen con respecto al grupo de pacientes clínicos con sospecha de simulación y al grupo de simuladores análogos. Las puntuaciones del SIMS que muestran un tamaño del efecto grande son la puntuación Total y las escalas Deterioro neurológico y Trastornos afectivos, por este orden de magnitud.

El análisis de la figura 6.1 permite observar visualmente la similitud en el perfil entre el grupo control y clínico sin sospecha de simulación y la diferencia con las puntuaciones de los grupos de simulación (pacientes clínicos con sospecha de simulación y simuladores análogos). Es interesante señalar como el grupo de simulación análogo obtiene puntuaciones significativamente más altas que el grupo de pacientes clínicos con sospecha de simulación, como se advierte en la tabla 6.1 y en la figura 6.1.

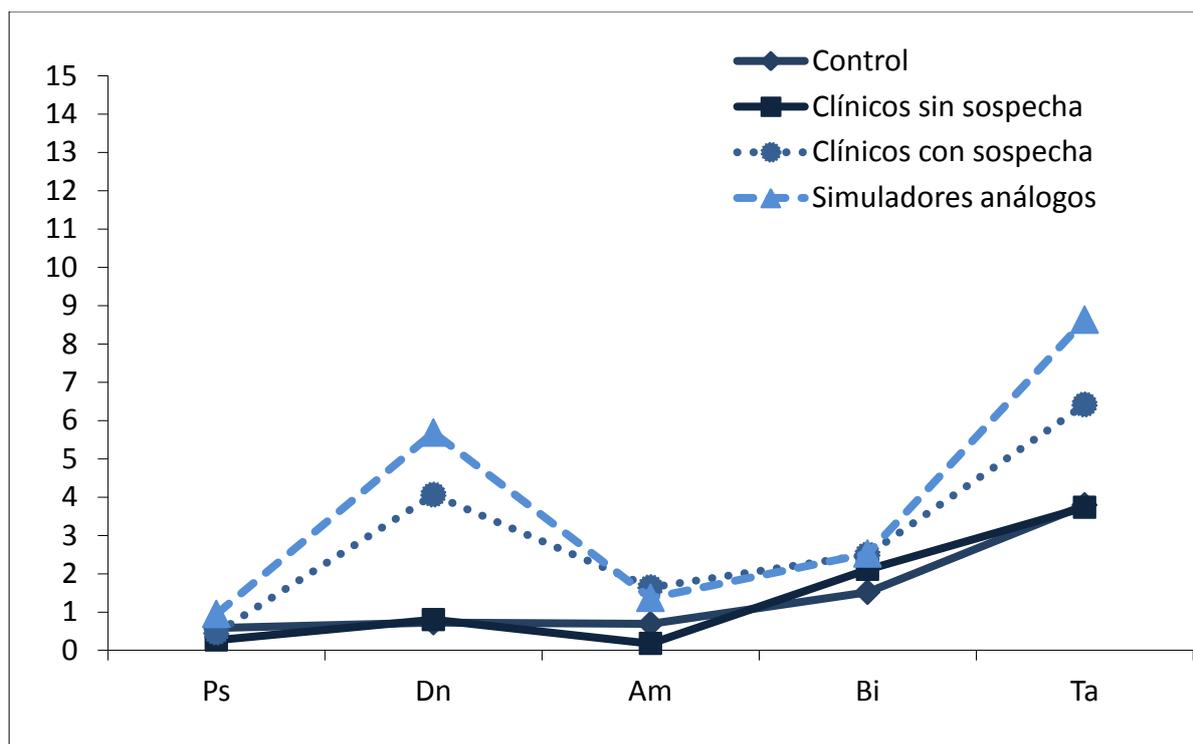


Figura 6.1. Puntuaciones medias para el grupo control, clínico sin sospecha de simulación, clínico con sospecha de simulación y simulador análogo en las escalas del SIMS

6.3.2. Utilidad diagnóstica

Se estudió la utilidad diagnóstica de los distintos puntos de corte para la puntuación total del SIMS y las escalas Deterioro neurológico y Trastornos afectivos en la discriminación entre los grupos

clínicos sin y con sospecha de simulación (se excluyeron al grupo control y al grupo de simuladores análogos de estos estudios de utilidad diagnóstica). Se consideró un amplio rango de posibles puntos de corte sugeridos por investigaciones previas.

Se emplearon curvas COR para examinar la precisión clasificatoria de las escalas del SIMS a la hora de diferenciar el grupo con sospecha de simulación del grupo sin sospecha de simulación. Las curvas COR se calculan como función de la sensibilidad y la especificidad y su rendimiento global puede ser evaluado con el área bajo la curva (ABC). Adicionalmente, se calcularon el poder predictivo positivo y el poder predictivo negativo (PPP y PPN) estimados de acuerdo a diferentes tasas base de prevalencia (0,15, 0,30 y 0,50) tal y como es recomendado en la literatura (Sellbom et al., 2010). El área bajo la curva de la puntuación total del SIMS fue de 0,84 (ET=0,040), de la escala Deterioro neurológico fue de 0,83 (ET=0,040) y de la escala Trastornos afectivos 0,81 (ET=0,044). En las tablas 6.2, 6.3 y 6.4 se muestran la utilidad de los distintos puntos de corte propuestos en cada una de las escalas.

Tabla 6.2

Estadísticos de utilidad diagnóstica de la puntuación Total en el SIMS en la discriminación entre pacientes clínicos con y sin sospecha de simulación.

Punto de corte	Sen.	Esp.	%Gl	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
> 6	0,90	0,44	0,68	0,22	0,96	0,41	0,91	0,62	0,81
> 7	0,85	0,58	0,72	0,26	0,96	0,46	0,90	0,67	0,79
> 8	0,77	0,71	0,74	0,32	0,95	0,53	0,88	0,73	0,76
> 9	0,75	0,80	0,77	0,40	0,95	0,62	0,88	0,79	0,76
> 10	0,67	0,91	0,79	0,57	0,94	0,76	0,87	0,88	0,73
> 11	0,58	0,91	0,74	0,53	0,92	0,73	0,83	0,87	0,68
> 12	0,54	0,93	0,73	0,58	0,92	0,77	0,83	0,89	0,67
> 13	0,52	0,98	0,74	0,82	0,92	0,92	0,83	0,96	0,67
> 14	0,46	0,98	0,71	0,80	0,91	0,91	0,81	0,96	0,64
> 15	0,38	1,00	0,68	1,00	0,90	1,00	0,79	1,00	0,62
> 16	0,33	1,00	0,65	1,00	0,89	1,00	0,78	1,00	0,60

Nota. Sen. = sensibilidad, Esp. = especificidad, %Gl = porcentaje global de clasificaciones correctas, TP = Tasa de prevalencia, PPP = poder predictivo positivo, PPN = poder predictivo negativo.

Tabla 6.3

Estadísticos de utilidad diagnóstica de la escala Deterioro neurológico (Dn) en la discriminación entre pacientes clínicos con y sin sospecha de simulación.

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Dn > 0	0,90	0,53	0,72	0,25	0,97	0,45	0,92	0,66	0,84
Dn > 1	0,73	0,80	0,76	0,39	0,94	0,61	0,87	0,78	0,75
Dn > 2	0,52	0,91	0,71	0,51	0,92	0,72	0,82	0,85	0,66
Dn > 3	0,48	0,98	0,72	0,79	0,91	0,90	0,81	0,96	0,65
Dn > 4	0,44	0,98	0,70	0,78	0,91	0,89	0,80	0,95	0,63
Dn > 5	0,38	1,00	0,68	1,00	0,90	1,00	0,79	1,00	0,62

Nota. Sen. = sensibilidad, Esp. = especificidad, %GI = porcentaje global de clasificaciones correctas, TP = Tasa de prevalencia, PPP = poder predictivo positivo, PPN = poder predictivo negativo.

Tabla 6.4

Estadísticos de utilidad diagnóstica de la escala Trastornos afectivos (Af) en la discriminación entre pacientes clínicos con y sin sospecha de simulación.

Punto de corte	Sen.	Esp.	%GI	TP=15%		TP=30%		TP=50%	
				PPP	PPN	PPP	PPN	PPP	PPN
Ta > 2	0,98	0,24	0,62	0,19	0,99	0,36	0,96	0,56	0,92
Ta > 3	0,85	0,44	0,66	0,21	0,95	0,40	0,88	0,61	0,75
Ta > 4	0,75	0,73	0,74	0,33	0,94	0,55	0,87	0,74	0,75
Ta > 5	0,56	0,87	0,71	0,43	0,92	0,64	0,82	0,81	0,66
Ta > 6	0,48	0,96	0,71	0,66	0,91	0,82	0,81	0,92	0,65
Ta > 7	0,35	1,00	0,67	1,00	0,90	1,00	0,78	1,00	0,61

Nota. Sen. = sensibilidad, Esp. = especificidad, %GI = porcentaje global de clasificaciones correctas, TP = Tasa de prevalencia, PPP = poder predictivo positivo, PPN = poder predictivo negativo.

6.3.3. Relación con otras pruebas

Se analizó la correlación de las escalas del SIMS con las escalas del MMPI-2-RF obteniendo los resultados que se pueden observar en la tabla 6.5. Las escalas que mayores correlaciones mostraron con las escalas del MMPI-2-RF fueron, por orden, la puntuación total ($M r = 0,63$), Deterioro neurológico ($M r = 0,61$), Trastornos afectivos ($M r = 0,59$) y Trastornos amnésicos ($M r = 0,42$). Las escalas Psicosis y Baja inteligencia mostraron valores notablemente inferiores en sus correlaciones ($M r = 0,15$ y $M r = 0,17$ respectivamente).

Tabla 6.5

Correlaciones entre las escalas del SIMS y las escalas de exageración y minimización de síntomas del MMPI-2-RF

Escalas del SIMS	MMPI-2-RF Escalas de validez						
	Exageración de síntomas				Minimización de síntomas		
	F-r	Fp-r	Fs	FBS-r	RBS	L-r	K-r
Total SIMS	0,71**	0,53**	0,57**	0,66**	0,67**	0,00	-0,56**
Psicosis (Ps)	0,20*	0,21*	0,16	0,08	0,12	-0,05	-0,21*
Deterioro neurológico (Dn)	0,69**	0,46**	0,60**	0,64**	0,65**	0,04	-0,50**
Trastornos amnésicos (Am)	0,56**	0,38**	0,34**	0,37**	0,46**	-0,03	-0,42**
Baja inteligencia (Bi)	0,18*	0,14	0,16	0,22**	0,15	-0,01	-0,12
Trastornos afectivos (Af)	0,60**	0,51**	0,51**	0,67**	0,65**	0,00	-0,53**

Nota. N = 143. F-r = Infrecuencia; Fp-r = Psicopatología infrecuente; Fs = Quejas somáticas infrecuentes; FBS-r = Validez de los síntomas; RBS = Escala de sesgo de respuestas; L-r = Virtudes inusuales; K-r = Validez del ajuste; MMPI-2-RF = Inventario Multifásico de Personalidad de Minnesota-2 Reestructurado.

* p<0,05 ** p<0,01.

6.4. Discusión

Los ítems del SIMS fueron redactados originalmente para que resultaran inverosímiles o poco probables en muestras psicopatológicas o con trastornos neuropsicológicos, como han mostrado diversas investigaciones y se ha argumentado en el capítulo 3 de esta tesis. Esta investigación muestra que, adicionalmente, también resultan poco probables en poblaciones con dolor crónico lumbar, sin que se observen diferencias en las puntuaciones en el SIMS entre participantes de la población general sin patologías diagnosticadas y pacientes aquejados de dolor lumbar crónico que responden honestamente a la prueba. Los valores obtenidos en el SIMS en la muestra de pacientes clínicos con lumbalgia crónica son además similares a los encontrados en muestras de población general en la literatura científica (v.g., Cima, Hollnack, et al., 2003; Clegg et al., 2009; Edens et al., 1999; Gillard, 2010; Rogers et al., 2005; Widows y Smith, 2005) o en la propia muestra control del capítulo 5 de esta tesis.

Las diferencias en las puntuaciones del SIMS aparecen en aquellos sujetos con sospecha de simulación de lumbalgia, o instruidos específicamente para simular, que se diferencian significativamente tanto del grupo control de población general como del grupo de pacientes clínicos con lumbalgia. Además, el actual conocimiento de las puntuaciones medias en el SIMS en diversos trastornos también permite contrastar de forma adicional cuando una determinada elevación es más propia de otras patologías de mayor gravedad o mayor afectación. Así, estos grupos con sospecha de simulación de lumbalgia manifiestan un nivel de sintomatología psicopatológica y neurológica en el SIMS incluso superior al de muestras clínicas psiquiátricas referidas en la literatura científica

(Dandachi-FitzGerald et al., 2011; Merckelbach y Smith, 2003; Peters, Jelicic, et al., 2013) o con puntuaciones superiores en la escala Deterioro neurológico a pacientes clínicos con epilepsia (Benge et al., 2012). Estos datos de contraste provenientes de otras patologías e investigaciones es un dato adicional al diseño de grupos diagnósticos que refuerza la capacidad discriminativa de simulación del SIMS en lumbalgia crónica.

Las diferencias señaladas en el SIMS aparecen particularmente en la puntuación total y en las escalas Deterioro neurológico y Trastornos afectivos, aquellas más plausiblemente asociadas a la exageración de sintomatología de dolor crónico. Este conjunto de dos escalas, Deterioro neurológico y Trastornos afectivos, junto con la puntuación total, han sido las que mayor capacidad discriminativa han mostrado en esta investigación para diferenciar al grupo de pacientes clínicos con y sin sospecha de simulación, con tamaños del efecto grandes. Estos resultados son concordantes con los obtenidos en investigaciones previas con el SIMS, donde estas dos escalas junto con la puntuación total eran las que mayores tamaños del efecto obtenían en contexto de exageración de síntomas médicos (Benge et al., 2012; Merten et al., 2008).

Adicionalmente, estas escalas han sido las que mayores correlaciones han obtenido con las escalas del MMPI-2-RF y particularmente con aquellas escalas más efectivas en la detección de simulación de dolor, como pueden ser F-r, Fs, FBS-r y RBS (Wygant et al., 2011), obteniendo valores altos de correlación y respaldando de forma convergente su capacidad discriminativa.

El estudio de los posibles puntos de corte permite conocer los valores de utilidad diagnóstica de distintas puntuaciones tanto para el total como para las dos escalas mencionadas. Preliminarmente se podrían considerar como orientativos en este campo el punto de corte de >9 para la puntuación Total, > 2 para la escala Deterioro neurológico y >5 en Trastornos afectivos. Estos puntos de corte son sustancialmente inferiores a los señalados para la detección de simulación de psicopatología tanto en los estudios originales de Smith y Burger (1997) como en los realizados posteriormente. Esto es debido a que en la patología lumbar crónica el nivel de sintomatología basal psicopatológica, neurológica y cognitiva es sensiblemente menor que en otros trastornos psicopatológicos, tal y como cabría esperar. Esta necesidad de ajustar los puntos de corte al nivel de sintomatología basal y al contexto de evaluación ha sido señalado previamente por diferentes investigadores (Santamaría y González-Ordi, 2012; Wygant et al., 2007, 2010).

Este es el primer estudio en la literatura científica que aborda específicamente la capacidad del SIMS para detectar simulación de dolor crónico empleando un diseño de grupos conocidos. Serán necesarios estudios posteriores que confirmen estas evidencias con muestras más amplias y diversas,

así como que repliquen la posible utilidad de los puntos de corte sugeridos. El tamaño muestral empleado en este estudio (45 y 48 en los grupos clínico y simulador respectivamente) afecta notablemente a la precisión de las estimaciones de sensibilidad y especificidad realizadas. Por lo tanto, los puntos de corte sugeridos deberán ser considerados con cautela en espera de investigaciones adicionales con muestras más amplias que permitan replicar o refutar los resultados obtenidos. Serán necesarios también estudios futuros que aborden la capacidad del SIMS para detectar simulación en otras áreas relevantes de dolor crónico, como el esguince cervical o la fibromialgia, señaladas por los expertos como con altas tasas de prevalencia de simulación.

En su conjunto, los resultados obtenidos permiten disponer de evidencia preliminar de la utilidad del SIMS como herramienta de *screening* de apoyo en el campo médico-legal en la detección de simulación de dolor crónico lumbar. Estos resultados se añaden a las investigaciones previas realizadas con el SIMS en el campo médico y que también arrojaban resultados satisfactorios como el estudio de Benge et al. (2012) en epilepsia, de van Beilen et al. (2009) en trastornos neurológicos o de Merten et al. (2008) en esguince cervical.

6.5. Conclusiones

El correcto diagnóstico y tratamiento del dolor crónico musculoesquelético en los países occidentales es uno de los retos más importantes existentes a nivel social, económico y laboral. No en vano se considera que es la primera causa de baja laboral en los estudios realizados por la OMS (Alonso et al., 2011) o por la Agencia Europea de Salud y Seguridad en el trabajo (European Agency for Safety and Health at Work, 2010).

Dentro de los trastornos musculoesqueléticos destaca por su relevancia el dolor lumbar que específicamente en nuestro país es considerada la enfermedad con mayor prevalencia en población adulta de más de 20 años y con un significativo consumo de recursos, de acuerdo a datos de la Sociedad Española de Reumatología (Carmon, Gabriel, Ballina y Laffon, 2001). Los costes económicos asociados a lumbalgia a nivel europeo se han estimado que supongan entre el 1,7% y el 2,1% del producto interior bruto (Ekman, Johnell y Lidgren, 2005a). Así, por ejemplo, se calcula que en Alemania el coste por dolor lumbar es superior a 7.000 euros anuales por paciente, suponiendo el absentismo laboral causado por la dolencia el 75% de los costes por paciente (Juniper, Le y Mladi, 2009).

Su alta prevalencia y relativa inespecificidad hacen que sea también una de las dolencias con mayores tasas de simulación, tal y como se indicaba en el primer estudio empírico de esta tesis, con estimaciones cercanas al 50% de simulación en contexto médico-legal, hecho que reclama el empleo

de herramientas de evaluación de la presencia de simulación de dolor (Greve, Bianchini y Brewer, 2013).

Por su brevedad de aplicación y el amplio espectro de áreas de simulación evaluadas (física, emocional, cognitiva) el SIMS fue seleccionada como una herramienta potencialmente útil a este fin. Estudios previos en la literatura científica habían abordado su utilidad en detección de simulación psicopatológica, cognitiva y médica (Alwes, 2006; Benge et al., 2012, 2012; Clegg et al., 2009; Merten et al., 2008; van Beilen et al., 2009; Vitacco et al., 2006), si bien no había habido ningún estudio que abordara particularmente la capacidad del SIMS para detectar la simulación de dolor lumbar.

Los resultados de este estudio, el primero en emplear un diseño de grupos diagnósticos con el SIMS en detección de dolor crónico, han mostrado valores satisfactorios en su capacidad discriminativa entre grupos sin sospecha de simulación y con sospecha de simulación, con tamaños del efecto grandes tanto en la puntuación total del SIMS como en las escalas Deterioro neurológico y Trastornos afectivos. También han mostrado una adecuada convergencia con los resultados en las escalas de exageración de síntomas del MMPI-2-RF, una de las pruebas de referencia en la evaluación de simulación de quejas médicas (Anderson, 2011; Greiffenstein et al., 2013; Sellbom et al., 2012; Wygant et al., 2011).

Estos resultados han permitido el establecimiento de puntos de corte preliminares para estas puntuaciones en la determinación de sospecha de simulación, aunque deberán ser considerados con cautela dadas las limitaciones ya señaladas previamente en el apartado de discusión.

Si bien serán necesarios futuros estudios que repliquen los resultados aquí obtenidos con muestras más amplias y variadas, existen evidencias preliminares que respaldan el uso del SIMS en el cribaje de sospechas de simulación en dolor lumbar. El uso de esta herramienta puede contribuir a un mejor manejo de los recursos limitados del sistema de salud y determinar su correcta asignación a aquellos pacientes que lo necesiten.

7. Discusión general y conclusiones

En septiembre de 2011, la Organización de las Naciones Unidas (ONU) celebró la Cumbre Mundial sobre Enfermedades Crónicas, donde se reconoció que las enfermedades crónicas, entre ellas los trastornos mentales y los trastornos por dolor crónico, suponen un desafío de "proporciones epidémicas", con un grave impacto en el desarrollo socio-económico de los países (World Health Organization, 2011). Se trata de la segunda vez en la historia de la ONU que la Asamblea General convoca una cumbre para abordar un tema de salud (el otro tema de salud tratado ha sido el SIDA), por lo que se pueden deducir la magnitud y las repercusiones de estas patologías a nivel mundial, cuyo coste asociado se prevé que supere los 47 billones de dólares en los próximos 20 años de acuerdo a la previsiones del *Global Economic Burden of Non-communicable Diseases* (Bloom et al., 2012).

La OMS ha alertado de que el aumento de las enfermedades crónicas representa un enorme reto para los gobiernos, anunciando que para algunos países no es exagerado describir la situación como un desastre inminente a nivel de salud, social y económico (World Health Organization, 2011). El Consejo de la Unión Europea aprobó, en 2011, una serie de conclusiones dirigidas a disminuir el impacto de las enfermedades crónicas en los sistemas sanitarios y solicitando a los Estados miembros y a la Comisión Europea que implementen de manera urgente medidas concretas y coordinadas que permitan hacer frente a este problema (Consejo de la Unión Europea, 2011).

Dentro de las enfermedades crónicas destacan por su prevalencia aquellas relacionadas con los trastornos mentales y con los dolores musculoesqueléticos. En concreto, los problemas de salud mental ocupan el primer puesto de la carga económica derivada de las enfermedades más frecuentes (Bloom et al., 2012; Delclós et al., 2012) y constituyen el 13% de la carga global de enfermedades (Collins et al., 2011). En 2010, los trastornos mentales supusieron un coste económico para el total de la Unión Europea de 798.000 millones de euros. En España, los trastornos mentales suponen la segunda causa de baja laboral y el tercer grupo de enfermedades que genera más gasto al sistema sanitario público con un coste económico anual estimado de 7.019 millones de euros, cerca del 1% del Producto Interior Bruto (PIB) de nuestro país (Oliva-Moreno, López-Bastida, Montejo-González, Osuna-Guerrero y Duque-González, 2009).

En el caso de los dolores musculoesqueléticos los datos son igualmente alarmantes, particularmente en el caso de la lumbalgia. La lumbalgia es la segunda queja más prevalente en la

población general, por detrás del resfriado, estando cerca del 15% de las bajas laborales asociadas a trastornos lumbares y suponiendo un gasto anual en España cercano a los 6.000 millones de euros incluyendo gastos directos e indirectos debidos a pérdidas de productividad (López Pascual, Peydro de Moya y Garrido Jaén, 2009). En otros países de nuestro entorno los datos son igualmente llamativos y alertan del alto coste que supone este trastorno (Asche, Kirkness, McAdam-Marx y Fritz, 2007; Dagenais, Caro y Haldeman, 2008; Ekman, Johnell y Lidgren, 2005b; Ekman, Jönhagen, Hunsche y Jönsson, 2005; Maetzel y Li, 2002). La lumbalgia está considerada la primera causa de baja laboral y la que genera un mayor desembolso por compensaciones económicas tanto al ente público como a las compañías aseguradoras (González Viejo, 2001). Se estima que en España dos millones de personas no acuden a su puesto de trabajo al menos 1 día cada seis meses por quejas de dolores lumbares (Humbría Mendiola, Carmona, Peña Sagredo y Ortiz, 2002).

El grave impacto económico y social de estas enfermedades crónicas reclama la gestión eficiente de los recursos. Un paso esencial es el desarrollo de políticas centradas en el paciente para la promoción de la salud, la prevención primaria y secundaria, el tratamiento y la atención de las enfermedades crónicas, tal y como insta a hacer el Consejo de la Unión Europea en sus directrices de actuación (Consejo de la Unión Europea, 2011). Otro paso complementario es la correcta identificación de la posible simulación en las incapacidades temporales y permanentes por enfermedades crónicas.

Los estudios internacionales suelen establecer porcentajes de simulación de trastornos cercanos al 30% en las solicitudes de bajas laborales, incapacidades o indemnizaciones, oscilando entre el 10 y el 50% según el estudio (Ardolf et al., 2007; Brooks et al., 2012; Chafetz, 2008; Gervais et al., 2001; Greene, 2008; Greiffenstein et al., 2013; Greve et al., 2009; Jorge y Jorge, 2011; Kay y Morris-Jones, 1998; Larrabee, 2003b; Leavitt y Sweet, 1986; Miller, 2000; Mittenberg et al., 2002; Norris y May, 1998; Schmand et al., 1998; Stevens et al., 2008, 2008; Van Egmond et al., 2005; Van Hout et al., 2006; Walters et al., 1988), demostrando la relevancia potencial que puede tener este posible fraude en el monto económico y social indicado.

A pesar de la relevancia de la simulación a nivel económico y profesional, apenas existe investigación en España que indague sobre la percepción de los profesionales sobre cuál es su prevalencia, a qué perfil sociodemográfico suele vincularse y qué evidencias se emplean para su detección. Conocer esta información se consideró un primer paso necesario para poder abordar las áreas donde se percibía mayor necesidad de instrumentos para la detección de simulación y complementar las estrategias ya empleadas por los profesionales.

Los resultados obtenidos en esta tesis a partir de la percepción de los profesionales indican que, lejos de observar tasas inferiores de simulación en nuestro contexto, los valores obtenidos son iguales o superiores a los obtenidos en la literatura internacional destacando la mayor prevalencia de la simulación de trastornos musculoesqueléticos y trastornos afectivos o emocionales como se ha observado en otras investigaciones internacionales (McDermott y Feldman, 2007; Mittenberg et al., 2002).

En los análisis realizados en el primer estudio empírico de esta tesis han destacado como los trastornos percibidos como más simulados los relacionados con dolores crónicos (esguince cervical, lumbalgia, cervicalgia y fibromialgia) y con trastornos de la ansiedad y del estado de ánimo (depresión), todos ellos con porcentajes estimados de simulación cercanos o superiores al 50%. Este conjunto de patologías señaladas como con mayores tasas de simulación se caracterizan por su carácter crónico y perdurable en el tiempo, su marcada subjetividad, su limitada evidencia objetiva y su mayor prevalencia en la población. Estas características permiten partir de una sintomatología probablemente genuina en su origen (por su alta prevalencia en la población) que posteriormente es alargada en el tiempo como una simulación parcial debido a su carácter subjetivo y crónico y con el objetivo de mantener los beneficios económicos o sociales relacionados y establecidos en nuestro sistema legal (Blasco Saiz y Pallardo Durá, 2012).

Las patologías percibidas con menores tasas de simulación y valores estimados inferiores al 35% fueron las patologías dolorosas agudas (cervicalgia y lumbalgia aguda), los traumatismos craneoencefálicos (leve y moderado) y la fiebre de origen desconocido y colon irritable. Estas patologías incumplen alguna de las tres características señaladas en las patologías más simuladas, o bien son agudas, o bien tienen mayores evidencias objetivas o bien tienen una menor prevalencia en la población y, por lo tanto, se desconocen más sus síntomas y no dan lugar a una simulación parcial o persistencia de la presentación sintomática una vez finalizada la sintomatología genuina.

Los resultados del estudio también han mostrado una marcada asociación de la simulación a un perfil sociodemográfico ajustable al modelo adaptativo de simulación (González-Ordi et al., 2012) en el que la presencia de condiciones sociolaborales adversas aumenta la presencia de simulación. Así los profesionales encuestados consideran más prevalente la simulación conforme menores son los ingresos, en medianas edades (particularmente 46-55 años, habitualmente con mayores cargas familiares y sociales asociadas y más beneficiadas con disponer de una baja laboral) y en profesiones menos estimulantes y vocacionales. Estos resultados concuerdan con múltiples estudios de la literatura científica y manuales que asocian la presencia de la simulación a la relevancia del incentivo

económico o la satisfacción con el trabajo, entre otros factores (Aurrekoetxea Agirre et al., 2009; Barraquer, 1992; Benavides et al., 2007; Binder y Rohling, 1996; Cassidy et al., 2000; Castillo et al., 2003; Catalina-Romero et al., 2013; Fayad et al., 2004; Fransen et al., 2002; Frymoyer y Cats-Baril, 1987; Frymoyer, 1992; Gjesdal et al., 2004, p. 200; Karlsson et al., 2008; Lakke, Soer, Takken y Reneman, 2009; Mendiola, 2004; Øverland et al., 2008; Pincus, Burton, Vogel y Field, 2002; Quezada-Ortega et al., 2006; Robinson y Loeser, 2012; Steenstra et al., 2005).

Así, los profesionales perciben las mayores tasas de simulación asociadas a personas de mediana edad con trabajos poco estimulantes y bajos o moderados ingresos, que comienzan sufriendo originalmente algún tipo de lesión o trastorno y deciden prologarlo en el tiempo para obtener beneficios a nivel económico o laboral. Los trastornos más propensos a simulación serán aquellos más prevalentes en la población general, más subjetivos y difíciles de verificar objetivamente y más proclives al carácter crónico.

Ante esta acuciante necesidad profesional, llama la atención la heterogeneidad y falta de procedimientos sistemáticos y protocolizados entre los profesionales encuestados. Así, ninguno de los encuestados mencionó en ningún caso el uso de pruebas o protocolos objetivados o sistematizados como vías de sospecha de simulación o exageración, confiando fundamentalmente en criterios más subjetivos y relacionados esencialmente a los criterios de sospecha del DSM (discrepancia información objetiva y subjetiva, falta de cooperación, contexto médico-legal, aspectos antisociales...), criterios que han sido duramente criticados por diversos autores por su subjetividad y por el alto número de falsos positivos que pueden ir asociados (p.ej., Berry y Nelson, 2010; Gerson, 2002; Rogers, 2008a).

Tal y como se ha argumentado en esta tesis, los estudios disponibles hasta la fecha han remarcado de forma insistente la falibilidad del “ojo clínico” en la detección de simulación (p. ej., Aamodt y Custer, 2006 o DePaulo, Charlton, Cooper, Lindsay y Muhlenbruck, 1997) y la necesidad de emplear protocolos sistemáticos y objetivos que determinen la correcta asignación y eviten la comisión tanto de falsos positivos como de falsos negativos. Esto supone, en terminología de Rogers y Bender (2003), pasar de un primer estadio evolutivo de detección de la simulación basado en la “intuición” a un segundo o tercer estadio que emplea medidas específicas desarrolladas para la evaluación y determinación del grado de simulación.

Sin duda la labor de valoración de las incapacidades temporales y permanentes debidas a estas enfermedades resulta un reto lleno de complejidades y dificultades (Amengual, Puig Bausili, Gómez Montoro y Jiménez Jiménez, 2006) pero no es menos cierto que la introducción de protocolos y medidas objetivas de detección de simulación permitiría una mejora de la calidad de la evaluación y

algunos estudios realizados con la inclusión de métodos objetivos de detección del posible fraude ya se han mostrado como una fuente notable de eficiencia económica (v.g. López Pascual et al., 2009).

El uso de protocolos objetivos y sistemáticos permite, por una parte, el estudio científico de sus propiedades, su replicación, contraste y posible mejora o sustitución por otro protocolos, así como se evita la intromisión de posibles variables subjetivas y contaminadoras que la investigación ha mostrado como graves sesgos perceptivos en la detección de mentira (p. ej., véase el trabajo del equipo de Vrij en esta línea: Vrij et al., 1992, Vrij, 2000, 2004).

Los sesgos presentes en la evaluación subjetiva de la veracidad de los síntomas hacen necesaria la evaluación detallada y protocolizada como pieza fundamental en la detección del engaño o en la determinación del no engaño. En este proceso han jugado y juegan un papel fundamental los autoinformes como herramienta sistemática de recogida de información y toma de decisiones (Capilla Ramírez y González-Ordi, 2009; González-Ordi y Gancedo Rojí, 1999; González-Ordi et al., 2012; Rogers, 2008c; Sellbom y Bagby, 2008).

Si bien hasta la fecha en España se han realizado diversos estudios de validación de pruebas de detección de simulación en distintas áreas (v.g., Arce et al., 2009; Burton et al., 2012; Sánchez Crespo et al., 2008; Sánchez Crespo, et al., 2012; Zaldivar Basurto et al., 2008a, 2008b), no ha habido ninguna investigación realizada con pruebas de autoinforme de *screening* de simulación combinando diseños análogos, de grupos diagnósticos y de medidas autosuficientes.

Esto es especialmente importante ya que, aunque las pruebas generales han mostrado satisfactorios resultados, no son menos ciertas sus limitaciones, entre ellas el tiempo de aplicación que requieren y la formación especializada que exige su aplicación e interpretación (Schinka y Borum, 1993, 1994). Además, la mayoría de estas pruebas evalúan la presencia de simulación en un determinado dominio de funcionamiento, lo que puede resultar problemático ya que es difícil determinar en qué área va a presentarse la simulación (cognitiva, psicopatológica o somática), algo que puede obligar a emplear desde un principio un amplio conjunto de pruebas extensas y complejas para poder hacer una barrido completo inicial de la posible simulación.

El SIMS (Smith y Burger, 1997; Widows y Smith, 2005) ha destacado dentro de estas pruebas de *screening* por la amplitud de áreas en que rastrea la presencia de simulación (Wisdom et al., 2010) así como la investigación acumulada en la literatura científica sobre su utilidad y capacidad discriminativa, tal y como se evidenciaba en el capítulo 3 de esta tesis. Los resultados disponibles en la literatura científica han mostrado adicionalmente que ha alcanzado tamaños del efecto iguales a

superiores a pruebas más largas y complejas (v.g., Edens et al., 2007; Laffoon, 2009; Vitacco et al., 2006). Como prueba de *screening* permite un acercamiento eficiente en términos de tiempo y esfuerzo, posibilitando un amplio rastreo de diversas áreas potenciales de simulación.

Los datos obtenidos con el SIMS en esta tesis han mostrado su utilidad como un primer *screening* de los diferentes dominios fundamentales de simulación. Así, se ha evaluado su capacidad discriminativa y utilidad tanto en la detección de simulación de psicopatología en diferentes áreas sintomatológicas (trastornos psicóticos, deterioro neurológico y trastornos afectivos) como en la detección de simulación de lumbalgia crónica. Para ello se han utilizado diferentes diseños de investigación, complementando estudios análogos con diseños autosuficientes, grupos diagnósticos y muestras clínicas de referencia. En todos los diseños analizados, se han observado tamaños del efecto grandes ($d > 1,25$) a la hora de diferenciar significativamente entre pacientes clínicos genuinos y simuladores a partir de la puntuación total del SIMS en las diferentes condiciones alegadas. Esto supone un claro respaldo a la capacidad de la puntuación total del SIMS para la detección de sospecha de simulación en población española en las diferentes áreas sintomatológicas estudiadas en esta investigación. Estos resultados concuerdan con los obtenidos en la literatura científica en diferentes diseños y áreas sintomatológicas con tamaños del efecto que oscilan entre 0,62 y 2,42 (Alwes, 2006; Buijs, 2011; Cima, Hollnack, et al., 2003; Clegg et al., 2009; Dearth, 2007; Efendov, 2007; González-Ordi et al., 2010; Lewis et al., 2002; Merckelbach y Smith, 2003; Merten et al., 2007; Vitacco et al., 2006; Wisdom et al., 2010).

Los resultados obtenidos en esta tesis han mostrado como es recomendable que la puntuación total del SIMS sea el punto de inicio de cualquier análisis de resultados ya que sistemáticamente es la que mayor capacidad discriminativa ha mostrado en los diferentes contextos y áreas de simulación, por lo que es una buena síntesis de la posible sospecha de simulación de forma global, independientemente del contexto específico o las áreas sintomáticas alegadas. Posteriormente, el análisis más detallado de cada condición realizado en los diversos estudios de esta tesis ha permitido advertir la sensibilidad de cada escala al contexto específico de simulación, así la escala Psicosis es la que mayor tamaño del efecto mostraba en la condición análoga de simulación de trastorno psicótico, la escala Trastorno afectivo en la condición de simulación de depresión, la escala Deterioro neurológico, junto con Trastornos afectivos, en la condición de simulación de consecuencias tras un accidente, o ambas escalas, Deterioro neurológico y Trastornos afectivos, en la simulación de lumbalgia, representando una combinación de quejas afectivas y somáticas poco creíbles. En cada una de estas condiciones mencionadas el tamaño del efecto de cada una de estas escalas era también grande ($d > 1,25$) a la hora de establecer diferencias estadísticamente significativas entre la muestra clínica de referencia y la muestra de simuladores. Estos resultados

avalan el posterior análisis detallado del perfil para poder detectar el área de quejas o síntomas poco creíbles que alega la persona evaluada y se suman a las diversas investigaciones en la literatura científica que han mostrado la relación entre determinadas escalas del SIMS y determinadas áreas de simulación de síntomas tanto en psicopatología (Alwes, 2006; Cima, Hollnack, et al., 2003; Clegg et al., 2009; Edens et al., 1999; G. P. Smith y Burger, 1997; Vitacco et al., 2006) como en problemas médicos o somáticos (Benge et al., 2012; Merten et al., 2008).

El análisis de los posibles puntos de corte a emplear para la detección de simulación en psicopatología ha mostrado la limitada especificidad de los propuestos originalmente por Smith y Burger (1997), algo ya señalado de forma sistemática en la literatura científica (Alwes, 2006; Cima, Hollnack, et al., 2003; Clegg et al., 2009; Dandachi-FitzGerald et al., 2011; Edens et al., 1999, 2007; Lewis et al., 2002; Merckelbach y Smith, 2003; Peters et al., 2012; Rogers, Hinds, et al., 1996; Vitacco et al., 2006; Wisdom et al., 2010), descartando mayoritariamente como punto de corte para la puntuación total el valor de >14 y optando por valores superiores como >16 (Rogers, Hinds, et al., 1996; Merckelbach y Smith, 2003; Lewis et al., 2002; Cima, Hollnack, et al., 2003), >19 (Clegg et al., 2009) o >23 (Wisdom et al., 2010).

Los resultados obtenidos en esta tesis en los estudios de detección de simulación en psicopatología también sugerirían tentativamente considerar las puntuaciones por debajo de 16 como indicativas de ausencia de simulación psicopatológica (≤ 16 "Paciente clínico, sin sospecha de exageración de síntomas de psicopatología") y puntuaciones por encima de 16 como indicadores de sospecha de simulación de psicopatología en diferentes niveles: >16 posible sospecha de exageración de síntomas; >23 probable sospecha de exageración de síntomas; >40 clara sospecha de exageración de síntomas, empleando los niveles de gradientes de certeza sugeridos por Berry y Nelson (2010).

El análisis de los resultados de utilidad diagnóstica de las escalas en la detección de simulación en psicopatología, arroja resultados similares a los comentados previamente para la puntuación total. Los puntos de corte sugeridos originalmente por los autores originales (>1 en Ps, >2 en Dn, Am y Bi y >5 en Ta) muestran una alta sensibilidad pero una muy baja especificidad. Este hecho ya ha sido observado también en las escasas investigaciones con muestras clínicas que han indagado en este aspecto tanto en contextos civiles como penales (Clegg et al., 2009; Vitacco et al., 2006). Los resultados obtenidos en esta tesis, sugieren preliminarmente un aumento de los puntos de corte para su uso en población española en la detección de exageración de síntomas de psicopatología. Serán necesarias investigaciones futuras que determinen exactamente los valores más recomendables, si bien tentativamente se podrían sugerir el punto de corte de >2 para la escala

Psicosis, de >3 para las escalas Deterioro neurológico, Trastornos Amnésicos y Baja inteligencia y >7 para la escala Trastornos afectivos. Estos puntos de corte permitirían una mejora de la especificidad, aun manteniendo su condición de puntos de cribaje que requieren la posterior confirmación por una batería más completa y exhaustiva.

Cabe señalar respecto a los estudios de detección de exageración de síntomas en psicopatología que los resultados hallados han mostrado oscilaciones en la sensibilidad y especificidad según la sintomatología alegada (trastorno psicótico, deterioro neurológico o trastorno afectivo) o el contexto (forense criminal o forense civil). Estos resultados posiblemente deriven en futuras investigaciones que establezcan puntos de corte diferentes según el contexto o área sintomatológica psicopatológica alegada tanto para la puntuación total como para las puntuaciones en las escalas. No obstante, dado el carácter preliminar de este estudio, la ausencia de investigación previa con el SIMS en esta dirección y la incertidumbre al determinar en la práctica qué sintomatología alegará una persona en su evaluación, hacen aconsejable no hacer una propuesta específica a partir de los resultados obtenidos en esta tesis y demorar su formulación a la espera de futuras investigaciones. Como señalaba (Alwes, 2006) esto debería ser una línea de investigación futura prioritaria con el SIMS.

En esta misma línea, los resultados obtenidos en esta tesis en la detección de exageración de síntomas de lumbalgia crónica claramente abogan por puntos de corte notablemente inferiores a los señalados anteriormente para detección de exageración de síntomas de psicopatología. Esto es debido a que en la patología lumbar crónica el nivel de sintomatología basal psicopatológica, neurológica y cognitiva es sensiblemente menor que en otros trastornos psicopatológicos, tal y como cabría esperar. Así, los valores obtenidos en el SIMS en la muestra de pacientes clínicos con lumbalgia crónica sin sospecha de simulación son similares a los encontrados en muestras de población general en la literatura científica (Cima, Hollnack, et al., 2003; Clegg et al., 2009; Edens et al., 1999; Gillard, 2010; Rogers et al., 2005; Widows y Smith, 2005) o en las diversas muestras de control analizadas en esta tesis. Las escasas investigaciones realizadas con el SIMS en contexto médico han mostrado valores basales más bajos a los obtenidos en contexto psicopatológico, tal y como cabría esperar (p. ej., van Beilen et al., 2009)

Los resultados obtenidos en esta tesis sugerirían considerar preliminarmente como puntos de corte para la sospecha de exageración de síntomas en lumbalgia crónica los valores de >9 para la puntuación Total, > 2 para la escala Deterioro neurológico y >5 en Trastornos afectivos. Dada la ausencia de investigación previa en esta área con el SIMS, se consideró conveniente no proporcionar diversos puntos de corte de acuerdo al nivel de certeza sino sugerir preliminarmente estos posibles

valores que deberán ser replicados en investigaciones futuras para ir articulando un cuerpo empírico que permita delimitar en un futuro los posibles niveles de certeza o interpretación de puntuaciones del SIMS en este contexto.

Adicionalmente a los estudios sobre tamaño del efecto y utilidad diagnóstica de los diferentes puntos de corte, se ha dispuesto también de estudios correlacionales entre el SIMS y el MMPI-2-RF tanto en el área psicopatológica como de lumbalgia crónica. Los resultados obtenidos han permitido evidenciar una alta convergencia entre ambas escalas, con correlaciones promedio elevadas tanto en el estudio de detección de exageración de síntomas de psicopatología ($M r = 0,66$) como en el de lumbalgia ($M r = 0,63$), de forma coherente con lo obtenido en otras investigaciones que han relacionado el SIMS con el MMPI-2 (Carter et al., 2009; Cima, Hollnack, et al., 2003; Heinze y Purisch, 2001; Leung, 2010; Widows y Smith, 2005).

A nivel de las escalas del SIMS, ha podido advertirse la convergencia entre las escalas específicas del SIMS y del MMPI-2-RF destinadas a la detección de áreas similares de exageración de síntomas. Así, destaca la correlación de la escala Psicosis del SIMS (Ps) con la escala Fp-r del MMPI-2-RF destinada a fines similares; la correlación de la escala Trastornos amnésicos del SIMS (Am) con la escala RBS del MMPI-2-RF, creada específicamente para la detección de exageración de problemas de memoria o la escala Deterioro neurológico del SIMS (Dn) con las escalas Quejas somáticas infrecuentes (Fs), problemas de memoria (RBS) y la escala de exageración de síntomas (F-r). Particularmente, en el estudio realizado con lumbalgia, las escalas del SIMS con mayor capacidad discriminativa, puntuación total, Deterioro neurológico y Trastornos afectivos, han sido las que mayores correlaciones han obtenido con las escalas del MMPI-2-RF más efectivas en la detección de simulación de dolor, como son F-r, Fs, FBS-r y RBS (Wygant et al., 2011), obteniendo valores altos de correlación y respaldando de forma convergente su capacidad discriminativa. Los resultados obtenidos en estos estudios correlacionales han mostrado una adecuada convergencia entre el SIMS y las escalas de exageración de síntomas del MMPI-2 y del MMPI-2-RF, las pruebas de referencia en la evaluación de exageración tanto en el área psicopatológica (Gervais et al., 2011; Marion et al., 2011; Rogers, Gillard, et al., 2011; Sellbom y Bagby, 2010; Sellbom et al., 2010; Wygant et al., 2009, 2010, 2011; Youngjohn et al., 2011) como en el área del dolor crónico (Anderson, 2011; Greiffenstein et al., 2013; Sellbom et al., 2012; Wygant et al., 2011).

En su conjunto, los resultados obtenidos en los diversos estudios empíricos que componen esta tesis permiten disponer de evidencia preliminar de la utilidad del SIMS como herramienta de *screening* de exageración de síntomas psicopatológicos y de lumbalgia en población española. Estos

resultados se añaden a las investigaciones previas realizadas con el SIMS internacionalmente tanto en el campo de la psicopatología (Alwes, Clark, Berry y Granacher, 2008; Clegg et al., 2009; Dandachi-FitzGerald et al., 2011; Dearth, 2007; Edens et al., 2007; Efendov, 2007; Heinze y Purisch, 2001; Laffoon, 2009; Lewis et al., 2002; Merckelbach y Smith, 2003; Peters, Jelicic, et al., 2013; G. P. Smith y Burger, 1997; Vitacco et al., 2006; Wisdom et al., 2010) como en el campo médico (Benge et al., 2012; Merten et al., 2008; van Beilen et al., 2009) y que también arrojaban resultados satisfactorios.

El impacto de disponer de una herramienta de detección de exageración de síntomas, va más allá de la detección de simulación. Tal y como señalan Boone (2007) y Berry y Nelson (2010) nuestras herramientas difícilmente podrán indicarnos si es un intento consciente de obtención de beneficios externos o, por el contrario, es una somatización o intento inconsciente de llamada de atención y auxilio, en todo caso lo que sí hace es advertirnos de la escasa credibilidad de los síntomas. Así, es interesante señalar la amplia presencia de exageración de síntomas vinculada a muy diferentes aspectos y su impacto en el tratamiento y la mejora de estos pacientes. Independientemente de cuál sea el origen o causa de dicha exageración, su correcta detección puede ayudarnos a orientar el tratamiento y evitar esfuerzos terapéuticos inútiles y costosos pero con poca probabilidad de éxito.

A este respecto cabe volver a citar el estudio de Van Egmond et al. (2005) en el que un 42% de los pacientes clínicos albergaban expectativas de beneficios específicos diferentes de la propia mejoría, mostrando este grupo una menor mejoría que el resto que no albergaba dichas expectativas o beneficios secundarios. Lo más interesante en este caso era que los clínicos que trataban estos pacientes desconocían en su mayoría la presencia de estas expectativas y solo experimentaban las dificultades de avance con este grupo de pacientes. Este estudio resalta tanto el amplio abanico de las ganancias secundarias ya no propiamente la simulación, su efecto en el tratamiento y el escaso conocimiento y consciencia por parte de los profesionales que tratan con los pacientes. Serán deseables investigaciones futuras que puedan ahondar en el papel del SIMS como indicador más general de posibles ganancias secundarias que retrasen el progreso en el tratamiento y que puedan resultar útiles al clínico, más allá del marco concreto de la detección de la simulación.

Son muchas las limitaciones de los estudios presentados en esta tesis, tal y como se han señalado en los diferentes apartados de discusión de cada estudio y son también muchos los interrogantes que quedan pendientes de responder al respecto del SIMS y que serán objeto de líneas de investigación futura. En primer lugar será necesaria una línea de estudios análogos que investigue el efecto del entrenamiento y la preparación de los participantes en la capacidad discriminativa del SIMS. La disposición de tiempos más amplios de preparación, incluyendo información sobre las estrategias empleadas para detectar a los simuladores, el acceso libre a Internet como vía de

obtención de información y la introducción de notables incentivos positivos y negativos pueden replicar condiciones más análogas al contexto de simulación real y permitir obtener información sobre su efecto diferencial en la capacidad discriminativa del SIMS. El trabajo reciente de Veltri y Williams (2013) sobre el efecto interactivo del tipo de entrenamiento y la sintomatología simulada junto con el trabajo del grupo de Jelicic con el SIMS pueden resultar puntos de partida interesantes (Jelicic et al., 2011, 2006; Jelicic, Merckelbach y van Bergen, 2004; Jelicic et al., 2013; Jelicic, Merckelbach, et al., 2007; Merckelbach et al., 2010; Merckelbach, Smeets y Jelicic, 2009).

Otra segunda línea de investigación futura será la realización de estudios análogos y con muestras clínicas de referencia en el área de problemas cognitivos, uno de los aspectos no abordados en los estudios realizados en esta tesis. Adicionalmente, será deseable la inclusión de diseños autosuficientes o con grupos diagnósticos que complemente los estudios análogos en esta área, así como su correlación con pruebas bien establecidas y adaptadas en España para la detección de simulación de problemas de memoria como el TOMM (Tombaugh, 1996).

Una tercera línea de investigación será la realización de estudios con grupos diagnósticos tanto en el contexto forense civil como, particularmente, en el campo penal. Los estudios en contexto forense civil podrán delimitar en qué medida se replican los resultados hallados en esta tesis a partir del diseño autosuficiente que fue empleado; los estudios en contexto forense penal con grupos diagnósticos podrán delimitar la utilidad del SIMS en esta área. En ambos casos los resultados que se obtengan permitirán establecer la necesidad o conveniencia de ajustar los puntos de corte sugeridos del total y de sus escalas a cada contexto o sintomatología alegada. Asimismo podrá resultar interesante replicar el estudio de percepción de los profesionales sobre simulación con psicólogos forenses civiles y penales, así como con neuropsicólogos en ambos campos, para evidenciar qué patologías se perciben como más simuladas en dichos contextos, con qué frecuencia se estiman que ocurren y qué perfil sociodemográfico y estrategias de detección se emplean. El contraste con los resultados obtenidos en esta tesis a partir de médicos permitirá obtener un panorama más ajustado y global en población española.

Una cuarta línea será la extensión de la investigación realizada a otras patologías de dolor crónico con alta prevalencia de simulación como el esguince cervical o la fibromialgia. Ya se han comenzado el diseño de diferentes diseños de investigación en estas áreas que esperamos aporten luz a la utilidad del SIMS en estas patologías y su integración dentro de protocolos de detección de exageración de síntomas en estos contextos. También será objeto de esta línea de investigación adentrarse en la compleja relación entre los trastornos somatomorfos, facticios y la simulación y el

papel del SIMS en esta tarea, siguiendo muchas de las reflexiones planteadas recientemente por Merten y Merckelbach (2013)

Finalmente, el uso de las medidas de simulación en adolescentes sigue siendo un área necesitada de investigación (Sweet y Guidotti Breting, 2013). En particular, solo ha habido dos estudios que estudiaran específicamente la utilidad del SIMS en esta población (Dearth, 2007; Rogers, Hinds, et al., 1996). Serían necesarios más estudios que aborden su capacidad discriminativa y la conveniencia de ajustar los puntos de corte sugeridos hasta la fecha para su empleo en adultos.

Si bien serán necesarios futuros estudios que repliquen los resultados aquí obtenidos con muestras más amplias y variadas y que lleven a cabo las líneas de investigación señaladas previamente, los resultados aportados en esta tesis ofrecen evidencias que respaldan el uso del SIMS en el cribaje de sospechas de simulación de psicopatología y dolor lumbar. El uso de esta herramienta puede contribuir a un mejor manejo de los recursos limitados del sistema de salud y determinar su correcta asignación a aquellos pacientes que lo necesiten.

Extended summary

In September 2011, the United Nations Organization (UN) held the World Summit on Chronic, Non-communicable Diseases, where it was acknowledged that chronic conditions, among them mental disorders and chronic pain disorders, constituted a challenge of "epidemic proportions", with serious impact on the socio-economic development of countries (World Health Organization, 2011). This was only the second time in the history of the UN that the General Assembly has organized a summit to address a health issue (the first was on the subject of AIDS). From this we can deduce the worldwide magnitude and repercussions of these pathologies, whose associated cost will top 47 billion dollars in the next twenty years according to the forecasts of the *Global Economic Burden of Non-communicable Diseases* (Bloom et al., 2012).

The severe economic and social impact of these chronic conditions brings an urgent demand for the efficient management of resources. An essential measure is the development of patient-centred policies for the promotion of health, of primary and secondary prevention, and of the treatment and care of chronic diseases, as urged by the Council of Europe in its action guidelines (Council of the European Union, 2011); a complementary step would involve the correct identification of possible malingering with regard to temporary and permanent incapacity due to chronic conditions.

International studies generally report percentages of malingering close to 30% in applications for sick leave, work incapacity or compensation, though the figures range from 10% to 50% depending on the study (Chafetz, 2008; Gervais et al., 2001; Mittenberg et al., 2002; Schmand et al., 1998; Stevens et al., 2008, 2008). This is an indication of the potential relevance of such fraud in relation to the estimated economic and social costs referred to above.

However, despite the relevance of malingering at the economic and work-related level, there has been scarcely any research in Spain on professionals' perceptions about its prevalence and the socio-demographic profile with which it tends to be associated, or indeed on the kinds of evidence they consider for its detection. Gathering information on these aspects was a first objective of this dissertation, and a necessary one for being able to identify the areas with the greatest perceived need for instruments to detect malingering and to complement the strategies already employed by professionals.

With this aim, the research was based on a sample of 161 doctors from 26 Spanish provinces, who were given a questionnaire regarding which pathologies they perceived as most commonly feigned, the typical socio-demographic profile of malingerers (age, sex, income, occupation, etc.), and the strategies they used for detecting malingering.

The results obtained in this first empirical study revealed that malingering rates in Spain – always according to the perception of professionals –, far from being lower than those of other countries, were actually equal to or higher than those reported in the international literature. A notable finding was the greater prevalence of malingering in relation to musculo-skeletal disorders and affective or emotional disorders.

The analyses carried out reveal that the conditions perceived as most commonly feigned are those related to chronic pain (neck sprain, lower-back pain, neck pain and fibromyalgia) and anxiety or mood disorders (depression), all with estimated percentages, according to the doctors' perceptions, of close to or over 50%. The pathologies perceived with lower rates of malingering – under 35% – were those involving acute pain (acute neck and lower-back pain), traumatic brain injury (mild and severe), fever of unknown origin and irritable bowel syndrome.

The results of the first empirical study also revealed a marked association between malingering and a socio-demographic profile in line with the adaptive model of malingering (González-Ordi et al., 2012), in which the presence of adverse social and working conditions increases the likelihood of malingering. Thus, the doctors in the survey considered malingering to be more prevalent the lower the person's income, in middle-aged individuals (particularly in the age range 46-55, typically with associated family and social burden and more likelihood of benefiting from being on sick leave) and in those working in less stimulating or vocational occupations. According to these professionals' perceptions, prototypical malingerers would be characterized by having some original injury or disorder and deciding to simulate its prolongation over time in order to obtain financial or work-related benefits.

Given the pressing need of professionals to assess malingering, the results of the third part of the survey – referring to the strategies doctors use for detecting it – are quite surprising and remarkable, revealing a marked lack of systematic and protocolized procedures, according to the professionals in this sample. None of the doctors mentioned the use of objective or systematic tests or protocols for the detection of malingering or exaggeration, trusting basically in criteria that were more subjective, and essentially related to the DSM suspicion criteria (discrepancies between objective and subjective information, lack of cooperation, medico-legal context, antisocial aspects, etc.) – criteria that have been strongly criticized by various authors for their subjectivity and for the

large numbers of false positives they can generate (e.g., Berry & Nelson, 2010; Gerson, 2002; Rogers, 2008a).

Assessment of temporary and permanent incapacity due to injury or illness is undoubtedly a complex and difficult task (e.g., Amengual et al., 2006); however, it is no less true that the introduction of protocols and objective measures for the detection of malingering would help to improve the quality of such assessment.

Although several studies have been carried out in Spain for the validation of tests for detecting malingering in various contexts (e.g., Arce et al., 2009; Burton et al., 2012; Sánchez Crespo et al., 2008, 2012), there has been scarcely any research in this country with self-report screening instruments for malingered symptoms.

This is particularly important, since although the general tests have yielded satisfactory results, they do have certain limitations, among them the time necessary for their administration and the specialized training required for their application and interpretation, which seriously limits the scope of their use (Schinka & Borum, 1993, 1994).

The SIMS (Widows & Smith, 2005) stands out among these screening instruments for malingering given the range of areas in which it can be used for detecting malingered symptoms (Wisdom et al., 2010); moreover, there are considerable accumulated research findings in the scientific literature on its utility and discriminative capacity (e.g., Edens et al., 2007; Laffoon, 2009; Vitacco et al., 2006).

The second objective of this thesis, and its principal purpose, was to determine the utility of the SIMS in Spanish population. To this end, three complementary empirical studies were designed to explore its discriminative capacity in the area of psychopathology and chronic pain, combining the use of analogue, self-sufficient measures and known-groups designs.

The first study included three groups (control $n=174$, analogue $n=159$ and clinical $n=145$) and three malingering conditions for the analogue group: feigning of psychotic disorder to avoid imprisonment or reduce the sentence, feigning of adverse consequences of a road accident in order to obtain financial compensation, and feigning of depression in order to be granted sick leave. Its objective was to determine the capacity of the SIMS for detecting participants who had been instructed to simulate in each of these three specific psychopathological conditions and overall.

The second study set out to test the efficacy of the SIMS in real medico-legal contexts with individuals on sick leave due to psychopathological problems ($n= 207$), categorized in terms of presence or absence of malingering according to an external criterion, the Fp scale of the MMPI-2, which had shown satisfactory properties for this purpose in previous studies (Rogers et al., 2009). Its convergent validity with the MMPI-2-RF (Ben-Porath and Tellegen, 2008) was also analyzed.

The aim of the third study was to analyze the capacity and utility of the SIMS scales for identifying non-credible somatic complaints related to chronic lower-back pain, using for this purpose – and for the first time in this field – a design with known diagnostic groups. To this end, the study compared SIMS results for, on the one hand, individuals without pathology (controls, $n=30$) and clinical patients suffering from chronic lower-back pain ($n=45$), and on the other, clinical patients with suspected malingering ($n=48$) and analogue malingerers instructed to feign lower-back pain in relation to a claim for financial compensation ($n=20$). In addition, the convergence of the SIMS with the MMPI-2-RF was analyzed.

The data obtained in these three empirical studies have shown the utility of the SIMS in Spanish samples as an instrument for initial screening in the different main areas of malingering. The analyses carried out yielded large effect sizes ($d>1.25$) when it came to overall significant differentiation between genuine clinical patients and malingerers based on SIMS total score. These results concur with those obtained in the scientific literature with different designs and in different symptomatological areas, with effect sizes ranging from 0.62 to 2.42 (Alwes, 2006; Buijs, 2011; Cima, Hollnack, et al., 2003; Clegg et al., 2009; Dearth, 2007; Efendov, 2007; González-Ordi et al., 2010; Lewis et al., 2002; Merckelbach & Smith, 2003; Merten et al., 2007; Vitacco et al., 2006; Wisdom et al., 2010).

The results obtained in this thesis as regards studies on the detection of malingered psychopathology also suggest the possibility of considering SIMS total scores lower than 16 points as indicating the absence of psychopathological malingering, and scores above 16 as indicating suspicion of malingered psychopathology at different levels: >16 possible suspicion of exaggeration of symptoms; >23 probable suspicion of exaggeration of symptoms; >40 clear suspicion of exaggeration of symptoms, using the levels of certainty gradients suggested by Berry and Nelson (2010).

The results obtained in the detection of exaggerated symptoms of chronic lower-back pain would endorse the use of cut-off points lower than those currently indicated for the detection of psychopathological symptom exaggeration. This is due to the fact that in the case of chronic lower-

back pain the baseline level of psychopathological, neurological and cognitive symptomatology is considerably lower than in other psychopathological disorders, as would be expected. Thus, the SIMS values obtained here in the sample of clinical patients with chronic lower-back pain not suspected of malingering are similar to those found in samples of the general population in the scientific literature (Cima, Hollnack, et al., 2003; Clegg et al., 2009; Edens et al., 1999; Gillard, 2010; Rogers et al., 2005; Widows & Smith, 2005) or in the various control samples analyzed in this thesis. The results obtained here would suggest a preliminary consideration of total-score values of >9 for suspicion of symptom exaggeration in chronic lower-back pain.

In addition to the studies on effect size and diagnostic utility of the different cut-off points, correlational studies were carried between the SIMS and MMPI-2-RF with regard to both psychopathology and chronic lower-back pain. The results obtained reveal a high level of convergence between the two scales, with high average correlations both in the study of the detection of psychopathological symptom exaggeration ($M r = 0,66$) and in that of lower-back pain ($M r = 0,63$); this is in line with the findings from previous research looking at correlations between the SIMS and the MMPI-2 (Carter et al., 2009; Cima, Hollnack, et al., 2003; Heinze & Purisch, 2001; Leung, 2010; Widows & Smith, 2005).

In relation to the SIMS and MMPI-2-RF scales, convergence was observed between the specific scales for the detection of similar areas of symptom exaggeration. Thus, among the notable correlations was that between the Psychosis scale from the SIMS (P) and the Fp-r from the MMPI-2-RF, which address similar aspects; between the Amnesic disorders scale from the SIMS (AM) and the RBS scale from the MMPI-2-RF, created specifically for detecting exaggeration of memory problems; and between the Neurologic Impairment scale from the SIMS (NI) and the Infrequent somatic complaints (Fs), Memory problems (RBS) and Exaggeration of symptoms (F-r) scales. In the study on lower-back pain in particular, the SIMS scales with the best discriminative capacity, Total score, Neurologic Impairment (NI) and Affective disorders (AF), were those that yielded the highest correlations with the most effective of the MMPI-2-RF scales for detecting pain simulation, that is, F-r, Fs, FBS-r and RBS (Wygant et al., 2011), and this supports, in convergent fashion, its discriminative capacity.

The results obtained in these correlational studies reveal adequate convergence between the SIMS and the exaggeration of symptoms scales of the MMPI-2 and MMPI-2-RF, the instruments of reference in the assessment of exaggeration in the areas of both psychopathology (Gervais et al., 2011; Marion et al., 2011; Rogers, Gillard, et al., 2011; Sellbom & Bagby, 2010; Sellbom et al., 2010;

Wygant et al., 2009, 2010, 2011; Youngjohn et al., 2011) and chronic pain (Anderson, 2011; Greiffenstein et al., 2013; Sellbom et al., 2012; Wygant et al., 2011).

On the whole, the findings from the empirical studies reported in this dissertation provide evidence of the utility of the SIMS as a screening instrument for exaggeration of psychopathological and lower-back pain symptoms in Spanish population. As such, they contribute to the body of research carried out with the SIMS internationally, in the fields of both psychopathology (Alwes, Clark, Berry y Granacher, 2008; Clegg et al., 2009; Dandachi-FitzGerald et al., 2011; Dearth, 2007; Edens et al., 2007; Efendov, 2007; Heinze & Purisch, 2001; Laffoon, 2009; Lewis et al., 2002; Merkelbach & Smith, 2003; Peters, Jelacic, et al., 2013; Smith & Burger, 1997; Vitacco et al., 2006; Wisdom et al., 2010) and medicine (Benge et al., 2012; Merten et al., 2008; van Beilen et al., 2009), and which also yielded satisfactory results.

Although further research will be needed to replicate the findings obtained here with larger and more varied samples, the results contributed by the present work provide evidence supporting the use of the SIMS as a screening instrument for suspicion of malingering in relation to psychopathology and lower-back pain in Spanish population. This tool can indeed contribute to the arsenal of professional resources that will enable better use of the limited resources of our health system and their assignment to those patients who really need them.

Keywords: Malingering, Feigning, SIMS, screening, disability, anxiety, depression, lower-back pain.

Referencias bibliográficas

- Aamodt, M. G., y Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner*, 15(1), 6–11.
- Achenbach, T. M., Krukowski, R. A., Dumenci, L., y Ivanova, M. Y. (2005). Assessment of Adult Psychopathology: Meta-Analyses and Implications of Cross-Informant Correlations. *Psychological Bulletin*, 131(3), 361–382. doi:10.1037/0033-2909.131.3.361
- Aguerrevere, L. E., Greve, K. W., Bianchini, K. J., y Meyers, J. E. (2008). Detecting malingering in traumatic brain injury and chronic pain with an abbreviated version of the Meyers Index for the MMPI-2. *Archives of Clinical Neuropsychology*, 23(7-8), 831–838. doi:10.1016/j.acn.2008.06.008
- Aikman, G. G., y Souheaver, G. T. (2008). Use of the Personality Assessment Inventory (PAI) in neuropsychological testing of psychiatric outpatients. *Applied Neuropsychology*, 15(3), 176–183. doi:10.1080/09084280802324283
- Allen, L. M., Conder, R. L., Green, P., y Cox, D. R. (1997). *CARB'97 manual for the Computerized Assessment of Response Bias*. Durham: NC: CogniSyst, Inc.
- Alonso, J., Petukhova, M., Vilagut, G., Chatterji, S., Heeringa, S., Üstün, T. B., ... Kessler, R. C. (2011). Days out of role due to common physical and mental conditions: results from the WHO World Mental Health surveys. *Molecular Psychiatry*, 16(12), 1234–1246. doi:10.1038/mp.2010.101
- Alwes, Y. R. (2006). *The utility of the Structured Inventory of Malingered Symptomatology as a screen for the feigning of neurocognitive deficit and psychopathology in a civil forensic sample* (Tesis de maestría). Recuperado de http://uknowledge.uky.edu/cgi/viewcontent.cgi?article=1397&context=gradschool_theses

- Alwes, Y. R., Clark, J., Berry, D. T. R., y Granacher, R. (2008). Screening for feigning in a civil forensic setting. *Journal of Clinical and Experimental Neuropsychology*, 30(2), 133–140. doi:10.1080/13803390701260363
- Amengual, G. M., Puig Bausili, L., Gómez Montoro, J., y Jiménez Jiménez, A. J. (2006). Aspectos médico-legales de las incapacidades laborales. *Medicina Clínica*, 126(17), 671–675.
- American Psychiatric Association. (2007). *Diagnostic and statistical manual of mental disorders : DSM-IV-TR* (4. ed., revision del texto.). Arlington VA: Autor.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5*. Arlington VA: Autor.
- Anderson, J. L. (2011). *A Multi-Method Assessment Approach to the Detection of Malingered Pain: Association with the MMPI-2 Restructured Form* (Tesis de maestría). Recuperado de <http://encompass.eku.edu/cgi/viewcontent.cgi?article=1018&context=etd>
- Arbisi, P. A., y Ben-Porath, Y. S. (1995). An MMPI-2 Infrequent Response Scale for Use With Psychopathological Populations: The Infrequency-Psychopathology Scale, F(p). *Psychological Assessment*, 7(4), 424–431.
- Arbisi, P. A., y Ben-Porath, Y. S. (1998). The ability of Minnesota Multiphasic Personality Inventory–2 validity scales to detect fake-bad responses in psychiatric inpatients. *Psychological Assessment*, 10(3), 221–228.
- Arbisi, P. A., Polusny, M. A., Erbes, C. R., Thuras, P., y Reddy, M. K. (2011). The Minnesota Multiphasic Personality Inventory-2 Restructured Form in National Guard Soldiers Screening Positive for Posttraumatic Stress Disorder and Mild Traumatic Brain Injury. *Psychological Assessment*, 23(1), 203–214. doi:10.1037/a0021339

- Arce, R., del Carmen Pampillón, M., y Fariña, F. (2002). Desarrollo y evaluación de un procedimiento empírico para detección de la simulación de enajenación mental en el contexto legal. *Anuario de Psicología*, *33*(3), 385–408.
- Arce, R., Fariña, F., Carballal, A., y Novo, M. (2009). Creación y validación de un protocolo de evaluación forense de las secuelas psicológicas de la violencia de género. *Psicothema*, *21*(2), 241–247.
- Archer, E. M., Hagan, L. D., Mason, J., Handel, R. W., y Archer, R. P. (2012). MMPI-2-RF Characteristics of Custody Evaluation Litigants. *Assessment*, *19*(1), 14–20. doi:10.1177/1073191110397469
- Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., y Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment*, *87*(1), 84–94.
- Ardolf, B. R., Denney, R. L., y Houston, C. M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist*, *21*(6), 899–916.
- Armstrong, D. (2008, 5 de marzo). Malingerer test roils personal-injury law: “Fake Bad Scale” bars real victims, its critics contend. *The Wall Street Journal*, p. A1.
- Asche, C. V., Kirkness, C. S., McAdam-Marx, C., y Fritz, J. M. (2007). The societal costs of low back pain: data published between 2001 and 2007. *Journal of Pain & Palliative Care Pharmacotherapy*, *21*(4), 25–33.
- Aslan, C. (2010). *A comparison of two psychological tests: The MMPI-2-Scale 4 and Restructured Clinical (RC) Scale 4 using an incarcerated sample* (Tesis doctoral). Recuperado de ProQuest Dissertations and Theses. (UMI No. 3411484)

- Aurrekoetxea Agirre, J. J., Sanzo Ollakarizketa, J. M., Zubero Oleagoitia, E., y Alamillo Gascón, E. (2009). Repetición de procesos de incapacidad temporal según diagnóstico. *Atención Primaria*, 41(8), 439–445. doi:10.1016/j.aprim.2008.11.002
- Ayearst, L. E. (2011). *The creation and construct validity of MMPI-2-RF based personality disorder scales* (Tesis doctoral). Recuperado de ProQuest Dissertations & Theses. (UMI No NR64880)
- Baer, R. A., Rinaldo, J. C., y Berry, D. T. R. (2003). Self-report distortions (including faking, lying, malingering, social desirability). En R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment* (Vol. 2, pp. 861–866). Londres: SAGE.
- Bagby, R. M., Nicholson, R. A., Bacchiochi, J. R., Ryder, A. G., y Bury, A. S. (2002). The Predictive Capacity of the MMPI-2 and PAI - Validity Scales and Indexes to Detect Coached and Uncoached Feigning. *Journal of Personality Assessment*, 78(1), 69–86. doi:10.1207/S15327752JPA7801_05
- Baldessarini, R. J., Finklestein, S., y Arana, G. W. (1983). The predictive power of diagnostic tests and the effect of prevalence of illness. *Archives of General Psychiatry*, 40(5), 569–573.
- Barraquer, L. I. (1992). El síndrome postconmocional: ¿Cuánto de orgánico y cuánto de psicógeno? *Archivos de Neurobiología*, 55, 193–196.
- Bauer, L., y McCaffrey, R. J. (2006). Coverage of the Test of Memory Malingering, Victoria Symptom Validity Test, and Word Memory Test on the Internet: Is test security threatened? *Archives of Clinical Neuropsychology*, 21(1), 121–126. doi:10.1016/j.acn.2005.06.010
- Beaber, R. J., Marston, A., Michelli, J., y Mills, M. J. (1985). A brief test for measuring malingering in schizophrenic individuals. *The American Journal of Psychiatry*, 142(12), 1478–1481.
- Beaulieu, A. C. (2008). *Development of a malingering screening tool for female inmates* (Tesis de maestría). Recuperado de ProQuest Dissertations & Theses. (UMI No. 3322240)

- Belter, R. W., y Piotrowski, C. (2001). Current Status of Doctoral-Level Training in Psychological Testing. *Journal of Clinical Psychology, 57*(6), 717–726. doi:10.1002/jclp.1044
- Ben-Porath, Y. S., y Tellegen, A. (2008). *MMPI-2-RF Manual for Administration, Scoring, and Interpretation*. Minneapolis: University of Minnesota Press.
- Ben-Shakhar, G., y Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology, 88*(1), 131–151.
- Benavides, F. G., Plana, M., Serra, C., Domínguez, R., Despuig, M., Aguirre, S., ... Gimeno, D. (2007). Incapacidad temporal por contingencia común: papel de la edad, el sexo, la actividad económica y la Comunidad Autónoma. *Revista Española de Salud Pública, 81*(2), 183–190.
- Bender, S. D., y Rogers, R. (2004). Detection of neurocognitive feigning: development of a multi-strategy assessment. *Archives of Clinical Neuropsychology, 19*(1), 49–60.
- Benge, J. F., Wisdom, N. M., Collins, R. L., Franks, R., LeMaire, A., y Chen, D. K. (2012). Diagnostic utility of the Structured Inventory of Malingered Symptomatology for identifying psychogenic non-epileptic events. *Epilepsy & Behavior, 24*(4), 439–444. doi:10.1016/j.yebeh.2012.05.007
- Bernard, L. C. (1990). Prospects for faking believable memory deficits on neuropsychological tests and the use of incentives in simulation research. *Journal of Clinical and Experimental Neuropsychology, 12*(5), 715–728. doi:10.1080/01688639008401014
- Berry, D. T. R., y Granacher, R. P. (2009). Feigning of psychiatric symptoms in the context of documented severe head injury and preserved motivation on neuropsychological testing. En J. Morgan y J. J. Sweet (Eds.), *Neuropsychology of Malingering Casebook* (pp. 170–179). Nueva York: Taylor & Francis.

- Berry, D. T. R., y Nelson, N. W. (2010). DSM-5 and Malingering: a Modest Proposal. *Psychological Injury and Law*, 3(4), 295–303. doi:10.1007/s12207-010-9087-7
- Berry, D. T. R., Wetter, M. W., Baer, R. A., Widiger, T. A., Sumpter, J. C., Reynolds, S. K., y Hallam, R. A. (1991). Detection of random responding on the MMPI-2: Utility of F, back F, and VRIN scales. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3(3), 418–423. doi:10.1037/1040-3590.3.3.418
- Bianchini, K. J., Greve, K. W., y Glynn, G. (2005). On the diagnosis of malingered pain-related disability: lessons from cognitive malingering research. *The Spine Journal: Official Journal of the North American Spine Society*, 5(4), 404–417. doi:10.1016/j.spinee.2004.11.016
- Binder, L., y Rohling, M. (1996). Money matters: a meta-analytic review of the effects of financial incentives on recovery after closed-head injury. *American Journal of Psychiatry*, 153(1), 7–10.
- Blanchard, D. D., McGrath, R. E., Pogge, D. L., y Khadivi, A. (2003). A Comparison of the PAI - and MMPI-2 As Predictors of Faking Bad in College Students. *Journal of Personality Assessment*, 80(2), 197–205. doi:10.1207/S15327752JPA8002_08
- Blasco Saiz, J. L., y Pallardo Durá, L. (2012). Trastornos adaptativos y de la personalidad. En H. González-Ordi, P. Santamaría, y P. Capilla Ramírez (Eds.), *Estrategias de detección de la simulación. Un manual clínico multidisciplinar*. (pp. 677–702). Madrid: TEA Ediciones.
- Blasco Saiz, J. L., y Pallardó Durá, L. (2013). Detección de exageración de síntomas mediante el SIMS y el MMPI-2-RF en pacientes diagnosticados de trastorno mixto ansioso-depresivo y adaptativo en el contexto medicolegal: un estudio preliminar. *Clínica Y Salud*, 24(3), 177–183. doi:10.1016/S1130-5274(13)70019-7
- Bloom, D. E., Cafiero, E., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., ... Weiss, J. (2012). *The Global Economic Burden of Noncommunicable Diseases* (PGDA Working Paper

- No. 8712). Program on the Global Demography of Aging. Recuperado de <http://ideas.repec.org/p/gdm/wpaper/8712.html>
- Boccaccini, M. T., Murrie, D. C., y Duncan, S. A. (2006). Screening for malingering in a criminal-forensic sample with the Personality Assessment Inventory. *Psychological Assessment, 18*(4), 415–423. doi:10.1037/1040-3590.18.4.415
- Boone, K. B. (2007a). 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). Nueva York: Guilford Press.
- Boone, K. B. (Ed.). (2007b). *Assessment of feigned cognitive impairment*. Nueva York: Guilford Press.
- Boone, K. B. (2011). Clarification or Confusion? A Review of Rogers, Bender, and Johnson's A Critical Analysis of the MND Criteria for Feigned Cognitive Impairment: Implications for Forensic Practice and Research. *Psychological Injury and Law, 4*(2), 157–162. doi:10.1007/s12207-011-9106-3
- Brage, S., Nygard, J. F., y Tellnes, G. (1998). The gender gap in musculoskeletal-related long term sickness absence in Norway. *Scandinavian Journal of Public Health, 26*(1), 34–43. doi:10.1177/14034948980260010901
- Brage, S., Sandanger, I., y Nygard, J. F. (2007). Emotional Distress as a Predictor for Low Back Disability. *Spine, 32*(2), 269–274. doi:10.1097/01.brs.0000251883.20205.26
- Breakwell, G. M., Smith, J. A., y Wright, D. B. (2012). *Research Methods in Psychology* (4ª Ed.). Londres: SAGE.
- Brooks, L., Johnson-Greene, D., Lattie, E., y FERENCE, T. (2012). The Relationship Between Performances on Neuropsychological Symptom Validity Testing and the MCMI-III in Patients

- With Fibromyalgia. *The Clinical Neuropsychologist*, 26(5), 816–831.
doi:10.1080/13854046.2012.662999
- Buijs, G. (2011). *Het Gebruik van Symptoomvaliditeitstests bij het Neuropsychologisch Onderzoek bij ADHD*. Universiteit van Tilburg, Tilburg, Holanda. Recuperado de <http://arno.uvt.nl/show.cgi?fid=122770>
- Burchett, D. L., y Ben-Porath, Y. S. (2010). The Impact of Overreporting on MMPI-2-RF Substantive Scale Score Validity. *Assessment*, 17(4), 497–516. doi:10.1177/1073191110378972
- Burton, V., Vilar-López, R., y Puente, A. E. (2012). Measuring Effort in Neuropsychological Evaluations of Forensic Cases of Spanish Speakers. *Archives of Clinical Neuropsychology*, 27(3), 262–267.
doi:10.1093/arclin/acs026
- Bush, S. S., Ruff, R. M., Troster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., ... Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity: NAN Policy & Planning Committee. *Archives of Clinical Neuropsychology*, 20(4), 419–426.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A. y Kaemmer, B. (1989). *Minnesota multiphasic personality inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.
- Capilla Ramírez, P. y González-Ordi, H. (2009). Protocolo para la detección de la simulación del dolor en la práctica clínica: estudio de casos. *Trauma*, 20(4), 255–263.
- Carmon, L., Gabriel, R., Ballina, J. y Laffon, A. (2001). Proyecto EPISER 2000: Prevalencia de enfermedades reumáticas en la población española. Metodología, resultados del reclutamiento y características de la población. *Revista Española de Reumatología*, 28(1), 18–25.

- Carter, A. ., Tharshini Chandra, B. S. y Holden, R. (2009, octubre). The use of forensic tests for the assessment of psychological symptom magnification in worker's compensation-seeking head injury claimants. Comunicación presentada en *The Trauma, Psychological Injury, Law: Controversies and Advances*, Toronto, Ontario.
- Casado Morales, M. I., Moix, J. y Vidal Fernández, J. (2008). Etiología, cronificación y tratamiento del dolor lumbar. *Clínica y Salud*, 19(3), 379–392.
- Casado Morales, M. I. y Urbano, M. P. (2001). Valoración clínica de los factores psicológicos que intervienen en el dolor lumbar crónico. *Revista de La Sociedad Española de Dolor*, 8, 48–69.
- Cassidy, J. D., Carroll, L. J., Cote, P., Lemstra, M., Berglund, A. y Nygren, A. (2000). Effect of Eliminating Compensation for Pain and Suffering on the Outcome of Insurance Claims for Whiplash Injury. *N Engl J Med*, 342(16), 1179–1186. doi:10.1056/NEJM200004203421606
- Castillo, M. S., Arias, M., Anglada, R., Lleget, I., Ruiz, A., Escribá, J. y Gil, M. (2003). Estudio epidemiológico de la lumbalgia. Análisis de factores predictivos de incapacidad. *Rehabilitación*, 37, 3–10.
- Catalina-Romero, C., Martínez-Muñoz, P., Quevedo-Aguado, L., Ruiz-Moraga, M., Fernández-Labandera, C. y Calvo-Bonacho, E. (2013). Predictores de la duración de la incapacidad temporal por contingencias comunes en los trastornos de ansiedad. *Gaceta Sanitaria*, 27(1), 40–46. doi:10.1016/j.gaceta.2011.12.008
- Cercy, S., Schretlen, D. y Brandt, J. (1997). Simulated amnesia and the pseudo-memory phenomena. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2ª ed., pp. 85–107). Nueva York: Guilford Press.

- Chafetz, M. D. (2008). Malingering on the Social Security Disability Consultative Exam: Predictors and Base Rates. *The Clinical Neuropsychologist*, 22(3), 529–546.
doi:10.1080/13854040701346104
- Chalmers, I., Hedges, L. V. y Cooper, H. (2002). A brief history of research synthesis. *Evaluation & the Health Professions*, 25(1), 12–37.
- Cima, M., Hollnack, S., Kremer, K., Knauer, E., Schellbach-Matties, R., Klein, B. y Merckelbach, H. (2003). Strukturierter Fragebogen Simulierter Symptome. *Der Nervenarzt*, 74(11), 977–986.
doi:10.1007/s00115-002-1438-5
- Cima, M., Merckelbach, H., Hollnack, S. y Knauer, E. (2003). Characteristics of psychiatric prison inmates who claim amnesia. *Personality and Individual Differences*, 35(2), 373–380.
doi:10.1016/S0191-8869(02)00199-X
- Cima, M. y Van Oorsouw, K. (2013). The relationship between psychopathy and crime-related amnesia. *International Journal of Law and Psychiatry*, 36(1), 23–29.
doi:10.1016/j.ijlp.2012.11.004
- Clark, T. C., Oslund, S. R. y Hopwood, C. J. (2010). PAI assessment in medical settings. En M. Blais, M. Baity y C. Hopwood (Eds.), *Clinical application of the Personality Assessment Inventory*. Nueva York: Routledge.
- Clegg, C., Fremouw, W. y Mogge, N. L. (2009). Utility of the Structured Inventory of Malingered Symptomatology (SIMS) and the Assessment of Depression Inventory (ADI) in screening for malingering among outpatients seeking to claim disability. *Journal of Forensic Psychiatry & Psychology*, 20(2), 239–254. doi:10.1080/14789940802267760
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.

- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Collins, P. Y., Patel, V., Joestl, S. S., March, D., Insel, T. R., Daar, A. S., ... Fairburn, C. (2011). Grand challenges in global mental health. *Nature*, 475(7354), 27–30.
- Conners, C. K. (2000). *Conners' Continuous Performance Test II (CPT II V. 5)*. North Tonawanda, NY: Multi-Health Systems Inc.
- Consejo de la Unión Europea. (2011). *Conclusions on "Innovative approaches for chronic disease in public health and healthcare systems."* Bruselas: Consejo de la Unión Europea. Recuperado de http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/lisa/118282.pdf
- Crawford, E. F., Greene, R. L., Dupart, T. M., Bongar, B. y Childs, H. (2006). MMPI–2 Assessment of Malingered Emotional Distress Related to a Workplace Injury: A Mixed Group Validation. *Journal of Personality Assessment*, 86(2), 217–231. doi:10.1207/s15327752jpa8602_11
- Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement*, 6, 475–494.
- Dagenais, S., Caro, J. y Haldeman, S. (2008). A systematic review of low back pain cost of illness studies in the United States and internationally. *The Spine Journal*, 8(1), 8–20. doi:10.1016/j.spinee.2007.10.005
- Dandachi-FitzGerald, B. y Merckelbach, H. (2013). Feigning ≠ Feigning a Memory Deficit: The Medical Symptom Validity Test as an Example. *Journal of Experimental Psychopathology*, 4, 46–53. doi:10.5127/jep.025511
- Dandachi-FitzGerald, B., Ponds, R. W. H. M. y Merten, T. (2013). Symptom Validity and Neuropsychological Assessment: A Survey of Practices and Beliefs of Neuropsychologists in Six European Countries. *Archives of Clinical Neuropsychology*, 28(8), 771–783. doi:10.1093/arclin/act073

- Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J. V. y Merckelbach, H. (2011). Cognitive Underperformance and Symptom Over-Reporting in a Mixed Psychiatric Sample. *The Clinical Neuropsychologist*, 25(5), 812–828. doi:10.1080/13854046.2011.583280
- Dearth, C. M. S. (2007). Cross-validation of malingering tests for adolescent forensic evaluations. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 68(4-B), 2643.
- Delclós, J., Gimeno, D., Torá, I., Martínez, J. M., Manzanera, R., Jardí, J., ... Benavides, F. G. (2012). Distribución de la duración de la incapacidad temporal por contingencia común por diagnóstico médico (Cataluna, 2006-2008). *Gaceta Sanitaria*, 27(1), 81–83.
- Delis, D. C. y Wetter, S. R. (2007). Cogniform Disorder and Cogniform Condition: Proposed diagnoses for excessive cognitive symptoms. *Archives of Clinical Neuropsychology*, 22(5), 589–604. doi:10.1016/j.acn.2007.04.001
- DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J. y Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review*, 1(4), 346–357.
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M. y Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, 70(5), 979–995.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K. y Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74–118.
- Derogatis, L. R. (1996). *SCL-90-R: Symptom Checklist-90-R: Administration, scoring, and procedures manual*. Minneapolis, MN: NCS Pearson.
- Dionysus, K. E., Denney, R. L. y Halfaker, D. A. (2010). Detecting Negative Response Bias with the Fake Bad Scale, Response Bias Scale, and Henry–Heilbronner Index of the Minnesota Multiphasic

- Personality Inventory-2. *Archives of Clinical Neuropsychology*, 26(2), 81–88.
doi:10.1093/arclin/acq096
- Drob, S. L., Meehan, K. B. y Waxman, S. E. (2009). Clinical and Conceptual Problems in the Attribution of Malingering in Forensic Evaluations. *Journal of the American Academy of Psychiatry and the Law Online*, 37(1), 98–106.
- Duellman, R. M. y Bowers, T. G. (2004). Use of the Personality Assessment Inventory (PAI) in Forensic and Correctional Settings: Evidence for Concurrent Validity. *International Journal of Forensic Psychology*, 1(2), 42–57.
- Echeburúa, E., Amor, P. J. y Corral, P. (2003). Autoinformes y entrevistas en el ámbito de la psicología clínica forense: limitaciones y nuevas perspectivas. *Análisis y Modificación de Conducta*, 29(126), 503–522.
- Edens, J. F., Cruise, K. R. y Buffington-Vollum, J. K. (2001). Forensic and correctional applications of the personality assessment inventory. *Behavioral Sciences & the Law*, 19(4), 519–543.
doi:10.1002/bsl.457
- Edens, J. F., Otto, R. K. y Dwyer, T. (1999). Utility of the structured inventory of malingered symptomatology in identifying persons motivated to malingering psychopathology. *The Journal of the American Academy of Psychiatry and the Law*, 27(3), 387–396.
- Edens, J. F., Poythress, N. G. y Watkins-Clay, M. M. (2007). Detection of Malingering in Psychiatric Unit and General Population Prison Inmates: A Comparison of the PAI, SIMS, and SIRS. *Journal of Personality Assessment*, 88(1), 33–42. doi:10.1207/s15327752jpa8801_05
- Edens, J. F., Ruiz, M. y Staff, P. A. R. (2005). *PAI Interpretive Report for Correctional Settings*. Lutz, FL: Psychological Assessment Resources.

- Efendov, A. A. (2007). *Incremental predictive validity of the TSI, MMPI-2, and SIMS in identifying coached and uncoached feigning of PTSD in workplace accident victims* (Tesis doctoral). Recuperado de <http://www.collectionscanada.gc.ca/obj/thesescanada/vol2/002/NR21761.PDF>
- Ekman, M., Johnell, O. y Lidgren, L. (2005a). The economic cost of low back pain in Sweden in 2001. *Acta Orthopaedica*, 76(2), 275–284. doi:10.1080/00016470510030698
- Ekman, M., Johnell, O. y Lidgren, L. (2005b). The economic cost of low back pain in Sweden in 2001. *Acta Orthopaedica*, 76(2), 275–284.
- Ekman, M., Jönhagen, S., Hunsche, E. y Jönsson, L. (2005). Burden of illness of chronic low back pain in Sweden: a cross-sectional, retrospective study in primary care setting. *Spine*, 30(15), 1777–1785.
- Elhai, J. D., Butcher, J. J., Reeves, A. N., Baugher, S. N., Gray, M. J., Jacobs, G. A., ... Arbisi, P. (2007). Varying Cautionary Instructions, Monetary Incentives, and Comorbid Diagnostic Training in Malingered Psychopathology Research. *Journal of Personality Assessment*, 88(3), 328–337. doi:10.1080/00223890701332136
- Essig, S. M., Mittenberg, W., Petersen, R. S., Strauman, S. y Cooper, J. T. (2001). Practices in forensic neuropsychology: Perspectives of neuropsychologists and trial attorneys. *Archives of Clinical Neuropsychology*, 16(3), 271–291. doi:10.1016/S0887-6177(99)00065-7
- European Agency for Safety and Health at Work. (2010). *OSH in figures: work-related musculoskeletal disorders in the EU - Facts and figures*. Luxembourg: Office for Official Publ. of the Europ. Communities. Recuperado de <https://osha.europa.eu/en/publications/reports/TERO09009ENC/>

- Farkas, M. R., Rosenfeld, B., Robbins, R. y van Gorp, W. (2006). Do tests of malingering concur? Concordance among malingering measures. *Behavioral Sciences & the Law*, 24(5), 659–671. doi:10.1002/bsl.730
- Faust, D. y Ackley, M. A. (1998). Did you think it was going to be easy? Some methodological suggestions for the investigation and development of malingering detection techniques. En C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation. Critical issues in neuropsychology* (pp. 1–54). Nueva York: Plenum Press.
- Fayad, F., Lefevre-Colau, M. M., Poiraudau, S., Fermanian, J., Rannou, F., Wlodyka Demaille, S., ... Revel, M. (2004). [Chronicity, recurrence, and return to work in low back pain: common prognostic factors]. *Annales de réadaptation et de médecine physique: revue scientifique de la Société française de rééducation fonctionnelle de réadaptation et de médecine physique*, 47(4), 179–189. doi:10.1016/j.annrmp.2004.01.005
- Fishbain, D. A., Cutler, R., Rosomoff, H. L. y Rosomoff, R. S. (1999). Chronic pain disability exaggeration/malingering and submaximal effort research. *The Clinical Journal of Pain*, 15(4), 244–274.
- Forbey, J. D., Lee, T. T. C. y Handel, R. W. (2010). Correlates of the MMPI-2-RF in a College Setting. *Psychological Assessment*, 22(4), 737–744. doi:10.1037/a0020645
- Franco Nicolás, M. y Vivo Molina, J. M. (2007). *Análisis de curvas ROC. Principios básicos y aplicaciones*. Madrid: La Muralla.
- Fransen, M., Woodward, M., Norton, R., Coggan, C., Dawe, M. y Sheridan, N. (2002). Risk factors associated with the transition from acute to chronic occupational back pain. *Spine*, 27(1), 92–98.

- Frederick, R. I. (2000). Mixed group validation: a method to address the limitations of criterion group validation in research on malingering detection. *Behavioral Sciences & the Law*, 18(6), 693–718. doi:10.1002/bsl.432
- Freeman, T., Powell, M. y Kimbrell, T. (2008). Measuring symptom exaggeration in veterans with chronic posttraumatic stress disorder. *Psychiatry Research*, 158(3), 374–380. doi:10.1016/j.psychres.2007.04.002
- Frymoyer, J. W. (1992). Predicting disability from low back pain. *Clinical Orthopaedics and Related Research*, 279, 101–109.
- Frymoyer, J. W. y Cats-Baril, W. (1987). Predictors of low back pain disability. *Clinical Orthopaedics and Related Research*, (221), 89–98.
- Gaines, M. V. (2009). *An examination of the combined use of the PAI and the M-FAST in detecting malingering among inmates* (Tesis doctoral). Recuperado de http://repositories.tdl.org/ttu-ir/bitstream/handle/2346/10347/Gaines_Michelle_diss.pdf?sequence=1
- García Silgo, M. y Robles Sánchez, J. I. (2010). Accesit Premio Fidel Pagés Miravé 2010.-Estudio de las Bajas Temporales para el Servicio por motivos psicológicos: Mejoras en la detección de psicopatología y simulación de trastorno mental en las Fuerzas Armadas. *Sanidad Militar: Revista de Sanidad de Las Fuerzas Armadas de España*, 66(3), 154–162.
- Geraerts, E., Jelicic, M. y Merckelbach, H. (2006). Symptom Overreporting and Recovered Memories of Childhood Sexual Abuse. *Law and Human Behavior*, 30(5), 621–630. doi:10.1007/s10979-006-9043-y
- Gerson, A. R. (2002). Beyond DSM-IV: A meta-review of the literature on malingering. *American Journal of Forensic Psychology*, 20(1), 57–70.

- Gervais, R. O., Ben-Porath, Y. S. y Wygant, D. B. (2009). Empirical Correlates and Interpretation of the MMPI-2-RF - Cognitive Complaints (COG) Scale. *The Clinical Neuropsychologist*, 23(6), 996–1015. doi:10.1080/13854040902748249
- Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B. y Green, P. (2007). Development and Validation of a Response Bias Scale (RBS) for the MMPI-2. *Assessment*, 14(2), 196 –208. doi:10.1177/1073191106295861
- Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B. y Sellbom, M. (2010). Incremental Validity of the MMPI-2-RF Over-reporting Scales and RBS in Assessing the Veracity of Memory Complaints. *Arch Clin Neuropsychol*, 25(4), 274–284. doi:10.1093/arclin/acq018
- Gervais, R. O., Lees-Haley y Ben-Porath, Y. S. (2007, noviembre). *Predicting SVT performance with the MMPI-2-RF FBS-r, RBS, and F S scales*. Póster presentado en the Annual Meeting of the National Academy of Neuropsychology, Scottsdale, AZ.
- Gervais, R. O., Rohling, M. L., Green, P. y Ford, W. (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. *Archives of Clinical Neuropsychology*, 19(4), 475–487. doi:10.1016/j.acn.2003.05.001
- Gervais, R. O., Russell, A. S., Green, P., Allen, L. M., Ferrari, R. y Pieschl, S. D. (2001). Effort Testing in Patients with Fibromyalgia and Disability Incentives. *The Journal of Rheumatology*, 28(8), 1892–1899.
- Gervais, R. O., Wygant, D. B., Sellbom, M. y Ben-Porath, Y. S. (2011). Associations Between Symptom Validity Test Failure and Scores on the MMPI-2-RF Validity and Substantive Scales. *Journal of Personality Assessment*, 93(5), 508–517. doi:10.1080/00223891.2011.594132

- Giger, P., Merten, T., Merckelbach, H. y Oswald, M. (2010). Detection of Feigned Crime-Related Amnesia: A Multi-Method Approach. *Journal of Forensic Psychology Practice*, 10(5), 440. doi:10.1080/15228932.2010.489875
- Gillard, N. D. (2010). *Methodological issues in malingering research | The use of simulation designs* (Tesis de maestría). Recuperado de http://digital.library.unt.edu/ark:/67531/metadc33153/m2/1/high_res_d/thesis.pdf
- Gillis, J. R., Rogers, R. y Bagby, R. M. (1991). Validity of the M Test: Simulation-Design and Natural-Group Approaches. *Journal of Personality Assessment*, 57(1), 130–140. doi:10.1207/s15327752jpa5701_15
- Gironda, R. y Clark, M. (2009). Psychometric properties of the MMPI-2 restructured form. *The Journal of Pain*, 10(4), S11. doi:10.1016/j.jpain.2009.01.048
- Gjesdal, S., Ringdal, P. R., Haug, K. y Mæland, J. G. (2004). Predictors of disability pension in long-term sickness absence Results from a population-based and prospective study in Norway 1994–1999. *The European Journal of Public Health*, 14(4), 398–405. doi:10.1093/eurpub/14.4.398
- Göbber, J., Petermann, F., Piegza, M. y Kobelt, A. (2012). [Symptom validation in patients with migration background in psychosomatic medicine]. *Die Rehabilitation*, 51(5), 356–364. doi:10.1055/s-0032-1323669
- González Viejo, M. A. (2001). Coste de la compensación por incapacidad temporal por dolor lumbar en España. *Rehabilitación: Revista de La Sociedad Española de Rehabilitación Y Medicina Física*, 35(1), 28–34.
- González Viejo, M. A. y Condón Huerta, M. J. (2000). Incapacidad por dolor lumbar en España. *Med Clin (Barc)*, 114, 491–492.

- González-Ordi, H., Capilla Ramírez, P. y Matalobos Veiga, B. (2008). Simulación del dolor en el contexto médico-legal Malingered pain in the medico-legal setting. *Clínica y Salud, 19*(3), 393–415.
- González-Ordi, H. y Gancedo Rojí, M. (1999). Evaluación de la simulación de respuesta: diagnóstico, técnicas y procedimientos. En A. Ruano Hernández, J. M. Muñoz Céspedes y C. Cid Rojo (Eds.), *Psicología de la rehabilitación* (pp. 405–441). Madrid: Fundación MAPFRE-Medicina.
- González-Ordi, H. y Iruarrizaga Díez, I. (2005). Evaluación de las distorsiones de respuestas mediante el MMPI-2. *Papeles Del Psicólogo, 26*, 129–137.
- González-Ordi, H. y Santamaría, P. (2009). *Adaptación española del Inventario Estructurado de Simulación de Síntomas*. Madrid: TEA Ediciones.
- González-Ordi, H., Santamaría, P. y Capilla, P. (2012). La simulación como estilo de respuesta. En H. González Ordi, P. Santamaría y P. Capilla (Eds.), *Estrategias de detección de la simulación. Un manual clínico multidisciplinar*. (pp. 19–64). Madrid: TEA Ediciones.
- González-Ordi, H., Santamaría, P. y Fernández Marín, P. (2010). Precisión predictiva del Inventario Estructurado de Simulación de Síntomas –SIMS– en el contexto medicolegal. *eduPsykhé, 9*(1), 3–22.
- Goodwin, B. E., Sellbom, M. y Arbisi, P. A. (2013). Posttraumatic stress disorder in veterans: The utility of the MMPI–2–RF validity scales in detecting overreported symptoms. *Psychological Assessment, 25*(3), 671–678. doi:10.1037/a0032214
- Graue, L. O., Berry, D. T. R., Clark, J. A., Sollman, M. J., Cardi, M., Hopkins, J. y Werline, D. (2007). Identification of Feigned Mental Retardation Using the New Generation of Malingering Detection Instruments: Preliminary Findings. *The Clinical Neuropsychologist, 21*(6), 929–942. doi:10.1080/13854040600932137

- Green, P. (2004). *Green's medical symptom validity test (MSVT) for microsoft windows: User's manual*. Edmonton, Canada: Green's Publishing.
- Green, P. (2005). *Green's Word Memory Test for Microsoft Windows: User's manual*. Edmonton, Canada: Green's Publishing.
- Green, P., Lees-Haley, P. R. y Allen, L. M. (2003). The Word Memory Test and the Validity of Neuropsychological Test Scores. *Journal of Forensic Neuropsychology*, 2(3), 97–124. doi:10.1300/J151v02n03_05
- Greene, R. L. (2008). Malingering and Defensiveness on the MMPI-2. En R. Rogers (Ed.), *Clinical Assessment of Malingering and Deception* (3^a Ed., pp. 159–181). Nueva York: The Guilford Press.
- Greiffenstein, M. F., Fox, D. y Lees-Haley, P. R. (2007). The MMPI-2 Fake Bad Scale in detection of noncredible brain injury claims. En K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 210–235). Nueva York: Guilford Press.
- Greiffenstein, M. F., Gervais, R. O., Baker, W. J., Artiola, L. y Smith, H. (2013). Symptom Validity Testing in Medically Unexplained Pain: A Chronic Regional Pain Syndrome Type 1 Case Series. *The Clinical Neuropsychologist*, 27(1), 138–147. doi:10.1080/13854046.2012.722686
- 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. doi:10.1016/j.acn.2003.08.002
- Greve, K. W., Bianchini, K. J. y Brewer, S. T. (2013). The Assessment of Performance and Self-report Validity in Persons Claiming Pain-related Disability. *The Clinical Neuropsychologist*, 27(1), 108–137. doi:10.1080/13854046.2012.739646

- Greve, K. W., Ord, J. S., Bianchini, K. J. y Curtis, K. L. (2009). Prevalence of Malingering in Patients With Chronic Pain Referred for Psychologic Evaluation in a Medico-Legal Context. *Archives of Physical Medicine and Rehabilitation*, 90(7), 1117–1126. doi:10.1016/j.apmr.2009.01.018
- Grieve, R. y Mahar, D. (2009). The role of fluid and emotional intelligence in malingering. *Australian Journal of Psychology*, 62(2), 103–111. doi:10.1080/00049530903032836
- Haggerty, K. A., Frazier, T. W., Busch, R. M. y Naugle, R. I. (2007). Relationships Among Victoria Symptom Validity Test Indices and Personality Assessment Inventory Validity Scales in a Large Clinical Sample. *The Clinical Neuropsychologist*, 21(6), 917–928. doi:10.1080/13854040600899724
- Hall, H. V. y Pritchard, D. A. (1996). *Detecting malingering and deception: Forensic distortion analysis (FDA)*. Delray Beach, Fl.: St. Lucie Press.
- Hamilton, J. C., Feldman, M. D. y Cunnien, A. J. (2008). Factitious disorder in medical and psychiatric practices. En R. Rogers (Ed.), *Clinical Assessment of Malingering and Deception* (3ª ed., pp. 128–144). Nueva York: Guilford Press.
- Hammond, J. (2006, octubre). *Symptom validity and cognition among claimants for social security disability benefits: Utility of the Structured Inventory of Malingered Symptoms (SIMS) and Effect of Warning*. Poster presentado en el encuentro anual de The National Academy of Neuropsychology, San Antonio, Texas.
- Hankins, G. C., Barnard, G. W. y Robbins, L. (1993). The Validity of the M Test in a Residential Forensic Facility. *Journal of the American Academy of Psychiatry and the Law Online*, 21(1), 111–121.
- Harp, J. P., Jasinski, L. J., Shandera-Ochsner, A. L., Mason, L. H. y Berry, D. T. R. (2011). Detection of Malingered ADHD Using the MMPI-2-RF. *Psychological Injury and Law*, 4(1), 32–43. doi:10.1007/s12207-011-9100-9

- Harrison, C. (2008). *Assessing malingered responding: Concurrent validation of a forced-choice test using ink blot stimuli for the identification of malingered responses* (Tesis doctoral). The University of Mississippi, Oxford, Mississippi.
- Hathaway, S. R. y McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory*. Minneapolis, MN: University of Minnesota Press.
- Hawes, S. W. y Boccaccini, M. T. (2009). Detection of Overreporting of Psychopathology on the Personality Assessment Inventory: A Meta-Analytic Review. *Psychological Assessment, 21*(1), 112–124. doi:10.1037/n0015036
- Hedges, L. V. y Olkin, I. (1985). *Statistical methods for meta-analysis*. Nueva York: Academic Press.
- Heinze, M. C. (2003). Developing sensitivity to distortion: utility of psychological tests in differentiating malingering and psychopathology in criminal defendants. *Journal of Forensic Psychiatry & Psychology, 14*(1), 151–177. doi:10.1080/1478994031000077961
- Heinze, M. C. y Purisch, A. D. (2001). Beneath the Mask -- Use of Psychological Tests to Detect and Subtype Malingering in Criminal Defendants. *Journal of Forensic Psychology Practice, 1*(4), 23–52. doi:10.1300/J158v01n04_02
- Hellman, B. E. (2010). *Characteristics of the MMPI-2 Fp scale as a function of diagnosis in an active duty inpatient population of combat veterans* (Tesis doctoral). Recuperado de ProQuest Dissertations & Theses Database. (UMI No. 3470457)
- Hill, D. (2009). *Detecting Malingering in Correctional Settings: A Comparison of Several Psychological Tests* (Tesis doctoral). Recuperado de <http://commons.pacificu.edu/cgi/viewcontent.cgi?article=1227&context=spp>

- Hoelzle, J. B., Nelson, N. W. y Arbisi, P. A. (2012). MMPI-2 and MMPI-2-Restructured Form Validity Scales: Complementary Approaches to Evaluate Response Validity. *Psychological Injury and Law*, 5(3-4), 174–191. doi:10.1007/s12207-012-9139-2
- Hopwood, C. J., Morey, L. C., Rogers, R. y Sewell, K. (2007). Malingering on the Personality Assessment Inventory: identification of specific feigned disorders. *Journal of Personality Assessment*, 88(1), 43–48.
- Hopwood, C. J., Orlando, M. J. y Clark, T. S. (2010). The detection of malingered pain-related disability with the Personality Assessment Inventory. *Rehabilitation Psychology*, 55(3), 307–310. doi:10.1037/a0020516
- Hosmer, D. W. y Lemeshow, S. (2000). *Applied logistic regression*. Nueva York: Wiley-Interscience.
- Humbría Mendiola, A., Carmona, L., Peña Sagredo, J. L. y Ortiz, A. M. (2002). Impacto poblacional del dolor lumbar en España: resultados del estudio EPISER. *Revista Española de Reumatología*, 29(10), 471–478.
- Hunsley, J. y Meyer, G. J. (2003). The Incremental Validity of Psychological Testing and Assessment: Conceptual, Methodological, and Statistical Issues. *Psychological Assessment*, 15(4), 446–455. doi:10.1037/1040-3590.15.4.446
- Inda Caro, M., Lemos Giráldez, S., López Rodrigo, A. M. y Alonso Rionda, J. L. (2005). La simulación de enfermedad física o mental. *Papeles Del Psicólogo*, 26, 99–108.
- Ingram, P. B., Kelso, K. M. y McCord, D. M. (2011). Empirical Correlates and Expanded Interpretation of the MMPI-2-RF Restructured Clinical Scale 3 (Cynicism). *Assessment*, 18(1), 95–101. doi:10.1177/1073191110388147
- Inman, T. H., Vickery, C. D., Berry, D. T. ., Lamb, D. G., Edwards, C. L. y Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort

- given during neuropsychological testing: The letter memory test. *Psychological Assessment*, 10(2), 128–139.
- Insurance Research Council. (1993). *Public attitude monitor (survey)*. Oak Brook: Insurance Research Council.
- Iverson, G. L. (1995). Qualitative aspects of malingered memory deficits. *Brain Injury*, 9(1), 35–40.
doi:10.3109/02699059509004569
- Iverson, G. L. (2007). Identifying Exaggeration and Malingering. *Pain Practice*, 7(2), 94–102.
doi:10.1111/j.1533-2500.2007.00116.x
- Iverson, G. L. (2010). Detecting Exaggeration, Poor Effort, and Malingering in Neuropsychology. En A. M. Horton y L. C. Hartlage (Eds.), *The Handbook of Forensic Neuropsychology* (2ª ed., pp. 91–136). Nueva York: Springer.
- Jelicic, M., Ceunen, E., Peters, M. J. V. y Merckelbach, H. (2011). Detecting coached feigning using the test of Memory Malingering (TOMM) and the structured inventory of Malingered Symptomatology (SIMS). *Journal of Clinical Psychology*, 67(9), 850–855.
doi:10.1002/jclp.20805
- Jelicic, M., Hessels, A. y Merckelbach, H. (2006). Detection of Feigned Psychosis with the Structured Inventory of Malingered Symptomatology (SIMS): A Study of Coached and Uncoached Simulators. *Journal of Psychopathology and Behavioral Assessment*, 28(1), 19–22.
doi:10.1007/s10862-006-4535-0
- Jelicic, M., Merckelbach, H., Candel, I. y Geraerts, E. (2007). Detection of feigned cognitive dysfunction using special malinger tests: a simulation study in naïve and coached malingerers. *International Journal of Neuroscience*, 117(8), 1185–1192.
doi:10.1080/00207450600934697

- Jelicic, M., Merckelbach, H. y van Bergen, S. (2004). Symptom validity testing of feigned amnesia for a mock crime. *Archives of Clinical Neuropsychology*, 19(4), 525–531. doi:10.1016/j.acn.2003.07.004
- Jelicic, M., Peters, M. J. V., Leckie, V. y Merckelbach, H. (2007). Basic knowledge of psychopathology does not undermine the efficacy of the Structured Inventory of Malingered Symptomatology (SIMS) to detect feigned psychosis. *Netherlands Journal of Psychology*, 63(3), 98–101. doi:10.1007/BF03061071
- Jelicic, M., van Gaal, M. y Peters, M. J. V. (2013). Expert Knowledge Doesn't Help: Detecting Feigned Psychosis in People with Psychiatric Expertise Using the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of Experimental Psychopathology*, 4(1), 38–45. doi:10.5127/jep.022411
- Jiménez, F. y Sánchez, G. (2004, junio). La falsificación de las técnicas psicométricas: un estudio con el MMPI-2. Comunicación presentada en el I Congreso de Psicología Jurídica y Forense en red., Madrid. Recuperado de <http://gredos.usal.es/jspui/handle/10366/22552>
- Jiménez Gómez, F. y Sánchez Crespo, G. (2002). Sensibilidad al fingimiento de la Escala Psiquiátrica Fp de Arbisi y Ben-Porath (1995, 1998) en la adaptación española del MMPI-2. *Revista Iberoamericana de Diagnóstico Y Evaluación Psicológica*, 14(2), 117–134.
- Jorge, R. M. y Jorge, B. M. (2011). Low back pain simulation: diagnosis and prevalence. *Acta Ortopédica Brasileira*, 19(4), 181–183. doi:10.1590/S1413-78522011000400001
- Jung, B. y Reidenberg, M. M. (2007). Physicians being deceived. *Pain Medicine*, 8(5), 433–437. doi:10.1111/j.1526-4637.2007.00315.x
- Juniper, M., Le, T. K. y Mladsí, D. (2009). The epidemiology, economic burden, and pharmacological treatment of chronic low back pain in France, Germany, Italy, Spain and the UK: a literature-

- based review. *Expert Opinion on Pharmacotherapy*, 10(16), 2581–2592.
doi:10.1517/14656560903304063
- Kahneman, D. y Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39(4), 341–350. doi:10.1037/0003-066X.39.4.341
- Kane, A. W. (2010). Essentials of malingering assessment. En M. J. Ackerman (Ed.) *Essentials of forensic psychological assessment* (2ª ed., pp. 85–111). New Jersey: John Wiley & Sons.
- Karlsson, N. E., Carstensen, J. M., Gjesdal, S. y Alexanderson, K. A. E. (2008). Risk factors for disability pension in a population-based cohort of men and women on long-term sick leave in Sweden. *The European Journal of Public Health*, 18(3), 224–231. doi:10.1093/eurpub/ckm128
- Kashy, D. A. y DePaulo, B. M. (1996). Who lies? *Journal of Personality and Social Psychology*, 70(5), 1037–1051.
- Kay, N. R. M. y Morris-Jones, H. (1998). Pain clinic management of medico-legal litigants. *Injury*, 29(4), 305–308. doi:10.1016/S0020-1383(98)80211-8
- Khouri, R. (2011). *MMPI-2 RF vs. MMPI-2: Latinos with depression* (Tesis doctoral). Recuperado de ProQuest Dissertations & Theses. (UMI No. 3416813)
- Kleiner, M. (2002). *Handbook of polygraph testing*. Londres: Academic Press.
- Klonsky, E. D., Oltmanns, T. F. y Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology Science and Practice*, 9(3), 300–311.
- Kovacs, F. M. y Gil del Real, M. T. (2000). Un estudio necesario, pero un Comité de Expertos superfluo. *Med Clin (Barc)*, 115(16), 637–638.

- Kraemer, H. C., Periyakoil, V. S. y Noda, A. (2002). Tutorial in biostatistics: kappa coefficients in medical research. *Stat Med*, 21(1), 2109–2129.
- Kucharski, L. T., Falkenbach, D. M., Egan, S. S. y Duncan, S. (2006). Antisocial personality disorder and the malingering of psychiatric disorder: A study of criminal defendants. *International Journal of Forensic Mental Health*, 5(2), 195–204.
- Kucharski, L. T., Toomey, J. P., Fila, K. y Duncan, S. (2007). Detection of Malingering of Psychiatric Disorder With the Personality Assessment Inventory: An Investigation of Criminal Defendants. *Journal of Personality Assessment*, 88(1), 25–32. doi:10.1207/s15327752jpa8801_04
- Lachar, D. y Gruber, C. P. (1995). *Personality Inventory for Youth (PIY)*. Los Angeles, CA: Western Psychological Services.
- Laffoon, L. L. (2009). *Detecting Feigning in a Correctional Setting: A Comparison of Multiple Measures* (Tesis doctoral). Recuperado de <http://commons.pacificu.edu/cgi/viewcontent.cgi?article=1219&context=spp>
- Lakke, S. E., Soer, R., Takken, T. y Reneman, M. F. (2009). Risk and prognostic factors for non-specific musculoskeletal pain: A synthesis of evidence from systematic reviews classified into ICF dimensions. *Pain*, 147(1-3), 153–164. doi:10.1016/j.pain.2009.08.032
- Lally, S. J. (2003). What tests are acceptable for use in forensic evaluations? A survey of experts. *Professional Psychology: Research and Practice*, 34(5), 491–498. doi:10.1037/0735-7028.34.5.491
- Lange, R. T., Sullivan, K. A. y Scott, C. (2010). Comparison of MMPI-2 and PAI validity indicators to detect feigned depression and PTSD symptom reporting. *Psychiatry Research*, 176(2-3), 229–235. doi:10.1016/j.psychres.2009.03.004

- Larrabee, G. J. (2003a). Detection of Malingering Using Atypical Performance Patterns on Standard Neuropsychological Tests. *The Clinical Neuropsychologist*, 17(3), 410–425. doi:10.1076/clin.17.3.410.18089
- Larrabee, G. J. (2003b). Exaggerated MMPI-2 symptom report in personal injury litigants with malingered neurocognitive deficit. *Archives of Clinical Neuropsychology*, 18(6), 673–686. doi:10.1016/S0887-6177(02)00157-9
- Larrabee, G. J. (2008). Aggregation Across Multiple Indicators Improves the Detection of Malingering: Relationship to Likelihood Ratios. *The Clinical Neuropsychologist*, 22(4), 666–679. doi:10.1080/13854040701494987
- Larrabee, G. J., Greiffenstein, M. F., Greve, K. W. y Bianchini, K. J. (2007). Refining diagnostic criteria for malingering. En G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334–372). Nueva York: Oxford University Press.
- Leavitt, F. y Sweet, J. J. (1986). Characteristics and frequency of malingering among patients with low back pain. *Pain*, 25(3), 357–364. doi:10.1016/0304-3959(86)90239-3
- Lees-Haley, P. R., English, L. T. y Glenn, W. J. (1991). A Fake Bad Scale on the MMPI-2 for personal injury claimants. *Psychological Reports*, 68(1), 203–210.
- León Mateos, L., Jover Jover, J. A., Loza Santamaría, E. y Pérez Nieto, M. A. (2007). *Recuperación funcional en la lumbalgia incapacitante: eficacia de una intervención cognitivo-conductual temprana*. Madrid: Fundación MAPFRE.
- Leung, N. T. Y. (2010). *Validation of the Malingered Depression Scale and Malingered Mood Disorders Scale with Head Injured Compensation Claimants* (Tesis doctoral). University of Toronto, Toronto, Ontario.

- Lewis, J. L., Simcox, A. M. y Berry, D. T. R. (2002). Screening for feigned psychiatric symptoms in a forensic sample by using the MMPI-2 and the structured inventory of malingered symptomatology. *Psychological Assessment*, 14(2), 170–176.
- Lilienfeld,, S., Thames,, A. y Watts,, A. (2013). Symptom Validity Testing: Unresolved Questions, Future Directions. *Journal of Experimental Psychopathology*, 4(1), 78–87.
doi:10.5127/jep.028312
- Lipman, F. D. (1962). Malingering in personal injury cases. *Temple Law Quarterly*, 35, 141–162.
- Locke, D. E. C., Kirlin, K. A., Thomas, M. L., Osborne, D., Hurst, D. F., Drazkowski, J. F., ... Noe, K. H. (2010). The Minnesota Multiphasic Personality Inventory-2-Restructured Form in the epilepsy monitoring unit. *Epilepsy & Behavior*, 17(2), 252–258.
doi:10.1016/j.yebeh.2009.12.004
- Locke, D. E. C. y Thomas, M. L. (2010). Initial development of Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) scales to identify patients with psychogenic nonepileptic seizures. *Journal of Clinical and Experimental Neuropsychology*, 1–9.
doi:10.1080/13803395.2010.518141
- López Pascual, J., Peydro de Moya, M. F. y Garrido Jaén, J. D. (2009). Análisis del uso de herramientas de valoración funcional de las dolencias lumbares en el ámbito laboral. *Rehabilitación*, 43(1), 16–23.
- MacLaren, V. V. (2001). A quantitative review of the Guilty Knowledge Test. *Journal of Applied Psychology*, 86(4), 674–683.
- Maetzel, A. y Li, L. (2002). The economic burden of low back pain: a review of studies published between 1996 and 2001. *Best Practice & Research. Clinical Rheumatology*, 16(1), 23–30.
doi:10.1053/berh.2001.0204

- Marion, B. E., Sellbom, M. y Bagby, R. M. (2011). The Detection of Feigned Psychiatric Disorders Using the MMPI-2-RF Overreporting Validity Scales: An Analog Investigation. *Psychological Injury and Law*, 4, 1–12. doi:10.1007/s12207-011-9097-0
- Martin, R. C., Bolter, J. F., Todd, M. E., Gouvier, W. D. y Niccolls, R. (1993). Effects of sophistication and motivation on the detection of malingered memory performance using a computerized forced-choice task. *Journal of Clinical and Experimental Neuropsychology*, 15(6), 867–880. doi:10.1080/01688639308402604
- Masip, J., Garrido, E. y Herrero, C. (2004). La detección de la mentira mediante la medida de la tensión en la voz: una revisión crítica The detection of deception using voice stress analysers: A critical review. *Estudios de Psicología*, 25(1), 13–30.
- Masip, J., Sporer, S. L., Garrido, E. y Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law*, 11(1), 99–122.
- Matalobos Veiga, B. y González-Ordi, H. (2007). Detección de simulación en estrés postraumático y depresión. *EduPsykhé: Revista de Psicología Y Psicopedagogía*, 6(1), 21–48.
- McDermott, B. E., Dualan, I. V. y Scott, C. L. (2013). Malingering in the correctional system: Does incentive affect prevalence? *International Journal of Law and Psychiatry*, 36(3–4), 287–292. doi:10.1016/j.ijlp.2013.04.013
- McDermott, B. E. y Feldman, M. D. (2007). Malingering in the medical setting. *Psychiatric Clinics of North America*, 30(4), 645–662.
- McGrath, R. E., Mitchell, M., Kim, B. H. y Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, 136(3), 450–470. doi:10.1037/a0019216

- Meehl, P. E. y Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, 52(3), 194–216.
- Mendiola, A. H. (2004). Consulta monográfica de columna lumbar. Protocolo de investigación clínica ¿cómo son los pacientes con lumbalgia crónica inespecífica? *Revista Española de Reumatología*, 31(8), 453–461.
- Merckelbach, H., Jelicic, M. y Pieters, M. J. V. (2010). The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology*, 33(1), 131–139. doi:10.1080/13803395.2010.495055
- Merckelbach, H., Peters, M. J. V., Jelicic, M., Brands, I. y Smeets, T. (2006). Detecting malingering of Ganser-like symptoms with tests: A case study. *Psychiatry and Clinical Neurosciences*, 60(5), 636–638. doi:10.1111/j.1440-1819.2006.01571.x
- Merckelbach, H., Smeets, T. y Jelicic, M. (2009). Experimental simulation: type of malingering scenario makes a difference. *Journal of Forensic Psychiatry & Psychology*, 20(3), 378–386. doi:10.1080/14789940802456686
- Merckelbach, H. y Smith, G. P. (2003). Diagnostic accuracy of the Structured Inventory of Malingered Symptomatology (SIMS) in detecting instructed malingering. *Archives of Clinical Neuropsychology*, 18(2), 145–152. doi:10.1016/S0887-6177(01)00191-3
- Merten, T., Diederich, C. y Stevens, A. (2008). Vorgetäuschte Beschwerden nach Distorsionstrauma der Halswirbelsäule: eine experimentelle Simulationsstudie [Feigned symptoms after whiplash injury: An experimental analogue study]. *Aktuelle Neurologie*, 35, 8–15.
- Merten, T., Friedel, E. y Stevens, A. (2006). Eingeschränkte Kooperativität in der neurologisch-psychiatrischen Begutachtung. *Versicherungsmedizin*, 58, 19–21.

- Merten, T., Friedel, E. y Stevens, A. (2007). Die Authentizität der Beschwerdenschilderung in der neurologisch-psychiatrischen Begutachtung: eine Untersuchung mit dem Strukturierten Fragebogen Simulierter Symptome.[The authenticity of symptom report in forensic neuropsychiatric assessment: A study with the Structured Inventory of Malingered Symptomatology]. *Praxis Der Rechtspsychologie*, 17(2007), 140–154.
- Merten, T., Lorenz, R. y Schlatow, S. (2010). Posttraumatic Stress Disorder Can Easily Be Faked, but Faking Can Be Detected in Most Cases. *Germ J Psychiatr*, 13, 140–149.
- Merten, T. y Merckelbach, H. (2013). Symptom Validity Testing in Somatoform and Dissociative Disorders: A Critical Review. *Psychological Injury and Law*, 6, 122–137. doi:10.1007/s12207-013-9155-x
- Merten, T., Thies, E., Schneider, K. y Stevens, A. (2009). Symptom Validity Testing in Claimants with Alleged Posttraumatic Stress Disorder: Comparing the Morel Emotional Numbing Test, the Structured Inventory of Malingered Symptomatology, and the Word Memory Test. *Psychological Injury and Law*, 2(3-4), 284–293. doi:10.1007/s12207-009-9057-0
- Meyer, G. J. (2002). Implications of information gathering methods for a refined taxonomy of psychopathology. En L. E. Beutler y M. L. Malik (Eds.), *Rethinking the DSM: A psychological perspective* (pp. 69–105). Washington, D.C.: American Psychological Association.
- Meyers, J. (2002). A validity index for the MMPI-2. *Archives of Clinical Neuropsychology*, 17(2), 157–169. doi:10.1016/S0887-6177(00)00107-4
- Miller, H. A. (2000). The development of the Miller's Forensic Assessment of Symptoms Test: a measure of malingering mental illness. *Disertation Abstracts International: Section B. The Sciences & Engineering*, 60(8-B), 4238.

- Mittenberg, W., Patton, C., Canyock, E. M. y Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, 24(8), 1094–1102.
- Mogge, N. L. y LePage, J. P. (2004). The assessment of depression inventory (ADI): A new instrument used to measure depression and to detect honesty of response. *Depression and Anxiety*, 20(3), 107–113. doi:10.1002/da.20033
- Mogge, N. L., LePage, J. P., Bell, T. y Ragatz, L. (2010). The negative distortion scale: a new PAI validity scale. *Journal of Forensic Psychiatry & Psychology*, 21(1), 77–90. doi:10.1080/14789940903174253
- Mohíno Justes, S., Dolado Cuello, J., Arimany Manso, J., Ortega-Monasterio, L., Cuquerella Fuentes, A., Vilardell Molas, J. y Planchat Teruel, L. M. (2004). Perfiles psicométricos fingidos en relación con estilos de personalidad en reclusos. *Actas Esp Psiquiatr*, 32(5), 264–268.
- Moix, J., Cano, A. y Grupo español de trabajo del programa COST B13 de la comisión europea. (2006). Guía de práctica clínica para la lumbalgia inespecífica basada en la evidencia científica. *Ansiedad Y Estrés*, 12(1), 117–129.
- Moran, C. (2011). *MMPI-2-RF scale comparisons of American and Mexican women entering into substance abuse treatment: An international study* (Tesis doctoral). Recuperado de ProQuest Dissertations & Theses. (UMI No. 3415802)
- Morey, L. C. (1996). *An interpretive guide to the Personality Assessment Inventory (PAI)*. Odessa, FL: Psychological Assessment Resources.
- Morey, L. C. (2007). *Personality Assessment Inventory (PAI). Professional manual* (2ª Ed.). Lutz, FL: Psychological Assessment Resources.

- Morey, L. C. (2012). Detection of Response Bias in Applied Assessment: Comment on McGrath et al. (2010). *Psychological Injury and Law*, 5(3-4), 153–161. doi:10.1007/s12207-012-9131-x
- Morey, L. C. y Quigley, B. D. (2002). The Use of the Personality Assessment Inventory (PAI) in Assessing Offenders. *International Journal of Offender Therapy and Comparative Criminology*, 46(3), 333–349. doi:10.1177/0306624X02463007
- Morey, L. C., Warner, M. B. y Hopwood, C. J. (2007). The Personality Assessment Inventory: Issues in legal and forensic settings. En A. M. Goldstein (Ed.), *Forensic psychology: Emerging topics and forensic roles* (pp. 97–126). Hoboken, NJ: Wiley.
- Moss, A., Jones, C., Fokias, D. y Quinn, D. (2003). The mediating effects of effort upon the relationship between head injury severity and cognitive functioning. *Brain Injury*, 17(5), 377–387. doi:10.1080/0269905031000070125
- Muñiz Fernández, J. y Fernández Hermida, J. R. (2010). La Opinión de los Psicólogos Españoles sobre el Uso de los Tests. *Papeles Del Psicólogo*, 31(1), 108–121.
- Muñoz-Gómez, J. (2003). Epidemiología del dolor lumbar crónico. En C. Muriel Villoria (Ed.), *Abordajes terapéuticos en el dolor lumbar crónico*. (pp. 23–28). Madrid: Ed. Fundación Grünenthal.
- Naldony, E. y O'Dell, R. T. (2010). Measures of central tendency and intercorrelations within and between the SIMS and the M-FAST among medium security prison inmates. *American Journal of Forensic Psychology*, 28(4), 51–68.
- Nelson, N. W., Sweet, J. J., Berry, D. T. R., Bryant, F. B. y Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society*, 13(03), 440–449. doi:10.1017/S1355617707070373

- Nicholson, R. A., Mouton, G. J., Bagby, R. M., Buis, T., Peterson, S. A. y Buigas, R. A. (1997). Utility of MMPI-2 indicators of response distortion: Receiver operating characteristic analysis. *Psychological Assessment, 9*(4), 471–479.
- Nies, K. J. y Sweet, J. J. (1994). Neuropsychological assessment and malingering: A critical review of past and present strategies. *Archives of Clinical Neuropsychology, 9*(6), 501–552. doi:10.1016/0887-6177(94)90041-8
- Norris, M. P. y May, M. C. (1998). Screening for malingering in a correctional setting. *Law and Human Behavior, 22*(3), 315–323. doi:10.1023/A:1025706606774
- O'Connell, B. A. (1960). Amnesia and homicide. *British Journal of Delinquency, 10*, 262–276.
- O'Meara, J. y Shaffer, T. L. (1964). Obscenity in the Supreme Court: A Note on *Jacobellis v. Ohio*. *Notre Dame Lawyer, 40*(1), 1–12.
- Oliva-Moreno, J., López-Bastida, J., Montejo-González, A. L., Osuna-Guerrero, R. y Duque-González, B. (2009). The socioeconomic costs of mental illness in Spain. *The European Journal of Health Economics, 10*(4), 361–369. doi:10.1007/s10198-008-0135-0
- Ortiz-Tallo, M., Santamaría, P., Cardenal, V. y Sánchez, M. P. (2011). *Adaptación española del Inventario de Evaluación de la Personalidad (PAI)*. Madrid: TEA Ediciones.
- Ossipov, V. P. (1944). Malingering: The simulation of psychosis. *Bulletin of the Menninger Clinic, 8*(2), 39–41.
- Otto, R. K. y Heilbrun, K. (2002). The practice of forensic psychology. A look toward the future in light of the past. *The American Psychologist, 57*(1), 5–18.

- Øverland, S., Glozier, N., Henderson, M., Mæland, J. G., Hotopf, M. y Mykletun, A. (2008). Health status before, during and after disability pension award: the Hordaland Health Study (HUSK). *Occupational and Environmental Medicine*, 65(11), 769–773. doi:10.1136/oem.2007.037861
- Pankratz, L. y Erickson, R. C. (1990). Two views of malingering. *The Clinical Neuropsychologist*, 4(4), 379–389.
- Patrick, C. J. y Iacono, W. G. (1989). Psychopathy, threat, and polygraph test accuracy. *Journal of Applied Psychology*, 74(2), 347–355. doi:10.1037/0021-9010.74.2.347
- Peace, K. A. y Masliuk, K. A. (2011). Do Motivations for Malingering Matter? Symptoms of Malingered PTSD as a Function of Motivation and Trauma Type. *Psychological Injury and Law*, 4(1), 44–55. doi:10.1007/s12207-011-9102-7
- Pelfrey, W. V. (2004). The Relationship Between Malingerers' Intelligence and MMPI-2 Knowledge and Their Ability to Avoid Detection. *International Journal of Offender Therapy and Comparative Criminology*, 48(6), 649–663. doi:10.1177/0306624X04265085
- Peters, M. J. V., Jelicic, M., Moritz, S., Hauschildt, M. y Jelinek, L. (2013). Assessing the Boundaries of Symptom Over-Reporting Using the Structured Inventory of Malingered Symptomatology in a Clinical Schizophrenia Sample: Its Relation to Symptomatology and Neurocognitive Dysfunctions. *Journal of Experimental Psychopathology*, 4(1), 64–77. doi:10.5127/jep.023811
- Peters, M. J. V., van Oorsouw, K. I. M., Jelicic, M. y Merckelbach, H. (2013). Let's use those tests! Evaluations of crime-related amnesia claims. *Memory*, 21(5), 599–607. doi:10.1080/09658211.2013.771672
- Pincus, T., Burton, A. K., Vogel, S. y Field, A. P. (2002). A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. *Spine*, 27(5), E109–E120.

- Poizner, J. (2010). A cross-cultural examination of the MMPI-2-RF Substance Abuse scale: Comparing samples from Mexico and the United States. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 71(6-B), 394.
- Poythress, N. G., Edens, J. F. y Watkins, M. M. (2001). The relationship between psychopathic personality features and malingering symptoms of major mental illness. *Law and Human Behavior*, 25(6), 567–582.
- Puente, A. E. y Perez-Garcia, M. (2000a). Neuropsychological assessment of ethnic minorities: Clinical issues. En F. Paniagua y I. Cuellar (Eds.), *Handbook of multicultural mental health* (pp. 419–435). San Diego: Academic Press.
- Puente, A. E. y Perez-Garcia, M. (2000b). Psychological assessment of ethnic minorities. En G. Goldstein y M. Hersen (Eds.), *Handbook of psychological assessment* (3ª Edición., pp. 527–551). Amsterdam: Pergamon Press.
- Quezada–Ortega, M. R., Razo–Mondragón, J.-L.-P.-, Marín-Cotoñieto, I. A., Salinas-Tovar, S. y López-Rojas, P. (2006). Simulación en trabajadores que solicitan pensión por invalidez laboral. *Gaceta Médica de México*, 142(2), 109–112.
- Quijano, F. (1999). Incapacidad temporal/permanente. Situación actual en España. Cambios y consecuencias. En semFYC (Ed.), *Incapacidad temporal: encrucijada ética, clínica y de gestión*. (pp. 5–14). Barcelona: semFYC.
- Resnick, P. J. (1997). Malingering of posttraumatic disorders. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2ª ed., pp. 130–152). Nueva York: Guilford Press.
- Resnick, P. J. (2002). Malingering. En R. Rosner (Ed.), *Principles and practice of forensic psychiatry* (2ª ed., pp. 543–554). Nueva York: Chapman & Hall.

- Resnick, P. J. y Knoll, J. (2005). Faking it: How to detect malingered psychosis. *Current Psychiatry*, 4(11), 12–25.
- Resnick, P. J. y Zuchowski, S. J. (2007). Malingering and psychopathic disorders. En A. Felthous y H. Sass (Eds.), *The International Handbook of Psychopathic Disorders and the Law* (Vol. 2, pp. 461–471). Indianapolis, IN: John Wiley and Sons, Ltd.
- Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology*; *Journal of Personality and Social Psychology*, 51(3), 649–660.
- Robinson, J. P. y Loeser, J. D. (2012). Effects of Workers' Compensation Systems on Recovery from Disabling Injuries. En M. I. Hasenbring, A. C. Rusu y D. C. Turk (Eds.), *From Acute to Chronic Back Pain: Risk Factors, Mechanisms, and Clinical Implications* (pp. 355–376). Nueva York: Oxford University Press.
- Rogers, R. (1984). Towards an empirical model of malingering and deception. *Behavioral Sciences & the Law*, 2(1), 93–111. doi:10.1002/bsl.2370020109
- Rogers, R. (1990a). Development of a new classificatory model of malingering. *The Bulletin of the American Academy of Psychiatry and the Law*, 18(3), 323–333.
- Rogers, R. (1990b). Models of feigned mental illness. *Professional Psychology: Research and Practice*, 21(3), 182–188.
- Rogers, R. (1997). Researching dissimulation. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2ª ed., pp. 398–426). Nueva York: Guilford.
- Rogers, R. (2008a). An introduction to Response Styles. En R. Rogers (Ed.), *Clinical Assessment of Malingering and Deception* (3ª Ed., pp. 3–13). Nueva York: The Guilford Press.
- Rogers, R. (2008b). *Clinical assessment of malingering and deception* (3ª Ed.). New York: Guilford Press.

- Rogers, R. (2008c). Current status of Clinical Methods. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 391–410). Nueva York: The Guilford Press.
- Rogers, R. (2008d). Detection Strategies for Malingering and Defensiveness. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 411–434). Nueva York: The Guilford Press.
- Rogers, R. (2008e). Researching response styles. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 411–434). Nueva York: The Guilford Press.
- Rogers, R., Bagby, R. M. y Chakraborty, D. (1993). Feigning Schizophrenic Disorders on the MMPI--2: Detection of Coached Simulators. *Journal of Personality Assessment*, 60(2), 215. doi:10.1207/s15327752jpa6002_1
- Rogers, R., Bagby, R. M. y Dickens, S. E. (1992). *Structured Interview of Reported Symptoms: SIRS; Professional Manual*. Odessa: Psycholog. Assessment Resources.
- Rogers, R. y Bender, S. D. (2003). Evaluation of malingering and deception. En A. Goldstein y I. B. Weiner (Eds.), *Handbook of Psychology: Vol 11. Forensic Psychology* (pp. 109–132). New Jersey: John Wiley & Sons.
- Rogers, R. y Bender, S. D. (2012). Evaluation of Malingering and related response styles. En I. B. Weiner y R. K. Otto (Eds.), *Handbook of Psychology, Volume 11, Forensic Psychology* (2ª Edición., pp. 517–540). Nueva York: Wiley.
- Rogers, R., Bender, S. D. y Johnson, S. (2011a). A Critical Analysis of the MND Criteria for Feigned Cognitive Impairment: Implications for Forensic Practice and Research. *Psychological Injury and Law*, 4(2), 147–156. doi:10.1007/s12207-011-9107-2

- Rogers, R., Bender, S. y Johnson, S. (2011b). A Commentary on the MND Model and the Boone Critique: "Saying It Doesn't Make It So." *Psychological Injury and Law*, 4(2), 163–167. doi:10.1007/s12207-011-9108-1
- Rogers, R. y Correa, A. A. (2008). Determinations of Malingering: Evolution from Case-Based Methods to Detection Strategies. *Psychiatry, Psychology and Law*, 15(2), 213–223. doi:10.1080/13218710802014501
- Rogers, R. y Cruise, K. R. (1998). Assessment of malingering with simulation designs: threats to external validity. *Law and Human Behavior*, 22(3), 273–285.
- Rogers, R., Gillard, N. D., Berry, D. T. R. y Granacher, R. P. (2011). Effectiveness of the MMPI-2-RF Validity Scales for Feigned Mental Disorders and Cognitive Impairment: A Known-Groups Study. *Journal of Psychopathology and Behavioral Assessment*, 33(3), 355–367. doi:10.1007/s10862-011-9222-0
- Rogers, R., Gillard, N. D., Wooley, C. N. y Ross, C. A. (2012). The Detection of Feigned Disabilities The Effectiveness of the Personality Assessment Inventory in a Traumatized Inpatient Sample. *Assessment*, 19(1), 77–88. doi:10.1177/1073191111422031
- Rogers, R., Hinds, J. D. y Sewell, K. W. (1996). Feigning Psychopathology Among Adolescent Offenders: Validation of the SIRS, MMPI-A, and SIMS - PB - Routledge. *Journal of Personality Assessment*, 67(2), 244–257. doi:10.1207/s15327752jpa6702_2
- Rogers, R., Jackson, R. y Kaminski, P. (2005). Factitious Psychological Disorders: The Overlooked Response Style in Forensic Evaluations. *Journal of Forensic Psychology Practice*, 5(1), 21–41. doi:10.1300/J158v05n01_02
- Rogers, R., Ornduff, S. R. y Sewell, K. W. (1993). Feigning specific disorders: a study of the Personality Assessment Inventory (PAI). *Journal of Personality Assessment*, 60(3), 554–560.

- Rogers, R., Payne, J., Berry, D. y Granacher, R. (2009). Use of the SIRS in Compensation Cases: An Examination of Its Validity and Generalizability. *Law and Human Behavior*, 33(3), 213–224. doi:10.1007/s10979-008-9145-9
- Rogers, R. y Reinhardt, V. R. (1998). Secondary gain: A clinical construct in search of consensus. En G. P. Koocher, J. C. Norcross y S. S. Hill III (Eds.), *Psychologist's desk reference* (pp. 57–72). Oxford: Oxford University Press.
- Rogers, R., Salekin, R. T., Sewell, K. W., Goldstein, A. y Leonard, K. (1998). A comparison of forensic and nonforensic malingerers: A prototypical analysis of explanatory models. *Law and Human Behavior*, 22(4), 353–367. doi:10.1023/A:1025714808591
- Rogers, R., Sewell, K. W. y Gillard, N. D. (2010). *Structured Interview of Reported Symptoms, 2nd Edition (SIRS-2). Professional Manual*. Lutz, FL: Psychological Assessment Resources.
- Rogers, R., Sewell, K. W. y Goldstein, A. M. (1994). Explanatory models of malingering. *Law and Human Behavior*, 18(5), 543–552. doi:10.1007/BF01499173
- Rogers, R., Sewell, K. W., Martin, M. A. y Vitacco, M. J. (2003). Detection of Feigned Mental Disorders: A Meta-Analysis of the MMPI-2 and Malingering. *Assessment*, 10(2), 160–177. doi:10.1177/1073191103010002007
- Rogers, R., Sewell, K. W., Morey, L. C. y Ustad, K. L. (1996). Detection of feigned mental disorders on the personality assessment inventory: a discriminant analysis. *Journal of Personality Assessment*, 67(3), 629–640.
- 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). Nueva York: Kulwer Academic Press.

- Rohling, M. L., Larrabee, G. J., Greiffenstein, M. F., Ben-Porath, Y. S., Lees-Haley, P., Green, P. y Greve, K. W. (2011). A misleading review of response bias: Comment on McGrath, Mitchell, Kim, and Hough (2010). *Psychological Bulletin*, *137*(4), 708–712. doi:10.1037/a0023327
- Rosen, G. (2006). DSM's cautionary guideline to rule out malingering can protect the PTSD data base. *Journal of Anxiety Disorders*, *20*(4), 530–535. doi:10.1016/j.janxdis.2005.03.004
- Ruiz, M. A., Drake, E. B., Glass, A., Marcotte, D. y van Gorp, W. G. (2002). Trying to beat the system: Misuse of the Internet to assist in avoiding the detection of psychological symptom dissimulation. *Professional Psychology: Research and Practice*, *33*(3), 294–299. doi:10.1037/0735-7028.33.3.294
- Ruiz Téllez, A. (1999). El problema de la Incapacidad Temporal y la Atención Primaria. La Gran oportunidad. La gran amenaza. Propuesta de un cambio. En semFYC (Ed.), *Incapacidad Temporal: encrucijada ética, clínica y de gestión*. (pp. 16–57). Barcelona: semFYC.
- Samuel, R. Z. y Mittenberg, W. (2005). Determination of malingering in disability evaluations. *Primary Psychiatry*, *12*(12), 60–68.
- Sánchez Crespo, G. (2002). *Sensibilidad de las escalas e indicadores de validez en el perfil de personalidad del MMPI-2* (1ª ed.). Salamanca: Ediciones Universidad de Salamanca.
- Sánchez Crespo, G., Jiménez Gómez, F. y Ampudia Rueda, A. (2008). Detectando el perfil simulador en el MMPI-2. *Revista de Psicología*, *26*(2), 277–298.
- Sánchez Crespo, G., Jiménez Gómez, F., Ampudia Rueda, A. y Merino Barragán, V. (2012). In search of a fast screening method for detecting the malingering of cognitive impairment. *The European Journal of Psychology Applied to Legal Context*, *4*(2), 135–158.
- Santamaría, P. (2009). *Adaptación española del MMPI-2-RF*. Madrid: TEA Ediciones.

- Santamaría, P. (2012). Metodología de investigación en simulación. En H. González-Ordi, P. Santamaría y P. Capilla (Eds.), *Estrategias de detección de la simulación. Un manual clínico multidisciplinar*. (pp. 105–152). Madrid: TEA Ediciones.
- Santamaría, P. y González-Ordi, H. (2012). Evaluación mediante autoinforme. En H. González-Ordi, P. Santamaría y P. Capilla (Eds.), *Estrategias de detección de la simulación. Un manual clínico multidisciplinar*. (pp. 185–274). Madrid: TEA Ediciones.
- Santamaría, P., González-Ordi, H., Capilla, P., Blasco, J. L. y Pallardo Durá, L. (2010, July). *La exageración de síntomas en el MMPI-2-RF*. Comunicación presented at the VII Congreso Iberoamericano de Psicología, Oviedo, Asturias, España. Recuperado de http://www.teaediciones.com/teaasp/Ejemplos/Exageracion_MMPI-2-RF.pdf
- Schinka, J. A. y Borum, R. (1993). Readability of adult psychopathology inventories. *Psychological Assessment*, 5(3), 384–386.
- Schinka, J. A. y Borum, R. (1994). Readability of normal personality inventories. *Journal of Personality Assessment*, 62(1), 95–101.
- Schmand, B. y Lindeboom, J. (2005). *ASTM Amsterdam Short-Term Memory Test* (Vol. 24). Leiden: PITS Test publishers.
- Schmand, B., Lindeboom, J., Schagen, S., Heijt, R., Koene, T. y Hamburger, H. L. (1998). Cognitive complaints in patients after whiplash injury: the impact of malingering. *Journal of Neurology, Neurosurgery & Psychiatry*, 64(3), 339–343.
- Schretlen, D., Neal, J. y Lesikar, S. (2000). Screening for malingered mental illness in a court clinic. *American Journal of Forensic Psychology*, 18(1), 5–16.

- Schretlen, D., Wilkins, S. S., Van Gorp, W. G. y Bobholz, J. H. (1992). Cross-validation of a psychological test battery to detect faked insanity. *Psychological Assessment*, 4(1), 77–83. doi:10.1037/1040-3590.4.1.77
- Schroeder, R. W., Baade, L. E., Peck, C. P., VonDran, E. J., Brockman, C. J., Webster, B. K. y Heinrichs, R. J. (2011). Validation of MMPI-2-RF Validity Scales in Criterion Group Neuropsychological Samples. *The Clinical Neuropsychologist*, 26(1), 129–146. doi:10.1080/13854046.2011.639314
- Sellbom, M. y Bagby, R. M. (2008). Response styles on multi-scale inventories. En R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 182–206). Nueva York: The Guilford Press.
- Sellbom, M. y Bagby, R. M. (2010). Detection of Overreported Psychopathology With the MMPI-2 RF Form Validity Scales. *Psychological Assessment*, 22(4), 757–767. doi:10.1037/a0020825
- Sellbom, M., Toomey, J. A., Wygant, D. B., Kucharski, L. T. y Duncan, S. (2010). Utility of the MMPI-2-RF (Restructured Form) validity scales in detecting malingering in a criminal forensic setting: A known-groups design. *Psychological Assessment*, 22(1), 22–31. doi:10.1037/a0018222
- Sellbom, M., Wygant, D. B. y Bagby, M. (2012). Utility of the MMPI-2-RF in detecting non-credible somatic complaints. *Psychiatry Research*, 197(3), 295–301. doi:10.1016/j.psychres.2011.12.043
- Sharland, M. J. y Gfeller, J. D. (2007). A survey of neuropsychologists' beliefs and practices with respect to the assessment of effort. *Archives of Clinical Neuropsychology*, 22(2), 213–223. doi:10.1016/j.acn.2006.12.004
- Sharp, L. K. y Lipsky, M. S. (2002). Screening for depression across the lifespan: a review of measures for use in primary care settings. *American Family Physician*, 66(6), 1001–1008.

- Shemer, H. (2011). *Cross-cultural equivalence of the MMPI-2 RF: A comparison between an Israeli and an American normal sample* (Tesis doctoral). Recuperado de ProQuest Dissertations & Theses. (UMI No. 3415799)
- Slick, D. J., Hopp, G., Strauss, E. y Thompson, G. B. (1997). *VSVT, Victoria Symptom Validity Test: Version 1.0, Professional Manual*. Psychological Assessment Resources Odessa, FL.
- 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 (Neuropsychology, Development and Cognition: Section D)*, 13(4), 545–561. doi:10.1076/1385-4046(199911)13:04;1-Y;FT545
- Slick, D. J., Tan, J. E., Strauss, E. H. y Hultsch, D. F. (2004). Detecting malingering: a survey of experts? practices. *Archives of Clinical Neuropsychology*, 19(4), 465–473. doi:10.1016/j.acn.2003.04.001
- Smith, A. M. (2010). *Assessing Personality Disorders Using the MMPI-2-RF* (Tesis de maestría). Recuperado de http://rave.ohiolink.edu/etdc/view?acc_num=kent1279132568
- Smith, G. P. y Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law*, 25(2), 183–189.
- Steenstra, I. A., Verbeek, J. H., Heymans, M. W. y Bongers, P. M. (2005). Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. *Occupational and Environmental Medicine*, 62(12), 851–860. doi:10.1136/oem.2004.015842

- Steffan, J. S., Kroner, D. G. y Morgan, R. D. (2007). Effect of Symptom Information and Intelligence in Dissimulation: An Examination of Faking Response Styles by Inmates on the Basic Personality Inventory. *Assessment*, *14*(1), 22–34. doi:10.1177/1073191106295404
- Stevens, A., Friedel, E., Mehren, G. y Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research*, *157*(1-3), 191–200. doi:10.1016/j.psychres.2007.01.003
- Sweet, J. J. (1999). Malingering: differential diagnosis. En J. J. Sweet (Ed.), *Forensic neuropsychology: Fundamentals and practice* (pp. 255–285). Lisse, Holanda: Swets & Zeitlinger.
- Sweet, J. J. (2008). Neuropsychology and the law: Malingering assessment in perspective. En J. Morgan y J. J. Sweet (Eds.), *Neuropsychology of Malingering Casebook* (pp. 3–8). Nueva York: Psychology Press.
- Sweet, J. J. y Guidotti Breting, L. (2013). Symptom Validity Test Research: Status and Clinical Implications. *Journal of Experimental Psychopathology*, *4*(1), 6–19. doi:10.5127/jep.022311
- Tan, J. E., Slick, D. J., Strauss, E. y Hultsch, D. F. (2002). How'd They Do It? Malingering Strategies on Symptom Validity Tests. *The Clinical Neuropsychologist (Neuropsychology, Development and Cognition: Section D)*, *16*(4), 495–505. doi:10.1076/clin.16.4.495.13909
- Tellegen, A. y Ben-Porath, Y. S. (2008). *MMPI-2-RF technical manual*. Minneapolis: University of Minnesota Press.
- Tellegen, A., Ben-Porath, Y. S., McNulty, J. L., Arbisi, P. A., Graham, J. R. y Kaemmer, B. (2003). *The MMPI-2 restructured clinical (RC) scales: Development, validation, and interpretation*. Minneapolis: University of Minnesota Press.

- Teuber, H. L. (1969). Neglected aspects of the posttraumatic syndrome. En A. E. Walker, W. F. Caveness y M. Critchley (Eds.), *The late effects of head injury* (pp. 13–34). Springfield, Ill: C.C.Thomas.
- Theron, F. H., Spangenberg, J. J., Hugo, F. J., Emsley, R. A., Hemp, F. y Maritz, J. S. (2001). Validation of psychometric scales for malingering in a student sample. *SAJP*, 7(4), 96–102.
- Thomas, M. L. y Locke, D. E. C. (2010). Psychometric properties of the MMPI-2-RF Somatic Complaints (RC1) Scale. *Psychological Assessment*, 22(3), 492–503. doi:10.1037/a0019229
- Tombaugh, T. N. (1996). TOMM: Test of memory malingering. *North Tonawanda, NY: Multi-Health Systems*.
- Turner, H. M. y Bernard, R. M. (2006). Calculating and synthesizing effect sizes. *Contemporary Issues in Communication Science and Disorders*, 33, 42–55.
- Van Beilen, M., Griffioen, B. T., Gross, A. y Leenders, K. L. (2009). Psychological assessment of malingering in psychogenic neurological disorders and non-psychogenic neurological disorders: relationship to psychopathology levels. *European Journal of Neurology*, 16(10), 1118–1123. doi:10.1111/j.1468-1331.2009.02655.x
- Van Der Heijden, P. ., Egger, J. I. . y Derksen, J. J. L. (2010). Comparability of Scores on the MMPI-2-RF Scales Generated With the MMPI-2 and MMPI-2-RF Booklets. *Journal of Personality Assessment*, 92(3), 254–259. doi:10.1080/00223891003670208
- Van der Werf, S. P., Prins, J. B., Jongen, P. J., van der Meer, J. W. y Bleijenberg, G. (2000). Abnormal neuropsychological findings are not necessarily a sign of cerebral impairment: a matched comparison between chronic fatigue syndrome and multiple sclerosis. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, 13(3), 199–203.

- Van Egmond, J., Kummeling, I. y Balkom, T. (2005). Secondary gain as hidden motive for getting psychiatric treatment. *European Psychiatry: The Journal of the Association of European Psychiatrists*, 20(5-6), 416–421. doi:10.1016/j.eurpsy.2004.11.012
- Van Hout, M. S. E., Schmand, B., Wekking, E. M. y Deelman, B. G. (2006). Cognitive functioning in patients with suspected chronic toxic encephalopathy: evidence for neuropsychological disturbances after controlling for insufficient effort. *Journal of Neurology, Neurosurgery & Psychiatry*, 77(3), 296–303.
- Veltri, C. O. C. y Williams, J. E. (2013). Does the Disorder Matter? Investigating a Moderating Effect on Coached Noncredible Overreporting Using the MMPI-2 and PAI. *Assessment*, 20(2), 199–209. doi:10.1177/1073191112464619
- Verdejo, A., Alcazar-Corcoles, M. A., Gomez-Jarabo, G. A. y Perez-Garcia, M. (2004). Pautas para el desarrollo científico y profesional de la neuropsicología forense. *Revista de Neurología*, 39(1), 60–73.
- Victor, T. L., Boone, K. B. y Kulick, A. D. (2010). My head hurts just thinking about it. *Journal of the International Neuropsychological Society*, 16(06), 1151–1152. doi:10.1017/S1355617710000858
- Viglione, D. J., Wright, D. M., Dizon, N. T., Moynihan, J. E., DuPuis, S. y Pizitz, T. D. (2001). Evading Detection on the MMPI-2: Does Caution Produce More Realistic Patterns of Responding? *Assessment*, 8(3), 237–250. doi:10.1177/107319110100800301
- Vitacco, M. J., Rogers, R., Gabel, J. y Munizza, J. (2006). An Evaluation of Malingering Screens with Competency to Stand Trial Patients: A Known-Groups Comparison. *Law and Human Behavior*, 31(3), 249–260. doi:10.1007/s10979-006-9062-8

- Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. Chichester, Inglaterra: John Wiley & Sons.
- Vrij, A. (2004). Why professionals fail to catch liars and how they can improve. *Legal and Criminological Psychology*, 9(2), 159–181.
- Vrij, A., Dragt, A. y Koppelaar, L. (1992). Interviews with ethnic interviewees: Non-verbal communication errors in impression formation. *Journal of Community & Applied Social Psychology*, 2(3), 199–208.
- Walker, L. J. y Frimer, J. A. (2011). The science of moral development. *Social Development: Relationships in Infancy, Childhood, and Adolescence*, 235–262.
- Walters, G. D., Berry, D. T. R., Rogers, R., Payne, J. y Granacher, R. (2009). Feigned neurocognitive deficit: Taxon or dimension? *Journal of Clinical and Experimental Neuropsychology*, 31(5), 584–593. doi:10.1080/13803390802363728
- Walters, G. D., Rogers, R., Berry, D. T. R., Miller, H. A., Duncan, S. A., McCusker, P. J., ... Granacher, R. P. (2008). Malingering as a categorical or dimensional construct: The latent structure of feigned psychopathology as measured by the SIRS and MMPI-2. *Psychological Assessment*, 20(3), 238–247. doi:10.1037/1040-3590.20.3.238
- Walters, G. D., White, T. W. y Greene, R. L. (1988). Use of the MMPI to identify malingering and exaggeration of psychiatric symptomatology in male prison inmates. *Journal of Consulting and Clinical Psychology*, 56(1), 111–117. doi:10.1037/0022-006X.56.1.111
- Watson, C., Quilty, L. C. y Bagby, R. M. (2011). Differentiating Bipolar Disorder from Major Depressive Disorder Using the MMPI-2-RF: A Receiver Operating Characteristics (ROC) Analysis. *Journal of Psychopathology and Behavioral Assessment*, 1–7. doi:10.1007/s10862-010-9212-7

- Wedding, D. y Faust, D. (1989). Clinical judgment and decision making in neuropsychology. *Archives of Clinical Neuropsychology*, 4(3), 233–265.
- Weiss, P. A., Bell, K. J. y Weiss, W. U. (2009). Use of the MMPI-2 Restructured Clinical (RC) Scales in Detecting Criminal Malingering. *Journal of Police and Criminal Psychology*, 25(1), 49–55. doi:10.1007/s11896-009-9056-9
- West, S. G. y Finch, J. F. (1997). Personality measurement: Reliability and validity issues. En R. Hogan, J. Johnson y S. Briggs (Eds.), *Handbook of personality psychology* (pp. 143–164). San Diego, CA: Academic Press.
- Wetter, M. W. y Corrigan, S. K. (1995). Providing Information to Clients About Psychological Tests: A Survey of Attorneys' and Law Students' attitudes. *Professional Psychology: Research and Practice*, 26(5), 474–477.
- Whiteside, D., Dunbar-Mayer, P. y Waters, D. (2009). Relationship between Tomm performance and PAI validity scales in a mixed clinical sample. *The Clinical Neuropsychologist*, 23(3), 523–533. doi:10.1080/13854040802389169
- Whitney, K., Davis, J., Shepard, P. y Herman, S. (2008). Utility of the Response Bias Scale (RBS) and other MMPI-2 validity scales in predicting TOMM performance. *Archives of Clinical Neuropsychology*, 23(7-8), 777–786. doi:10.1016/j.acn.2008.09.001
- Widows, M. R. y Smith, G. P. (2005). *SIMS: Structured Inventory of Malingered Symptomatology. Professional manual*. Lutz, FL: PAR: Psychological Assessment Resources.
- Wisdom, N. M., Callahan, J. L. y Shaw, T. G. (2010). Diagnostic utility of the structured inventory of malingered symptomatology to detect malingering in a forensic sample. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 25(2), 118–125. doi:10.1093/arclin/acp110

- Wise, E. A. (2002). Relationships of personality disorders with MMPI-2 malingering, defensiveness, and inconsistent response scales among forensic examinees. *Psychological Reports, 90*(3; PART 1), 760–766.
- Wise, E. A. (2009). Selected MMPI-2 Scores of Forensic Offenders in a Community Setting. *Journal of Forensic Psychology Practice, 9*(4), 299–309. doi:10.1080/15228930902936048
- World Health Organization. (2011). *Global status report on noncommunicable diseases 2010: Description of the global burden of NCDs, their risk factors and determinants*. Ginebra: Autor. Recuperado de http://www.who.int/nmh/publications/ncd_report_full_en.pdf
- Wygant, D. B., Anderson, J. L., Sellbom, M., Rapier, J. L., Allgeier, L. M. y Granacher, R. P. (2011). Association of the MMPI-2 Restructured Form (MMPI-2-RF) Validity Scales with Structured Malingering Criteria. *Psychological Injury and Law, 4*(1), 13–23. doi:10.1007/s12207-011-9098-z
- Wygant, D. B., Ben-Porath, Y. S. y Arbisi, P. A. (2004). Development and initial validation of a scale to detect infrequent somatic complaints. Presented at the 39th Annual Symposium on Recent Developments of the MMPI-2/MMPI-A, Minneapolis, MN.
- Wygant, D. B., Ben-Porath, Y. S., Arbisi, P. A., Berry, D. T. R., Freeman, D. B. y Heilbronner, R. L. (2009). Examination of the MMPI-2 Restructured Form (MMPI-2-RF) Validity Scales in Civil Forensic Settings: Findings from Simulation and Known Group Samples. *Archives of Clinical Neuropsychology, 24*(7), 671–680. doi:10.1093/arclin/acp073
- Wygant, D. B., Sellbom, M., Ben-Porath, Y. S., Stafford, K. P., Freeman, D. B. y Heilbronner, R. L. (2007). The relation between symptom validity testing and MMPI-2 scores as a function of forensic evaluation context. *Archives of Clinical Neuropsychology, 22*(4), 489–499.

- Wygant, D. B., Sellbom, M., Gervais, R. O., Ben-Porath, Y. S., Stafford, K. P., Freeman, D. B. y Greene, R. L. (2010). Further Validation of the MMPI-2 and MMPI-2-RF Response Bias Scale: Findings From Disability and Criminal Forensic Settings. *Psychological Assessment*, 22(4), 745–756. doi:10.1037/a0020042
- Youngjohn, J. R. (1995). Confirmed Attorney Coaching Prior to Neuropsychological Evaluation. *Assessment*, 2(3), 279–283. doi:10.1177/1073191195002003007
- Youngjohn, J. R., Spector, J. y Mapou, R. L. (1998). Failure to Assess Motivation, Need to Consider Psychiatric Disturbance, and Absence of Objectively Verified Physical Pathology: Some Common Pitfalls in the Practice of Forensic Neuropsychology. *The Clinical Neuropsychologist*, 12(2), 233–236. doi:10.1076/clin.12.2.233.1997
- Youngjohn, J. R., Wershba, R., Stevenson, M., Sturgeon, J. y Thomas, M. L. (2011). Independent Validation of the MMPI-2-RF Somatic/Cognitive and Validity Scales in TBI Litigants Tested for Effort. *The Clinical Neuropsychologist*, 1–14. doi:10.1080/13854046.2011.554444
- Zaldivar Basurto, F., García Montes, J. M., López Rios, F., Molina Moreno, A. y Santiago, E. (2008a). Psicología jurídica: Familia y victimología. En F. J. Rodríguez, C. Bringas, F. Fariña, R. Arce y A. Bernardo (Eds.), *Adaptación y validación al castellano de la entrevista estructurada de síntomas informados –SIRS* (pp. 469–477). Oviedo: Ediciones de la Universidad de Oviedo. Recuperado de <http://gip.uniovi.es/T6EJD.pdf>
- Zaldivar Basurto, F., García Montes, J. M., López Rios, F., Molina Moreno, A. y Santiago, E. (2008b). Psicología jurídica: Familia y victimología. En F. J. Rodríguez, C. Bringas, F. Fariña, R. Arce y A. Bernardo (Eds.), *Adaptación y validación al castellano de la evaluación forense de Millar del test de síntomas (M-FAST)* (pp. 461–468). Oviedo: Ediciones de la Universidad de Oviedo. Recuperado de <http://gip.uniovi.es/T6EJD.pdf>

Apéndice A. Cuestionario de percepción de los profesionales de la salud de la prevalencia de simulación por trastornos en incapacidad temporal.

Datos del médico

Edad..... SEXO..... Años de ejercicio profesional.....
Especialidad..... Antigüedad en la empresa.....
Centro de trabajo..... Provincia desarrollo profesional.....

CUESTIONARIO

1-Indique, numerando de 1 a 10 las patologías en las que usted considera que se dan más casos de simulación (1=máxima prevalencia de simulación; 10 = mínima prevalencia de simulación):

- Fibromialgia /Síndrome de Fatiga Crónica
- TCE leve
- TCE moderado
- Esguince cervical
- Cervicalgia Aguda
- Cervicalgia Crónica
- Lumbalgia o ciática aguda
- Lumbalgia o ciática crónica
- Depresión
- Ansiedad
- Fiebre de origen desconocido
- Síndrome de colon irritable
- Otros cuadros no mencionados y que usted desee puntuar de 1 a 10
-
-

2-En las mismas patologías de la pregunta anterior estime, a su criterio, el porcentaje que hay de exageración de síntomas/simulación:

- Fibromialgia /Síndrome de Fatiga Crónica
- TCE leve
- TCE moderado
- Esguince cervical
- Cervicalgia Aguda
- Cervicalgia Crónica
- Lumbalgia o ciática aguda
- Lumbalgia o ciática crónica
- Depresión
- Ansiedad
- Fiebre de Origen desconocido

- Síndrome de colon irritable
- Otros cuadros no mencionados y que usted desee valorar
-

3-En qué tramos de edad piensa usted que se dan más casos de exageración de síntomas

simulación: numere de 1 (más casos) a 5 (menos casos)

- 16 a 25 años
- 26 a 35 años
- 36 a 45 años
- 46 a 55 años
- 55 al final de la vida laboral

4-Escriba las 5 profesiones en las que, desde su experiencia y en el medio en el que usted trabaja, se producen más casos de exageración de síntomas /simulación (la primera en la que usted ha observado más casos y así hasta 5)

- 1-.....
- 2-.....
- 3-.....
- 4-.....
- 5-.....

5- Qué salario medio mensual bruto piensa usted que tiene con más frecuencia el paciente que exagera/simula síntomas (numere de 1 a 6 de más frecuente a menos)

- menos de 1.000 euros/mes
- de 1.000 a 1.500 “
- de 1.500 a 2.000 ”
- de 2.000 a 2.500 “
- de 2.500 a 3.000 “
- más de 3.000 euros /mes

6-En qué tipo de contingencia cree usted que se produce más simulación: (marque con una x)

- Accidente de trabajo
- Enfermedad profesional
- Contingencia Común.

7-Entre los pacientes que usted piensa que simulan/exageran síntomas: escriba el porcentaje que estima usted que lo hace en las siguientes situaciones:

- % simulan el mecanismo de producción
- % perpetúan síntomas una vez que estos han mejorado
- % simulan la patología completa (no existe y la simulan)
- % simulan o exageran síntomas por que adoptan el rol de enfermo
- % simulan síntomas por que tiene así un beneficio externo (tribunal, incapacidad, invalidez, seguro coche....)
- % simulan por que existe un conflicto laboral

8-Entre los pacientes que usted cree que exageran/simulan síntomas (escriba el porcentaje que usted estime)

-----% son mujeres

-----% son varones

9-Escriba, por orden de importancia, los 4 datos que a usted le hacen sospechar la existencia de exageración de síntomas/simulación:

1-----

2-----

3-----

4-----

Apéndice B. Instrucciones a simuladores análogos de psicopatología

Condición de trastorno psicótico

“A continuación vas a responder a una prueba de evaluación psicológica. Imagina que has cometido un delito violento. La policía te ha detenido y ha establecido claramente tu culpabilidad en el hecho delictivo, de modo que no hay dudas sobre tu implicación. Tu única opción es intentar minimizar la condena alegando la existencia de un trastorno mental severo que hace que no fueras responsable de tus actos. Debes convencer a los peritos evaluadores que cuando hiciste lo que hiciste no eras tú, estabas como loco, no eras dueño de tus actos. Que en este momento te sientes extraño, como si no fueras tú, como si tu mente estuviera fuera de control y hubieras perdido el juicio.

Ahora te van a evaluar para comprobarlo. Deberás simular la existencia de dicho trastorno mental sin que se note que estás mintiendo ya que de ello dependerá el tiempo que vas a pasar en la cárcel y el tipo de prisión que se te asignará. Contesta a la prueba sin perder de vista esta intención, intentando que tus respuestas sean lo más convincentes posibles, pero sin que puedan detectar que estás mintiendo”.

Condición de daño neurológico

“A continuación vas a responder a una prueba de evaluación psicológica. Imagina que has tenido un accidente y aunque no tienes grandes secuelas, has solicitado una indemnización económica. Alegas que te ha afectado física, mental y psicológicamente.

Ahora te van a evaluar para comprobarlo. Deberás simular la existencia de dichas secuelas sin que se note que estás mintiendo ya que de ello dependerá la indemnización que vas a recibir. Contesta a la prueba sin perder de vista esta intención, intentando que tus respuestas sean lo más convincentes posibles, pero sin que puedan detectar que estás mintiendo”.

Condición de trastorno afectivo

“A continuación vas a responder a una prueba de evaluación psicológica. Imagina que quieres conseguir una baja laboral porque tienes serios problemas con tus jefes, encuentras muchas dificultades para desarrollar tu trabajo y sentirte realizado laboralmente. Para ello alegas problemas emocionales, como sentirte triste, angustiado, apático, sin ilusiones, sin esperanzas y con muchas dificultades para hacer cualquier cosa por muy sencilla que ésta sea. Necesitas urgentemente la baja y por ello no tienes más opción que intentar simular lo mejor que puedas tus problemas emocionales.

Ahora te van a evaluar para comprobar la existencia de dichos problemas. Deberás simularlos lo mejor posible ya que de ello dependerá que consigas o no la baja laboral, tan importante para ti en este momento. Contesta a la prueba sin perder de vista esta intención, intentando que tus respuestas sean lo más convincentes posibles, pero sin que puedan detectar que estás mintiendo”.

Apéndice C. Instrucciones a simuladores análogos de lumbalgia

“A continuación vas a responder a una prueba de evaluación psicológica.

Imagina que para acceder a una serie de ventajas económicas y laborales o conseguir un subsidio o pensión afirmas que te encuentras psicológica y físicamente muy afectado porque sufres dolores intensos en la parte baja de la espalda, que te impiden estar sentado en la misma posición mucho tiempo, coger pesos excesivos o hacer movimientos bruscos o intensos; por lo que te resulta muy difícil moverte y desplazarte a trabajar y, una vez en tu puesto de trabajo, mantener la concentración y la eficacia durante mucho tiempo, debido a que el dolor va siendo cada vez más intenso a medida que transcurre el día.

Ahora te van a evaluar para comprobarlo. Deberás simular la existencia de dichos problemas sin que se note que estás mintiendo ya que de ello dependerá la indemnización que vas a recibir. Contesta a la prueba sin perder de vista esta intención, intentando que tus respuestas sean lo más convincentes posibles, pero sin que puedan detectar que estás mintiendo”.