



Simulación de enfermedad mental incapacitante para obtener una incapacidad laboral temporal: detección y morfología

Bárbara G. Amado¹, Manuel Vilarinho^{2,3,*}, Magdalena Mainar², e Irene Ágreda²

¹ Universidad Internacional de Valencia, España

² Unidade de Psicoloxía Forense, Universidade de Santiago de Compostela, España

³ Departamento de Ciencia Política e Socioloxía, Facultade de Psicoloxía, Universidade de Santiago de Compostela, España

Resumen: *Antecedentes/Objetivo:* En las evaluaciones psicológicas cuyos resultados pueden conllevar la obtención de algún beneficio o ganancia debe sospecharse la posibilidad de simulación. Este es el caso de la incapacidad laboral temporal por causas psicológicas, casuística para la que se han estimado altas prevalencias de simulación. Por ello, se diseñó un estudio con el objetivo de examinar la utilidad del *Symptom Checklist-90-Revised* (SCL-90-R) en la detección de psicopatología simulada, motivada por el incentivo externo de obtener una incapacidad laboral temporal. *Método:* Un total de 182 participantes completaron el SCL-90-R en un diseño de simulación, bajo dos condiciones diferentes: una respuesta siguiendo instrucciones estándar y, otra, instrucciones de simulación de síntomas psicológicos con el fin de obtener una incapacidad laboral temporal. *Resultados:* Los resultados revelaron que la población trabajadora es, en su mayoría, capaz de simular tanto síntomas clínicos como trastornos mentales con una intensidad suficiente para justificar una incapacidad laboral temporal (ganancia secundaria). Para ello, emplearon dos estrategias principales: la agrupación indiscriminada de síntomas (simulación de deterioro en todas las dimensiones clínicas, incluso superior al observado en la población psiquiátrica) y la severidad de síntomas (simulación de una intensidad extrema de los síntomas, también superior a la de la población psiquiátrica). Finalmente, se observó que el género no desempeña un papel significativo en la simulación. *Conclusiones:* Se discuten las implicaciones de estos hallazgos para la evaluación forense. **Palabras clave:** Incapacidad laboral temporal. Psicopatología. Simulación. SCL-90-R. Agrupación indiscriminada de síntomas. Severidad de síntomas.

Title: Malingering of disabling mental illness to obtain a temporary work disability benefit: detection and morphology.

Abstract: *Background/Objective:* In psychological assessments where the results may lead to obtaining some benefit or gain, malingering should be suspected. This is the case for temporary work disability due to psychological causes, for which high prevalence of malingering have been estimated. Therefore, a study was designed to examine the utility of the SCL-90-R in detecting malingered psychopathology motivated by the external incentive of obtaining a temporary work disability. *Method:* A total of 182 participants completed the SCL-90-R in a simulation design, under two conditions: responding under standard instructions and under instructions for malingering of psychological cause to obtain a temporary work disability. *Results:* The results revealed that the working population is almost entirely capable of malingering both clinical symptoms and mental disorders with sufficient severity to obtain a temporary work disability (secondary gain). They employed two strategies: indiscriminant symptom endorsement (malingering impairment across all clinical dimensions, even greater than the psychiatric population) and symptom severity (malingering an extreme intensity, even greater than the psychiatric population). Finally, it was found that the gender factor does not play a significant role in malingering. *Conclusions:* The implications of the results for forensic assessments are discussed.

Keywords: Temporary work disability. Psychopathology. Malingering. SCL-90-R. Indiscriminant symptom endorsement. Symptom severity.

Introducción

La necesidad de sospechar simulación en evaluaciones clínicas con implicaciones forenses, en las que los resultados puedan condicionar la obtención de algún beneficio o ganancia, es una advertencia recurrente en la literatura científica (Catwright & Donkin, 2020; Rogers, 2018a), así como en las guías clínicas (American Psychiatric Association [APA], 2013). Ignorar esta prescripción implica que, en los juicios penales, la evaluación del daño psicológico podría no ser suficiente para enervar el principio de presunción de inocencia (ningún inocente ha de ser condenado; ya que un simulador puede ser diagnosticado como paciente); mientras que, en procesos civiles, podría conducir a la pérdida de la demanda, debido a la imposibilidad de demostrar la realidad del daño o de las secuelas clínicas, al no descartarse cualquier probabilidad de simulación. La estimación de la prevalencia de la simulación es un tema controvertido, debido a la ausencia de un criterio de clasificación válido (ground truth) y consensuado para su medición (p. e., impresiones clínicas, diagnós-

tico diferencial forense, test de validez del rendimiento, evaluaciones de discapacidad de la Administración de la Seguridad Social), así como a la alta variabilidad según el contexto de la evaluación (p. e., clínico, forense, trastorno de estrés posttraumático, traumatismo craneoencefálico, dolor crónico). En cualquier caso, las cifras son muy elevadas y se alejan del error aceptable en el ámbito forense (en la evaluación forense de incapacidades, el error Tipo II, falsos negativos, debe ser cero; Arce, 2017). Así, Larrabee (2003) estimó la prevalencia de la simulación en el contexto de la evaluación neuropsicológica en un 40±10%, y Young (2015) tasó una prevalencia general del 15±15%, con valores más altos en contextos forenses. Uno de los ámbitos de evaluación donde se sospecha una alta prevalencia de simulación es el de la incapacidad laboral, tanto temporal como permanente (Collado et al., 2014; Domínguez et al., 2017; Yoxall et al., 2018), siendo especialmente relevante la incapacidad laboral por causas psicológicas, ya que la naturaleza de las enfermedades que podrían justificarla, a diferencia de la mayoría de las patologías físicas, carece de una organicidad fácilmente demostrable que facilite su objetivación (Goldwaser y Goldwaser, 2018). De hecho, las patologías con mayor prevalencia de simulación en el ámbito de las incapacidades son los trastornos del estado de ánimo y el dolor crónico (Pina et al., 2022).

Considerando los costes económicos y sanitarios asocia-

* Correspondence address [Dirección para correspondencia]:
Manuel Vilarinho Vázquez. Facultade de Psicoloxía, Universidade de Santiago de Compostela (Spain). Email: manuel.vilarino@usc.es
(Article received: 04-06-2025; revised: 11-06-2025; accepted: 01-07-2025)

dos a la simulación (Pascual y Salces, 2018), la necesidad de controlar este fenómeno se torna elevada, cuando no imperativa, especialmente si parte de dichos costes puede estar sustentada en el fraude (Chafetz y Underhill, 2013). En psicología se han desarrollado diversas técnicas para la detección del engaño, algunas centradas en el análisis de la credibilidad del testimonio (Cabezas-García et al., 2022; Leal et al., 2023; Montes et al., 2024; Sanmarco et al., 2023; Selaya et al., 2024), y otras en la validez de los síntomas informados (González et al., 2022; Puente-López et al., 2024; Puente-López et al., 2023).

Los profesionales encargados de realizar estas evaluaciones identifican la simulación como un fenómeno habitual, al tiempo que reconocen la ausencia de procedimientos y criterios estandarizados para su detección (Domínguez et al., 2017). No obstante, existen pruebas de evaluación psicológica de síntomas clínicos que podrían ser utilizadas en este contexto. El MMPI-2 y su versión reformulada, el MMPI-2-RF, el PAI y el SCL-90-R, entre otros, cuentan con un importante respaldo empírico para su uso tanto clínico como forense (Graham, 2011; Greene, 2011; Hawes y Boccaccini, 2009; Redondo et al., 2019; Rogers et al., 2003; Sharf et al., 2017; Vilariño et al., 2018; Wolf y Miller, 2014; Yoxall et al., 2018), ya que cumplen con el doble propósito psicológico-forense: la evaluación del daño y el diagnóstico diferencial de simulación.

De los instrumentos antes referidos, el SCL-90-R destaca por su óptima respuesta a esta doble demanda. A través de 90 ítems se obtiene sencilla y rápidamente una impresión diagnóstica de la persona evaluada, y con base en sus tres índices globales (GSI, PSDI y PST) se puede estudiar la sospecha de simulación. De hecho, su utilidad para el estudio de la simulación se ha evidenciado en casuísticas de diversa índole. En esta línea, existen hallazgos que revelan su capacidad para discriminar entre simuladores y pacientes genuinos aquejados de cuadros clínicos caracterizados por el dolor (McGuire & Shores, 2001; Torres et al., 2010; Wallis & Bogduk, 1996), así como entre psicopatología real y simulada en diseños de simulación (Sullivan & King, 2010) y, específicamente, en casuísticas de violencia de género y acoso escolar (Arce et al., 2013; Vilariño et al., 2009). Es más, Martínez et al. (2011) encontraron que el SCL-90-R se mostraba útil para la detección de psicopatología simulada en evaluaciones de incapacidad laboral permanente, cuando era empleado de forma combinada con otras medidas.

Para la detección de la simulación se puede acudir al análisis de estilos de respuesta específicos (Rogers, 2018b). En este sentido, se han establecido dos grandes categorías: una centrada en patrones de síntomas improbables debido a su naturaleza atípica o inusual, que raramente se presentan en pacientes genuinos (e. g., síntomas raros, síntomas improbables, combinación de síntomas), y otra relacionada con presentaciones exageradas de síntomas, en las que la frecuencia e intensidad del sufrimiento superan las observadas en pacientes reales. Dentro de esta segunda categoría, altamente prevalente entre simuladores (van Impelen et al., 2014), se

enmarcan dos estrategias de simulación: la agrupación indiscriminada de síntomas, caracterizada por informar de un número elevado de síntomas que rara vez se da en pacientes genuinos, y la severidad de síntomas, consistente en asignar tal gravedad (extremadamente elevada) a los síntomas referidos que resulta inusual entre pacientes reales (Fariña et al., 2014; Vilariño et al., 2020; Vilariño et al., 2013). Ambos tipos de patrones son evaluados a través de los índices globales del SCL-90-R (GSI, PST, PSDI), los cuales han demostrado ser altamente eficaces tanto en la clasificación precisa de respuestas simuladas como en la discriminación entre respuestas honestas y simuladas (Arce et al., 2013; Lees-Haley, 1989; Vilariño et al., 2009; Wallis y Bogduk, 1996).

En las últimas décadas, se ha enfatizado la importancia de incorporar una perspectiva de género en la investigación científica, especialmente cuando el objeto de estudio involucra a personas. No obstante, es frecuente encontrar cierta confusión en los trabajos científicos respecto a los conceptos de sexo y género, en el sentido de que muchos estudios tienden a centrarse en las diferencias entre sexos, mientras que el género constituye un constructo que trasciende dichas distinciones. Conscientes de esta controversia, y sin ánimo de profundizar en ella, en el presente trabajo se asimilarán ambas variables bajo el término *género*, con el objetivo de analizar si existen diferencias entre hombres y mujeres en la simulación de síntomas. En este sentido, la literatura científica recoge abundosas pruebas de diferencias en diversas dimensiones del funcionamiento psicológico (Ching et al., 2014; Novo et al., 2015; Riecher-Rössler, 2017; Strömwall et al., 2014; Vinagre-González et al., 2023).

Sobre la base de la revisión precedente, se diseñó un estudio de simulación con el objetivo general de analizar la utilidad del SCL-90-R para la detección de sintomatología psicológica simulada motivada por un incentivo externo: la obtención de una incapacidad laboral temporal. A su vez, se pretenden alcanzar cuatro objetivos específicos. Primero, conocer si la población general es capaz de simular sintomatología psicológica que podría motivar una incapacidad laboral temporal en el SCL-90-R. Segundo, estimar y cuantificar la efectividad de los indicadores globales del SCL-90-R (GSI, PSDI y PST) en la detección de la simulación. Tercero, examinar las características de los patrones de respuesta simulada en relación con la presencia de estrategias de simulación y la tipología de síntomas simulados. Cuarto, estudiar el efecto del factor género en la simulación de sintomatología incapacitante para la práctica laboral.

Método

Participantes

Un total de 182 personas participaron en el estudio, con edades que variaban entre los 20 y los 65 años ($M = 41.87$, $SEM = 1.04$). La muestra estaba equiparada en género, $\chi^2(1) = 3.16$, ns , y edad, $t(180) = 0.60$, $p = .550$. Del total, 103 (56.6%) eran mujeres, cuyas edades oscilaban entre los 20 y

los 65 años ($M = 42.53$, $SEM = 1.33$), y 79 hombres, con edades entre los 20 y los 65 años ($M = 41.15$, $SEM = 1.67$).

Diseño y procedimiento

Se adoptó un diseño de simulación intra-sujeto (Rogers, 2018a) en el que los participantes respondieron al SCL-90-R bajo instrucciones de respuesta estándar e instrucciones de simulación. En la primera condición, cumplieron el SCL-90-R siguiendo las instrucciones del manual (respuesta honesta). En la segunda condición, los participantes fueron instruidos para que simularan una incapacidad laboral temporal, que no permanente, motivada por la causa que estimaran, como la consecución de una indemnización, descanso laboral, exigencias laborales, ambiente laboral negativo, etc. Se les concedió una semana para que prepararan la evaluación y no se les dio instrucción alguna sobre cómo simular. Se les solicitaba que respondieran con rapidez (primer impulso) y sin poder volver atrás. El tiempo de respuesta osciló entre los 12 y 15 minutos. Por último, se llevó a cabo un *de-briefing* dirigido a verificar el recuerdo, compromiso y comprensión de los participantes con la tarea. Además, se calculó el cambio clínico significativo para cada participante i.e., $CS = (X_{hr} - X_{mr})/SD_{Diff} > 1.96$, donde X_{hr} es la puntuación de respuesta honesta, X_{mr} es la puntuación de respuesta de simulación, y SD_{Diff} es el error estándar de la diferencia entre las dos puntuaciones. Así, se observó un cambio clínico significativo en todos los participantes. Por otro lado, se implementó un diseño inter-sujeto para comparar las puntuaciones entre hombres y mujeres con el objetivo de analizar los efectos del género. La participación fue voluntaria y sujeta a consentimiento informado. Los datos obtenidos fueron tratados acorde a la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales, 2018.

Como criterios de inclusión y exclusión se aplicaron los siguientes: estar o haber estado en condición de empleado por cuenta ajena; estar en edad laboral o no jubilados; ausencia de enfermedad física o trastorno mental que les incapacite para trabajar; que hubieran simulado el padecimiento de, al menos, 5 síntomas (simulación significativa) adicionales a los manifestados bajo instrucciones de respuesta honesta; y que ninguno de los índices globales advirtiera de sospecha de negación de síntomas o gravedad.

Instrumento de medida

El ajuste psicológico se midió con la adaptación española del *Symptom Check List 90-R* (SCL-90-R; Derogatis, 1994, 2002). El SCL-90-R se compone de 90 ítems (síntomas psicológicos, psiquiátricos y somáticos) al que los participantes responden en una escala de respuesta de cinco puntos que gradúa la intensidad del sufrimiento: *Nada* (0), *un poco* (1), *moderadamente* (2), *bastante* (3) y *muchísimo* (4). La medida se estructura en torno a 9 dimensiones sintomáticas [somatización (SOM), obsesión-compulsión (O-C), sensibilidad interpersonal (INT), depresión (DEP), ansiedad (ANS), hostili-

dad (HOS), ansiedad fóbica (FOB), ideación paranoide (PAR) y psicoticismo (PSI)], 7 ítems adicionales y 3 índices globales [Índice de Severidad Global (GSI), Total de Síntomas Positivos (PST) y el Malestar Referido a Síntomas Positivos (PSDI)].

Los índices globales tienen una utilidad diferente en el ámbito clínico (medida de la intensidad; Derogatis, 1994, 2002) y el forense (evaluación de la distorsión en las respuestas: simulación y disimulación; Arce, Fariña, & Vilariño, 2015; McGuire & Shores, 2001; Vilariño et al., 2009). En este trabajo se interpretaron las medidas en estos índices en términos de intensidad (evaluación clínica) para estimar la capacidad de simulación y la sospecha de simulación (evaluación forense), tomando como punto de corte una puntuación $T \geq 66.45$ (percentil 95; significatividad estadística, clínica y forense).

Análisis de datos

Se ejecutó la prueba t de student para muestras relacionadas con el fin de comparar las medias de los autoinformes de síntomas psicológicos obtenidos bajo las instrucciones estándar (respuesta honesta) y de acuerdo con las instrucciones de simulación (respuesta simulada). El tamaño del efecto se obtuvo con la d de Cohen para medidas repetidas [$d_m = (M_1 - M_2)/SD_{Diff}$], interpretándose cualitativamente como de tamaño pequeño ($d = 0.20$), moderado ($d = 0.50$), grande ($d = 0.80$) y más que grande ($d > 1.20$; Arce, Fariña, Seijo et al., 2015). El error del modelo empírico fue estimado con la Probabilidad de una Puntuación de Inferioridad (PIS; Arias et al., 2020), siendo trivial si $PIS < .05$ o el 95% intervalo tiene .05; entre trivial y común si $.05 < PIS < .50$; común si $PIS = .50$ (el CI tiene .50); entre común y normal si $.50 < PIS < .95$; y normal si el límite inferior de $PIS > .95$ o el 95% CI tiene .95.

Asimismo, y con el objetivo de comparar las medias de las respuestas bajo instrucciones de simulación con la población psiquiátrica se aplicó la prueba t de student para una muestra (valor de prueba: media de la población psiquiátrica en la estandarización española del SCL-90-R) computando el tamaño del efecto con la d de Cohen y cuantificando el error del modelo con PIS. También con el objetivo de comparar las medias entre hombres y mujeres en la medida de síntomas psicopatológicos, así como en las medidas de sospecha de simulación, se ejecutaron t para muestras independientes, el tamaño del efecto con d y el error del modelo estadístico con PIS. La prevalencia observada de casos clínicos ($T \geq 66.45$ i.e., percentil ≥ 95 significación clínica, estadística y forense en las dimensiones clínicas y los índices globales) se contrastó con una constante (.05, .50 y .95, prevalencia trivial, común y normal, respectivamente) estimando el valor de Zeta para la diferencia de proporciones. La magnitud del efecto fue medida con la b de Cohen, interpretándose la magnitud del tamaño del efecto como pequeño, moderado, grande y más que grande cuando $b = 0.20$, $b = 0.50$, $b = 0.80$

y $b > 1.20$, respectivamente (Arce, Fariña, Seijo et al., 2015; Cohen, 1988).

La asociación entre género y casuística de síntomas psicopatológicos ($T \geq 66.45$) y la sospecha de simulación ($T \geq 66.45$) en los índices globales, se estimó con chi cuadrado de Pearson (no se observó ninguna contingencia de frecuencia esperada de celdas < 5). El tamaño del efecto fue computado con la razón de prevalencia, interpretándose la magnitud como pequeña ($PR = 1.44$), moderada ($PR = 2.47$), grande ($PR = 4.25$) y más que grande ($PR > 8.82$; Arce, Fariña, Seijo et al., 2015), y cuantificándose el efecto con el Índice de Efecto Incremental (EII) i.e., el incremento (o reducción en negativos) en la probabilidad observada en una categoría sobre otra categoría (Vilariño et al., 2022).

Resultados

Evaluación de la capacidad para simular síntomas clínicos y trastornos para una incapacidad laboral temporal

Los resultados (ver Tabla 1) mostraron que los participantes tienen una capacidad de simulación de enfermedad mental significativa en las 9 dimensiones clínicas. Esta capacidad de simulación es de un tamaño del efecto más que grande ($d > 1.20$), siendo trivial el error del modelo estadístico ($PIS \leq .05$) a excepción de la dimensión obsesión-compulsión que es significativo, pero muy cercano a trivial (límite inferior del PIS 95% CI fue .069). En suma, la capacidad de la población de trabajadores de simulación de síntomas clínicos para motivar una incapacidad laboral temporal es prácticamente total.

Tabla 1

Comparación de medias de las respuestas honestas y simuladoras. Prueba *t* para muestras relacionadas.

Variables	<i>t</i>	M_{rs}	M_{th}	$d[95\% \text{ CI}]$	PIS[95% CI]
SOM	20.20***	2.33	0.83	1.50[1.54, 1.46]	.067[.041, .093]
O-C	17.31***	2.33	1.01	1.28[1.24, 1.32]	.100[.069, .131]
INT	21.00***	2.49	0.86	1.55[1.51, 1.59]	.061[.036, .086]
DEP	24.94***	2.69	0.84	1.85[1.81, 1.89]	.032[.014, .051]
ANS	26.03***	2.56	0.65	1.93[1.89, 1.97]	.027[.010, .044]
HOS	20.59***	2.46	0.74	1.35[1.31, 1.39]	.089[.060, .118]
FOB	23.01***	2.36	0.38	1.69[1.65, 1.73]	.046[.024, .068]
PAR	23.35***	2.68	0.85	1.74[1.70, 1.78]	.041[.021, .061]
PSI	26.10***	2.38	0.32	1.93[1.89, 1.97]	.027[.010, .044]

Nota. $g(182)$; M_{rs} : Media de las respuestas bajo instrucciones de simulación (respuestas de simulación); M_{th} : Media de las respuestas bajo instrucciones estándar (respuesta honesta); $d[95\% \text{ CI}]$: d de Cohen [95% Intervalo de Confianza]; PIS[95% CI]: Probabilidad de una Puntuación de Inferioridad [95% Intervalo de confianza]; *** $p < .001$.

Asimismo, los resultados (ver Tabla 2) también evidenciaron que los participantes disponían de la capacidad para simular intensidad clínica i.e., severidad global, total de síntomas positivos y malestar referido a síntomas positivos. La magnitud del efecto fue también más que grande y el modelo de error estadístico, trivial ($PIS \leq .05$).

Tabla 2

Comparación de los efectos del factor respuesta (simulación vs. honestidad) en el sufrimiento global informado. Prueba *t* para muestras relacionadas

Índice	<i>t</i>	M_{rs}	M_{th}	$d[95\% \text{ CI}]$	PIS[95% CI]
GSI	26.34***	2.50	0.75	1.94[1.90, 1.98]	.026[.009, .042]
PST	20.56***	80.01	45.05	1.74[1.70, 1.78]	.041[.020, .061]
PSDI	26.80***	2.75	1.41	2.00[1.96, 2.04]	.023[.008, .038]

Nota. $g(182)$; M_{rs} : Media de las respuestas bajo instrucciones de simulación (respuestas de simulación); M_{th} : Media de las respuestas bajo instrucciones estándar (respuesta honesta); $d[95\% \text{ CI}]$: d de Cohen [95% Intervalo de Confianza]; PIS[95% CI]: Probabilidad de una Puntuación de Inferioridad [95% Intervalo de confianza]; *** $p < .001$.

Los resultados del contraste de la probabilidad observada de casos clínicos ($T \geq 66.45$) con una prevalencia trivial (constante .05) mostró una habilidad significativa para simular en las nueve dimensiones sintomáticas (ver $Z_{.05}$ en la Tabla 3) con una magnitud de tamaño del efecto más que grande ($b_{.05} > 1.20$). Por tanto, la capacidad para simular psicopatología es altamente significativa. Además, esta capacidad es significativamente (ver $Z_{.50}$ en la Tabla 3) más que común ($> .50$) con una magnitud del efecto grande ($b > .80$) en somatización, obsesión-compulsión, depresión, ansiedad, hostilidad y ansiedad fóbica, y más que grande ($b > 1.20$) en sensibilidad interpersonal, ideación paranoide y psicoticismo. En cualquier caso, la capacidad para simular malestar clínico significativo no es normal ($Z_{.95}$ significativamente inferior a .95) en ninguna de las dimensiones (i. e., podría no ser aplicable a toda la población).

Adicionalmente, las respuestas de simulación revelaron una gravedad significativamente mayor en las nueve dimensiones sintomáticas con respecto a la población psiquiátrica (ver Tabla 4). Además, la magnitud del efecto fue más que grande ($d > 1.20$), mientras que el error del modelo estadístico resultante fue trivial ($PIS \leq .05$). Este hallazgo revela el empleo de la estrategia de simulación “agrupación indiscriminada de síntomas”, sin que sean capaces de discernir entre psicopatología con implicaciones para una incapacidad temporal (p. e., depresión) y otra potencialmente permanente (p. e., psicoticismo).

Tabla 3

Capacidad para simular: casos clínicos en las dimensiones clínicas

Variables	$f(p)$	Z.05	$h.05[95\% \text{ CI}]$	Z.50	$h.5[95\% \text{ CI}]$	Z.95	$h.95[95\% \text{ CI}]$
SOM	145(.797)	46.24***	2.48[2.42, 2.54]	8.01***	0.90[0.83, 0.97]	-9.47***	-0.69[-0.74, -0.64]
O-C	154(.846)	49.27***	2.66[2.60, 2.72]	9.34***	1.08[1.01, 1.15]	-6.44***	-0.51[-0.55, -0.47]
INT	159(.874)	51.00***	2.77[2.71, 2.83]	10.09***	1.20[1.13, 1.27]	-4.70***	-0.39[-0.43, -0.35]
DEP	150(.824)	47.91***	2.58[2.52, 2.64]	8.74***	1.00[0.93, 1.07]	-7.79***	-0.59[-0.64, -0.54]
ANS	155(.852)	49.64***	3.28[3.22, 3.34]	9.50***	1.11[1.04, 1.18]	-6.07***	-0.48[-0.52, -0.44]
HOS	144(.791)	45.87***	2.46[2.40, 2.52]	7.85***	0.88[0.81, 0.95]	-9.84***	-0.71[-0.76, -0.66]
FOB	157(.863)	50.32***	3.32[3.26, 3.38]	9.79***	1.15[1.08, 1.22]	-5.39***	-0.44[-0.48, -0.40]
PAR	167(.918)	53.73***	2.68[2.63, 2.73]	11.28***	1.40[1.33, 1.47]	-1.98*	-0.19[-0.22, -0.16]
PSI	162(.890)	52.00***	2.85[2.80, 2.90]	10.52***	1.27[1.20, 1.34]	-3.71***	-0.32[-0.35, -0.29]

Nota. $f(p)$: frecuencia observada (probabilidad observada); Z.05: Puntuación Zeta para la diferencia entre la probabilidad observada y una constante, un efecto trivial (.05); $h.05$: b de Cohen para Z.05; Z.50: Puntuación Zeta para la diferencia entre la probabilidad observada y una constante, un efecto común (.50); $h.50$: b de Cohen para Z.50; Puntuación Zeta para la diferencia entre la probabilidad observada y una constante, un efecto normal (.95); $h.95$: b de Cohen para Z.95; * $p < .05$; *** $p < .001$.

Tabla 4

Comparación de la media de las respuestas de simulación con la media de los valores de la población psiquiátrica (valor de prueba) en las dimensiones sintomáticas. Prueba t para una muestra

Variable	t	M_{rs}	M_{psiqui}	$d[95\% \text{ CI}]$	PIS[95% CI]
SOM	9.32***	2.33	1.67	2.45[2.41, 2.49]	.001[-.002, .004]
O-C	17.43***	2.52	1.42	2.96[2.92, 3.00]	.001[-.002, .004]
INT	8.78***	2.49	1.89	2.68[2.64, 2.72]	.004[-.002, .010]
DEP	16.14***	2.69	1.61	2.99[2.95, 3.03]	.001[-.002, .004]
ANS	13.13***	2.56	1.64	2.72[2.68, 2.76]	.003[-.003, .009]
HOS	15.45***	2.46	1.18	2.20[2.16, 2.24]	.014[.002, .026]
FOB	16.15***	2.36	1.04	2.13[2.09, 2.17]	.017[.004, .030]
PAR	19.90***	2.68	1.33	2.91[2.87, 2.95]	.001[-.002, .004]
PSI	17.68***	2.38	1.03	2.31[2.27, 2.35]	.010[-.000, .020]

Nota. $g(181)$. M_{rs} : Media de las respuestas bajo instrucciones de simulación (respuestas de simulación); M_{psiqui} : media de las respuestas de la población psiquiátrica; d de Cohen [95% Intervalo de Confianza]; PIS[95% CI]: Probabilidad de una Puntuación de Inferioridad [95% Intervalo de confianza]; *** $p < .001$.

Del mismo modo, la comparación de respuestas simuladas con población psiquiátrica en los índices globales de malestar (ver Tabla 5) también arrojó que, en condiciones de simulación, los participantes informaban de una gravedad significativamente mayor a la de la población psiquiátrica. Además, la magnitud del efecto fue más que grande en el PST y GSI (el 95% CI tuvo 1.20) y moderada en el PSDI ($d = 0.50$). El error del modelo estadístico (probabilidad de simuladores que informan de una severidad en los índices inferior a la media de la población psiquiátrica) fue trivial (PIS $\leq .05$) en el GSI y PST (esto es, los simuladores informan normalmente, $\geq 95\%$, de más severidad global y más síntomas positivos que la población psiquiátrica) y entre trivial y común ($.05 < \text{PIS} < .50$) en el PSDI (aproximadamente 1/3 de los simuladores informaban de menos severidad en los síntomas positivos que la media de la población psiquiátrica y, por el contrario, unos 2/3 de más severidad). En consecuencia, estos resultados avalan que, además de la agrupación indiscriminada de síntomas, los simuladores también siguen una estrategia de “severidad de síntomas”, como lo

demuestra el hecho de superar en frecuencia (síntomas en las 9 dimensiones clínicas) e intensidad (puntuaciones en los tres índices globales de malestar) a la población psiquiátrica.

Tabla 5

Comparación de la media de las respuestas de simulación con la media de los valores de la población psiquiátrica (valor de prueba) en los índices globales de malestar. Prueba t para una muestra

Índice	t	M_{rs}	M_{psiqui}	$d[95\% \text{ CI}]$	PIS[95% CI]
GSI	15.79***	2.49	1.49	1.17[1.11, 1.23]	.121[-.087, .155]
PST	24.60***	79.58	52.73	1.82[1.76, 1.88]	.034[.015, .053]
PSDI	7.20***	2.74	2.39	0.52[0.58, 0.46]	.302[.255, .349]

Nota. $g(181)$. M_{rs} : Media de las respuestas bajo instrucciones de simulación (respuestas de simulación); M_{psiqui} : Media de las respuestas de la población psiquiátrica; d de Cohen [95% Intervalo de Confianza]; PIS[95% CI]: Probabilidad de una Puntuación de Inferioridad [95% Intervalo de confianza]; *** $p < .001$.

Evaluación de la sensibilidad para detectar simulación de los índices globales de gravedad del malestar

El análisis de la sensibilidad de los índices globales de malestar (ver Tabla 6) reveló que el total de síntomas positivos (PST) clasifica normalmente (Capacidad para Detectar Correctamente Simulación [ASCM] = .95) respuestas simuladas; el índice de severidad global (GSI) entre común y normal ($.50 < \text{ASCM} < .95$) y el índice de malestar referido a síntomas positivos (PSDI), comúnmente (ASCM = .50). Adicionalmente, 18 protocolos (10.1%) fueron clasificados como sospechosos de simulación por un índice, 54 (30.2%) por dos índices, y 97 (54.2%) por tres índices. Aplicado el *gold standard* para la sospecha forense de simulación (la clasificación de dos índices como sospecha de simulación; Vilariño et al., 2009), la eficacia de las medidas de simulación es significativa $\chi^2(1) = 117.12$, $p < .001$, $b = 3.54$. Por tanto, la probabilidad de error en la correcta clasificación de las respuestas de simulación (falsos negativos) con el *gold standard* es prácticamente nula (PIS $< .001$).

Tabla 6*Clasificación de sospecha de simulación en los índices globales*

Índice	$f(p)$	Z.05	$b.05$ [IC 95%]	Z.50	$b.5$ [IC 95%]	Z.95	$b.95$ [IC 95%]
GSI	154(.846)	49.27***	2.66[2.60, 2.72]	9.34***	1.08[1.01, 1.15]	-6.44***	-0.51[-0.55, -0.47]
PST	163(.948)	55.59***	3.15[3.11, 3.19]	12.09***	1.57[1.50, 1.64]	-0.12	-0.02[-0.04, 0.00]
PSDI	97(.533)	29.90***	1.67[1.60, 1.74]	0.89	0.09[0.06, 0.12]	-25.81***	-1.50[-1.57, -1.43]

Nota. $f(p)$: frecuencia observada (probabilidad observada); Z.05: Puntuación Zeta para la diferencia entre la probabilidad observada y una constante, un efecto trivial (.05); $b.05$: b de Cohen para Z.05; Z.50: Puntuación Zeta para la diferencia entre la probabilidad observada y una constante, un efecto común (.50); $b.50$: b de Cohen para Z.50; Puntuación Zeta para la diferencia entre la probabilidad observada y una constante, un efecto normal (.95); $b.95$: b de Cohen para Z.95; *** $p < .001$.

Estudio de los efectos del género sobre la capacidad para simular

Los resultados de la comparación de medias en las dimensiones sintomáticas (ver Tabla 7) para el factor género, mostró que hombres y mujeres informaban de sintomatología similar en somatización, sensibilidad interpersonal, ansiedad, hostilidad, ansiedad fóbica, ideación paranoide y psicoticismo. Sin embargo, en las dimensiones de obsesión-compulsión y depresión, las mujeres informaban de un sufrimiento significativamente mayor, aunque el tamaño del efecto era pequeño y la probabilidad de error del modelo estadístico era entre trivial y común. Adicionalmente, tampoco se hallaron diferencias en los índices globales de malestar (ver Tabla 8).

Tabla 7*Comparación de medias en las dimensiones sintomáticas para el factor género. Prueba t para dos muestras*

Variable	t	p	M_{mujeres}	M_{hombres}	d [95% CI]	PIS[95% CI]
SOM	1.71	.089	2.43	2.19	0.25[0.19, 0.31]	.401[.351, .451]
O-C	2.25	.026	2.64	2.36	0.33[0.27, 0.39]	.371[.321, .421]
INT	1.32	.187	2.57	2.39	0.19[0.13, 0.25]	.525[.474, .576]
DEP	2.64	.009	2.84	2.49	0.39[0.33, 0.45]	.348[.299, .397]
ANS	1.88	.061	2.67	2.41	0.28[0.22, 0.34]	.390[.340, .440]
HOS	0.35	.726	2.49	2.43	0.05[-0.01, 0.11]	.480[.429, .531]
FOB	0.74	.458	2.42	2.29	0.12[0.06, 0.18]	.452[.4001, .503]
PAR	0.70	.487	2.72	2.63	0.10[0.04, 0.16]	.460[.409, .511]
PSI	-0.45	.655	2.35	2.42	-0.07[-0.13, -0.01]	.528[.477, .579]

Nota. $g(181)$. M_{mujeres} : Media del grupo de mujeres; M_{hombres} = Media del grupo de hombres; d de Cohen[95% Intervalo de Confianza]; PIS[95% CI]: Probabilidad de una Puntuación de Inferioridad [95% Intervalo de confianza]; *** $p < .001$.

Tabla 8*Comparación de medias en los índices globales de malestar para el factor género. Prueba t para dos muestras*

Índice	t	p	M_{mujeres}	M_{hombres}	d [95% CI]	PIS
GSI	1.57	.118	2.59	2.39	0.24[0.18, 0.30]	.405
PST	1.72	.087	81.52	78.04	0.25[0.19, 0.31]	.401
PSDI	1.20	.233	2.81	2.69	0.17[0.11, 0.23]	.433

Nota. $g(181)$. M_{mujeres} : Media del grupo de mujeres; M_{hombres} = Media del grupo de hombres; d de Cohen[95% Intervalo de Confianza]; PIS[95% CI]: Probabilidad de una Puntuación de Inferioridad [95% Intervalo de confianza]; *** $p < .001$.

Asimismo, la probabilidad observada de casos clínicos en las dimensiones sintomáticas fue la misma para hombres y mujeres (ver Tabla 9). De hecho, el efecto incremental debido a la categoría género es trivial (ver IEE).

Tabla 9*Asociación entre género y casos clínicos en las dimensiones sintomáticas*

Variable	$f(p)_{\text{hombres}}$	$f(p)_{\text{mujeres}}$	χ^2	p	PR[95% CI]	IEE[95% CI]
SOM	81(.786)	64(.811)	0.16	.715	0.97[0.83, 1.12]	-.031[-.056, -.001]
O-C	88(.854)	66(.835)	0.12	.836	1.02[0.90, 1.16]	.022[.001, .043]
INT	89(.864)	70(.886)	0.20	.822	0.98[0.87, 1.09]	.025[.002, .048]
DEP	83(.806)	67(.848)	0.55	.557	0.95[0.83, 1.09]	.050[.018, .081]
ANS	87(.845)	68(.861)	0.09	.835	0.98[0.87, 1.11]	.019[-.001, .039]
HOS	77(.748)	67(.848)	2.74	.140	0.88[0.76, 1.02]	.118[.071, .165]
FOB	88(.854)	69(.873)	0.14	.829	0.98[0.87, 1.10]	.022[.001, .043]
PAR	93(.903)	74(.937)	0.68	.588	0.96[0.89, 1.05]	.036[.008, .063]
PSI	91(.883)	71(.899)	0.11	.814	0.98[0.89, 1.09]	.018[-.001, .037]

Nota. $g(1)$. $f(p)_{\text{mujeres}}$: frecuencia de sospecha de simulación en mujeres (proporción observada); $f(p)_{\text{hombres}}$: frecuencia de sospecha de simulación en hombres (proporción observada); PR[95% CI]: Razón de Prevalencia[95% Intervalo de Confianza]; IEE[95% CI]: Índice del Efecto Incremental[95% Intervalo de Confianza].

Con respecto a las medidas de sospecha de simulación (ver Tabla 10), los resultados no mostraron diferencias entre hombres y mujeres. Globalmente, estos hallazgos revelan que el factor género no desempeña un efecto significativo ni en la simulación ni en la detección de ésta.

Tabla 10*Relación entre género y sospecha de simulación en los índices globales de malestar*

Índice	$f(p)_{\text{mujeres}}$	$f(p)_{\text{hombres}}$	χ^2	p	PR[95% CI]	IEE[95% CI]
GSI	85(.825)	69(.873)	0.80	.413	0.94[0.84, 1.07]	-.055[-.775, .033]
PST	93(.903)	70(.886)	0.14	.808	1.02[0.92, 1.13]	.019[-.001, .039]
PSDI	56(.544)	41(.519)	0.12	.766	1.05[0.79, 1.38]	.046[.016, .076]

Nota. $g(1)$. $f(p)_{\text{mujeres}}$: frecuencia de sospecha de simulación en mujeres (proporción observada); $f(p)_{\text{hombres}}$: frecuencia de sospecha de simulación en hombres (proporción observada); PR[95% CI]: Razón de Prevalencia[95% Intervalo de Confianza]; IEE[95% CI]: Índice del Efecto Incremental[95% Intervalo de Confianza].

Discusión

Este trabajo presenta limitaciones que deben ser contempladas en la generalización de los resultados. De este modo, debido al tamaño de la muestra y su procedimiento de selección, un muestreo no probabilístico, los hallazgos deben generalizarse con cautela. Es posible que las características de los participantes no reflejen completamente las de la población trabajadora, lo que puede influir en los resultados cuantitativos, aunque no sobre las conclusiones cualitativas (es decir, la aceptación o refutación de hipótesis). Además, aunque se empleó un diseño de simulación de alta fidelidad, los resultados cuantitativos difieren de los obtenidos en estudios de campo, pero no cualitativamente (Amado et al., 2016; De Marchi y Balboni, 2018). Otra limitación es la ausencia de una muestra de casos reales de incapacidad laboral temporal,

que hubiera permitido determinar la especificidad de los índices globales del SCL-90-R en la detección de la simulación, así como contrastar las medidas de las dimensiones sintomáticas entre ambas muestras.

Con estas limitaciones en mente, los resultados obtenidos evidencian que la población de trabajadores está capacitada para simular psicopatología con potencial para motivar una incapacidad laboral temporal, a la vez que le otorga una intensidad clínica. Así, el contraste entre las condiciones de realidad y de simulación reveló puntuaciones significativamente superiores, con un tamaño del efecto más que grande, en la de simulación en todas las dimensiones sintomáticas e índices globales. Por tanto, la muestra fue capaz de fingir síntomas que realmente no padecía. Esta conclusión se ve reforzada por los resultados emanados del estudio de casos, en el que se plasmó una habilidad altamente significativa para simular ser caso clínico en todas las dimensiones sintomáticas del SCL-90-R, con unos tamaños del efecto de grande a más que grande. Esta evidencia resulta congruente con la investigación previa que advierte acerca de la facilitación que las tareas de reconocimiento de síntomas ejercen sobre la simulación (Arce et al., 2013; Arce, Fariña y Vilariño, 2015; Fariña et al., 2014). En este sentido, las tareas de reconocimiento, como el SCL-90-R, incorporan entre sus ítems los síntomas que podrían fundamentar una incapacidad laboral temporal, de tal forma que a la persona evaluada se le demanda exclusivamente la misión de seleccionar aquellos que resultan necesarios para alcanzar una simulación exitosa y que, cuando le sea requerido, gradúe la intensidad del sufrimiento en la respuesta (Vilariño et al., 2020).

Aunque se ha verificado la capacidad para simular psicopatología incapacitante en el SCL-90-R, esta deja su rastro en el patrón de respuesta encontrado en el estudio. Concretamente, se evidenció el empleo de dos estrategias de simulación: agrupación indiscriminada de síntomas, y severidad de síntomas. En relación con la primera, el contraste de las puntuaciones obtenidas en la condición de simulación con los valores de población psiquiátrica reveló un padecimiento sintomático significativamente mayor y un tamaño del efecto más que grande para la muestra simuladora en todas dimensiones sintomáticas del SCL-90-R. En este sentido, los participantes no discernían entre la psicopatología con implicaciones para la incapacidad temporal y la permanente. Esta limitación para lograr una simulación con especificidad se vio reforzada en el estudio de casos, cuando los tamaños del efecto asociados con la capacidad de simular ser un caso clínico fueron más que grandes en las dimensiones sintomáticas con potencial para una incapacidad permanente (p. e., ideación paranoide y psicoticismo) en comparación con los tamaños del efecto grandes en las dimensiones sintomáticas asociadas con la incapacidad temporal (p. e., depresión, ansiedad, somatización...). Sin embargo, los datos reales sobre prevalencia revelan que los trastornos de ansiedad y depresión son los que subyacen a la mayoría de las bajas laborales por causas psicológicas (Pina et al., 2022). Por tanto, es posible que el afán por lograr simulaciones exitosas sea el que

subyazca este énfasis por informar de sufrimiento psicopatológico reconocido por su alta gravedad. Por su parte, la severidad de síntomas se evidenció en las puntuaciones significativamente más altas, con tamaños del efecto que varían de moderados a más que grandes, que la muestra simuladora obtuvo en contraste con la población psiquiátrica. En resumen, para simular, los participantes indicaron un alto número de síntomas como presentes, a los cuales atribuían una gravedad extrema. Este hallazgo es consistente con investigaciones previas que habían detectado el uso de tales estrategias (Arce et al., 2013; Rogers, 2018b; Rogers et al., 2003; Vilariño et al., 2009; Vilariño et al., 2013; Vilariño et al., 2020).

El análisis de los índices globales del SCL-90-R mostró diferentes capacidades de detección. El PST fue altamente efectivo en la detección de la simulación clasificando como “normalmente”. El GSI tuvo un rendimiento ligeramente inferior clasificando entre “común y normalmente”. En contraste, el PSDI clasificó “comúnmente”. Por otro lado, su uso combinado mediante la aplicación del gold standard para la sospecha forense de simulación (sospecha de simulación por dos o más índices) mostró una efectividad significativa en la detección de la simulación. En este sentido, el 84.4% de los participantes fueron clasificados como sospechosos de simulación según este estándar. Además, se verificó que la probabilidad de error del gold standard en la clasificación correcta de respuestas simuladas era prácticamente nula, lo que significa que prácticamente ningún simulador sería identificado como paciente genuino (falso negativo). Estos hallazgos revelan, en línea con investigaciones previas (Arce et al., 2013; Sullivan & King, 2010; Torres et al., 2010; Vilariño et al., 2009; Wallis & Bogduk, 1996), que el SCL-90-R es una herramienta útil para la detección de la simulación. Sin embargo, en este contexto de evaluación, no es suficiente aplicar el gold standard para concluir sobre la sospecha de simulación; también debe realizarse un diagnóstico diferencial de daño grave (Arce, 2017); ya que clasificar a un paciente real como simulador resultaría inaceptable (error forense punible), pues se negaría la incapacidad a un paciente genuino. Además, deben realizarse investigaciones con individuos con incapacidad laboral temporal para observar cómo se comporta el SCL-90-R en estas poblaciones y determinar la especificidad de los índices globales del instrumento con los puntos de corte utilizados en este estudio. En cualquier caso, para lograr la mayor efectividad posible en la clasificación correcta, especialmente en el diagnóstico diferencial de daño grave, es esencial recurrir a una evaluación multi-método y multi-medida (Greene, 2011; Vilariño et al., 2013).

Globalmente, el estudio del factor género no reveló diferencias ni en la simulación ni en la detección de esta. La tendencia dominante fue la ausencia de diferencias entre hombres y mujeres (i.e., dimensiones sintomáticas, frecuencia de casos clínicos, detección de la simulación por los índices globales). Solo en las dimensiones de obsesión-compulsión y depresión, las mujeres obtuvieron puntuaciones significativamente más altas en la condición de simulación, aunque con un tamaño del efecto pequeño. De acuerdo con la literatura

tura, la prevalencia de psicopatología es mayor en las mujeres (Riecher-Rössler, 2017), y específicamente en el caso de la sintomatología depresiva y obsesivo-compulsiva en la adultez (APA, 2013; Girgus & Yang, 2015). Es posible que esta contingencia haya contribuido a las diferencias encontradas, ya que la familiarización con niveles psicopatológicos basales más altos pudo haber derivado en una mayor simulación de dicha sintomatología.

En resumen y como conclusión, se podría afirmar, en primer lugar, que la muestra fue capaz de simular sintomatología psicopatológica en el SCL-90-R con potencial para justificar una incapacidad laboral temporal. En segundo lugar, para llevar a cabo esta simulación, recurrieron a las estrategias de agrupación indiscriminada de síntomas y de severidad de síntomas. En tercer lugar, los índices globales del SCL-90-R podrían ser útiles para controlar la simulación, aunque sería necesario conocer su especificidad mediante estudios con pacientes reales. En cualquier caso, su uso aislado no sería suficiente para la detección total de la simulación, siendo necesario realizar un diagnóstico diferencial de daño grave en los casos bajo sospecha de simulación. En cuarto lugar, el factor género no parece jugar un papel significativo en la simulación. Finalmente, pensando en la práctica profesional, el SCL-90-R podría ser útil en este contexto. Además de su fácil aplicación y corta duración, proporciona información útil para el diagnóstico clínico y la sospecha de simulación. Sin

embargo, debe utilizarse en combinación con otras técnicas dentro de protocolos multi-método y multi-medida.

Finalmente, con la investigación futura en mente, sería de interés llevar a cabo nuevos estudios que combinen el uso del SCL-90-R con otras técnicas de evaluación que incluyan una tarea de conocimiento de síntomas para determinar la efectividad de su uso combinado, como la entrevista clínico-forense (Vilariño et al., 2013). Además, en estudios futuros se podrían demandar tareas de simulación especificando tipos concretos de trastornos o síntomas psicopatológicos (por ejemplo, depresión, somatización, ansiedad...), en lugar de otorgar libertad al participante en la simulación. También sería deseable contar con muestras de participantes entrenados para conocer su efecto sobre las características de la simulación y la sensibilidad de los índices globales. Finalmente, como se había mencionado, para determinar la especificidad de los índices globales en la detección de simulación, se deberían desarrollar investigaciones que incorporaran muestras clínicas con patologías diversas.

Información complementaria

Conflicto de interés.- Los autores declaran que no existe un conflicto de interés.

Financiación.- Sin financiación.

Referencias

- Amado, B. G., Arce, R., Fariña, F., & Vilariño, M. (2016). CBCA reality criteria in adults: A meta-analytic review. *International Journal of Clinical and Health Psychology*, 16(2), 201-210. <http://dx.doi.org/10.1016/j.ijchp.2016.01.002>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author.
- Arce, R. (2017). Análisis de contenido de las declaraciones de testigos: Evaluación de la validez científica y judicial de la hipótesis y la prueba forense [Content analysis of the witness statements: Evaluation of the scientific and judicial validity of the hypothesis and the forensic proof]. *Acción Psicológica*, 14(2), 171-190. <https://doi.org/10.5944/ap.14.2.21347>
- Arce, R., Fariña, F., & Quinteiro, I. (2013). Evaluación forense de la simulación en casos de acoso escolar [Malingering forensic evaluation in bullying cases]. *Aula Abierta*, 41, 67-74.
- Arce, R., Fariña, F., Seijo, D., & Novo, M. (2015). Assessing impression management with the MMPI-2 in child custody litigation. *Assessment*, 22(6), 769-777. <http://dx.doi.org/10.1177/1073191114558111>
- Arce, R., Fariña, F., & Vilariño, M. (2015). Daño psicológico en casos de víctimas de violencia de género: Estudio comparativo de las evaluaciones forenses [Psychological injury in intimate partner violence cases: A contrastive analysis of forensic measures]. *Revista Iberoamericana de Psicología y Salud*, 6, 72-80. <https://doi.org/10.1016/j.rips.2015.04.002>
- Arias, E., Arce, R., Vázquez, M. J., & Marcos, V. (2020). Treatment efficacy on the cognitive competence of convicted intimate partner violence offenders. *Anales de Psicología/Annals of Psychology*, 36(3), 427-435. <https://doi.org/10.6018/analesps.428771>
- Cabezas-García, M., de la Peña Fernández, M. E., & Andreu Rodríguez, J. M. (2022). Factores psicosociales asociados a la credibilidad del testimonio en Abuso Sexual Infantil [Psychosocial factors associated with the assessment of testimony credibility in Child Sexual Abuse]. *Anales de Psicología / Annals of Psychology*, 38(2), 307-315. <https://doi.org/10.6018/analesps.418161>
- Catwright, A., & Donkin, R. (2020). Knowledge of depression and malingering: An exploratory investigation. *Europe's Journal of Psychology*, 16, 32-44. <https://doi.org/10.5964/ejop.v16i1.1730>
- Chafetz, M. & Underhill, J. (2013). Estimated costs of malingered disability. *Archives of Clinical Neuropsychology*, 28(7), 633-639. <http://dx.doi.org/10.1093/arclin/act038>
- Ching, H., Daffern, M., & Thomas, S. (2014). Clarifying the characteristics of youth who have engaged in appetitive violence. *Journal of Forensic Psychiatry & Psychology*, 25, 451-463. <https://doi.org/10.1080/14789949.2014.933860>
- Cohen, J. (1988). *Statistical power analysis for behavioral sciences* (2nd ed.). LEA.
- Collado, A., Gomez, E., Coscolla R., Sunyol, R., Rivera, J., Altarriba, E., Carbonell, J., & Castells, X. (2014). Work, family and social environment in patients with fibromyalgia in Spain: An epidemiological study: EPIFFAC study. *BMC Health Services Research*, 14(1), 513. <http://dx.doi.org/10.1186/s12913-014-0513-5>
- De Marchi, B., & Balboni, G. (2018). Detecting malingering mental illness in forensics: Known-group comparison and simulation design with MMPI-2, SIMS and NIM. *PeerJ*, 6, e5259. <https://doi.org/10.7717/peerj.5259>
- Derogatis, L. R. (1994). *SCL-90-R: Administration, scoring and procedures manual* (3th ed.). NCS Pearson.
- Derogatis, L. R. (2002). *SCL-90-R Manual*. TEA Ediciones.
- Dominguez, A., Gómez, A. M., de la Fuente, J. L., López, M. J., García, P., & López, R. (2017). Opiniones de los inspectores médicos acerca de la distorsión clínica en la valoración de la incapacidad laboral en España [Opinions of medical inspectors on clinical distortion in the assessment of work disability in Spain]. *ARS MEDICA Revista de Ciencias Médicas*, 42(1), 13-19. <http://dx.doi.org/10.11565/arsmed.v42i1.546>
- Fariña, F., Arce R., Vilariño, M., & Novo, M. (2014). Assessment of the standard forensic procedure for the evaluation of psychological injury in intimate-partner violence. *Spanish Journal of Psychology*, 17, e32. <https://doi.org/10.1017/sjp.2014.30>
- Girgus, J. S., & Yang, K. (2015). Gender and depression. *Current Opinion in Psychology*, 4, 53-60. <https://doi.org/10.1016/j.copsyc.2015.01.019>
- Goldwaser A. M., & Goldwaser E. L. (2019). *The forensic examination: A handbook for the mental health professional*. Springer.
- Gonzalez, B., Novo, R., & Farrajota, P. (2022). Detecting underreport in real-world assessment contexts: The utility of multiple-scale indicators. *Anales de Psicología/Annals of Psychology*, 38(3), 555-564. <https://doi.org/10.6018/analesps.502511>
- Graham, J. R. (2011). *MMPI-2: Assessing personality and psychopathology* (5th ed.). Oxford University Press.

- Greene, R. L. (2011). *The MMPI-2/MMPI-2-RF: An interpretive manual* (3rd ed.). Allyn & Bacon.
- Hawes, S. W., & Boccacini, M. T. (2009). Detection of overreporting of psychopathology on the Personality Assessment Inventory: A meta-analytic review. *Psychological Assessment*, 21, 112-124. <https://doi.org/10.1037/a0015036>
- Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *Clinical Neuropsychologist*, 17(3), 410-425. <https://doi.org/10.1076/clin.17.3.410.18089>
- Leal, S., Vrij, A., Deeb, H., Burkhardt, J., Dabrowna, O., & Fisher, R. P. (2023). Verbal cues to deceit when lying through omitting information: Examining the effect of a model statement interview protocol. *European Journal of Psychology Applied to Legal Context*, 15, 1-8. <https://doi.org/10.5093/ejpal2023a1>
- Lees-Haley, P. R. (1989). Malingering emotional distress on the SCL-90-R: Toxic exposure and cancerphobia. *Psychological Reports*, 65, 1203-1208. <https://doi.org/10.2466/pr0.1989.65.3f.1203>
- Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. (2018). *Boletín Oficial del Estado*, 294, 119788-119857. <https://www.boe.es/boe/dias/2018/12/06/pdfs/BOE-A-2018-16673.pdf>
- Martínez, N., Orihuela, M. J., & Abeledo, M. (2011). Evaluación de simulación de incapacidad laboral permanente mediante el Sistema de Evaluación Global (SEG) [Assessing malingering of a permanent work incapacity through the Global Evaluation System (GES)]. *Revista Iberoamericana de Psicología y Salud*, 2(2), 207-224.
- McGuire, B. E., & Shores, E. A. (2001). Simulated pain on the Symptom Checklist 90-Revised. *Journal of Clinical Psychology*, 57(12), 1589-1596. <https://doi.org/10.1002/jclp.1121>
- Montes, A., Sanmarco, J., Cea, B., Marcos, V., & Seijo, D. (2024). Examining the persistence of central and peripheral information in honest witness. *Revista Iberoamericana de Psicología y Salud*, 15, 96-104. <https://doi.org/10.23923/j.rips.2024.02.079>
- Novo, M., Díaz-Saco, P., Amado, B. G., & Vilarinho, M. (2015). Efectos de contexto en la formación de juicios en un caso de violación [Contextual effects on judgment-making in a case of rape]. *Anuario de Psicología Jurídica*, 25, 39-45. <https://doi.org/10.1016/j.api.2015.02.002>
- Pascual, R., & Salces, L. (2019, March 24). Las bajas laborales costaron 14.400 millones en salarios y prestaciones en 2018 [Medical leaves cost €14.4 billion in wages and benefits in 2018]. *Cinco Días*. https://cincodias.elpais.com/cincodias/2019/03/22/economia/155327924_3_546916.html
- Pina, D., Puente-López, E., Ruiz-Hernández, J. A., Llor-Esteban, B., & Aguerrevere, L. E. (2022). Self-report measures for symptom validity assessment in whiplash-associated disorders. *European Journal of Psychology Applied to Legal Context*, 14(2), 73-81. <https://doi.org/10.5093/ejpal2022a7>
- Puente-López, E., Pina, D., Daugherty, J. C., Pérez-García, M., & Merten, T. (2024). Simulación y validez de la información de síntomas psicopatológicos en España: conceptos, métodos y desafíos [Malingering and validity of reported psychopathological symptoms in Spain: concepts, methods, and challenges]. *Revista Iberoamericana de Psicología y Salud*, 15, 66-79. <https://doi.org/10.23923/j.rips.2024.02.077>
- Puente-López, E., Pina, D., López-Nicolás, R., Iguacel, I., & Arce, R. (2023). The Inventory of Problems-29 (IOP-29): A systematic review and bivariate diagnostic test accuracy meta-analysis. *Psychological Assessment*, 35, 339-352. <https://doi.org/10.1037/pas0001209>
- Redondo, L., Fariña, F., Seijo, D., Novo, M., & Arce, R. (2019). A meta-analytical review of the responses in the MMPI-2/MMPI-2-RF clinical and restructured scales of parents in child custody dispute. *Anales de Psicología/Annals of Psychology*, 35(1), 156-165. <http://dx.doi.org/10.6018/analesps.35.1.338381>
- Riecher-Rössler, A. (2017). Sex and gender differences in mental disorders. *The Lancet Psychiatry*, 4(1), 8-9. [https://doi.org/10.1016/S2215-0366\(16\)30348-0](https://doi.org/10.1016/S2215-0366(16)30348-0)
- Rogers, R. (2018a). An introduction to response styles. In R. Rogers & S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4th ed., pp. 3-17). Guilford Press.
- Rogers, R. (2018b). Detection strategies for malingering and defensiveness. In R. Rogers & S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4th ed., pp. 3-17). The Guilford Press.
- Rogers, R., Sewell, K. W., Martin, M. A., & Vitacco, M. J. (2003). Detection of feigned mental disorders: A meta-analysis of the MMPI-2 and malingering. *Assessment*, 10, 160-177. <http://dx.doi.org/10.1177/1073191103252349>
- Sanmarco, J., Selaya, A., Marcos, V., & Montes, Á. (2023). Contrasting the efficacy of the content analysis categories of the Global Evaluation System to discern between experienced memories and video observed memories [Contrasting the efficacy of the content analysis categories of the Global Evaluation System to discern between experienced memories and video observed memories]. *Revista Iberoamericana de Psicología y Salud*, 14, 55-67. <https://doi.org/10.23923/j.rips.2023.02.066>
- Selaya, A., Vilarinho, M., & Arce, R. (2024). In search of an empirical definition of a social model for the assessment of the quality of memory. *Revista Iberoamericana de Psicología y Salud*, 15, 12-17. <https://doi.org/10.23923/j.rips.2024.01.071>
- Sharf, A. J., Rogers, R., Williams, M. M., & Henry, S. A. (2017). The effectiveness of the MMPI-2-RF in detecting feigned mental disorders and cognitive deficits: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment*, 39, 441-455. <http://dx.doi.org/10.1007/s10862-017-9590-1>
- Strömwall, L. A., Landström, S., & Alfredsson, H. (2014). Perpetrator characteristics and blame attributions in a stranger rape situation. *European Journal of Psychology Applied to Legal Context*, 6, 63-67. <https://doi.org/10.1016/j.ejpal.2014.06.002>
- Sullivan, K., & King, J. (2010). Detecting faked psychopathology: A comparison of two tests to detect malingered psychopathology using a simulation design. *Psychiatry Research*, 176, 75-81. <https://doi.org/10.1016/j.psychres.2008.07.013>
- Torres, X., Bailles, E., Collado, A., Taberner, J., Gutierrez, F., Peri, J. M., ... & Valdes, M. (2010). The symptom checklist-revised (SCL-90-R) is able to discriminate between simulation and fibromyalgia. *Journal of Clinical Psychology*, 66, 774-790. <https://doi.org/10.1002/jclp.20698>
- van Impelen, A., Merckelbach, H., Jelicic, M., & Merten, T. (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. *Clinical Neuropsychologist*, 28(8), 1336-1365. <https://doi.org/10.1080/13854046.2014.984763>
- Vilarinho, M., Amado, G., Martín-Peña, J., & Vázquez, M. J. (2020). La simulación del acoso laboral en el LIPT-60: Implicaciones para la evaluación psicológica forense [Feigning mobbing in the LIPT-60: Implications for forensic psychological assessment]. *Anuario de Psicología Jurídica*, 30, 83-91. <https://doi.org/10.5093/api2019a19>
- Vilarinho, M., Amado, B. G., Seijo, D., Selaya, A., & Arce, R. (2022). Consequences of child maltreatment victimization in internalizing and externalizing mental health problems. *Legal and Criminological Psychology*, 27, 182-193. <https://doi.org/10.1111/lcrp.12212>
- Vilarinho, M., Amado, B. G., Vázquez, M. J., & Arce, R. (2018). Psychological harm in women victims of intimate partner violence: Epidemiology and quantification of injury in mental health markers. *Psychosocial Intervention*, 27, 145-152. <https://doi.org/10.5093/pi2018a23>
- Vilarinho, M., Arce, R., & Fariña, F. (2013). Forensic-clinical interview: Reliability and validity for the evaluation of psychological injury. *European Journal of Psychology Applied to Legal Context*, 5(1), 1-21. <https://journals.copmadrid.org/ejpal/archivos/articulo20180219092235.pdf>
- Vilarinho, M., Fariña, F., & Arce, R. (2009). Discriminating real victims from feigners of psychological injury in gender violence: Validating a protocol for forensic settings. *European Journal of Psychology Applied to Legal Context*, 1(2), 221-243. https://zaguan.unizar.es/record/78802/files/texto_completo.pdf
- Vinagre-González, A. M., Puente-López, E., Aguilar-Cárceles, M. M., Aparicio-García, M. E., & Loinaz, I. (2023). Differences between men and women in the acceptance of gender roles and stereotypes in intimate partner violence. *Revista Iberoamericana de Psicología y Salud*, 14, 75-82. <https://doi.org/10.23923/j.rips.2023.02.068>
- Wallis, B. J., & Bogduk, N. (1996). Faking a profile: Can naive subjects simulate whiplash responses? *Pain*, 66, 223-227. [https://doi.org/10.1016/0304-3959\(96\)03044-8](https://doi.org/10.1016/0304-3959(96)03044-8)
- Wolf, E. J., & Miller, M. W. (2014). The Minnesota Multiphasic Personality Inventory-2 Restructured Form and posttraumatic stress disorder: Forensic applications and considerations. *Psychological Injury and Law*, 7, 143-152. <https://doi.org/10.1007/s12207-014-9193-z>
- Young, G. (2015). Malingering in forensic disability-related assessments: Prevalence 15 ± 15%. *Psychological Injury and Law*, 8, 188-199. <https://doi.org/10.1007/s12207-015-9232-4>
- Yoxall, J., Bahr, M., & O'Neill, T. (2017). Faking bad in workers compensation psychological assessments: Elevation rates of negative distortion scales on the personality assessment inventory in an Australian sample. *Psychiatry, Psychology and Law*, 24, 682-693. <https://doi.org/10.1080/13218719.2017.1291295>