

PRECISIÓN DIAGNÓSTICA Y MINUSVALORACIÓN DE SÍNTOMAS EN EL MMPI-2-RF

Guadalupe Sánchez-Crespo^{1*}, Amada Ampudia-Rueda** y Fernando Jiménez-Gómez*

**Facultad de Psicología, Universidad de Salamanca; **Facultad de Psicología,
Universidad Nacional Autónoma de México*

RESUMEN. Antecedentes. La detección de la simulación al contestar el Minnesota Multiphasic Personality Inventory (MMPI-2/RF) ha sido un tema que ha suscitado el interés de los investigadores por la elaboración de un grupo de escalas de validez que detectan el estilo de respuesta de las personas evaluadas, llegando a configurar una “segunda generación” de escalas detectoras de la simulación y disimulación de respuestas en el MMPI-2. El objetivo de este estudio se centra en analizar la precisión diagnóstica de las principales escalas simuladoras de minusvaloración de síntomas en la versión reestructurada (MMPI-2-RF). **Método.** Se trata de un diseño experimental en el que se establecen dos grupos, control (Sincero, con 309 participantes) y experimental (Simulador, con 278), asignados al azar, sin evidencia de patología alguna. Al grupo Simulador se le instruyó para que contestaran al cuestionario ofreciendo una imagen positiva de sí mismo. A las escalas de Validez propuestas (*L-r* y *K-r*) en el MMPI-2-RF (adaptación española) se les añadieron las escalas de Deseabilidad Social (*Wsd*) de Wiggins (1959), la de Otro engaño, de Nichols y Greene (1991) (*Odec*), la Superlativa (*S*), de Butcher y Han (1995), y la Deseabilidad social (*ESD*) de Edwards (1953). **Resultados.** Los resultados se muestran con el análisis de la fiabilidad (α de Cronbach), las diferencias estadísticas y “*d*” de Cohen y su precisión diagnóstica a través de la curva ROC (*Receiver Operating Characteristic*) comparando los grupos sincero y simulador. **Conclusiones.** Se señalan las escalas que aportaron mayor precisión diagnóstica que el *L-r* y el *K-r* en la detección de los simuladores de minimización de síntomas.

Palabras clave: MMPI-2-RF, simulación de minusvaloración de síntomas, Precisión diagnóstica, Curva ROC.

ABSTRACT. The detection of the simulation to the reply the Minnesota Multiphasic Personality Inventory (MMPI / 2 / RF) has been a theme that has oversubscribed the interest of them researchers by the elaboration of a group of scales of validity that detected the style of response of them people evaluated, coming to configure a “second generation” of scales detection of it simulation and dissimulation of answers in the MMPI-2. The objective of this study is focused in analyze the precision diagnostic of them main scales simulating of underestimation of symptoms in the version restructured (MMPI-2-RF). Method. Is of a design experimental in which are established two groups, control (sincere, with 309 participating) and experimental

¹ Correspondencia: Guadalupe Sánchez-Crespo. Universidad de Salamanca. Facultad de Psicología. Avda. de la Merced, 101, 37005.- Salamanca. E-mail: lupes@usal.es

(Simulator, with 278), assigned to the random, without evidence of Pathology any. The Simulator Group was instructed that they reply to the questionnaire providing a positive image of itself. To validity scales (I-r and k-r) proposals in the MMPI-2-RF (Spanish adaptation) were added to the scale of Social desirability (Wsd) of Wiggins (1959), the other deception, Nichols y Greene (1991) (Odec), the superlative (S), Butcher y Han (1995), and the social desirability (ESD) of Edwards (1953). Results. These are shown with the reliability analysis (α of Cronbach), statistical differences and "d" of Cohen and his diagnostic accuracy through the ROC (Receiver Operating Characteristic) curve comparing groups sincere and Simulator. Conclusions. Is point them scales that provided greater precision diagnostic that them I-r and k-r in the detection of them simulators of minimization of symptoms.

Keyword: MMPI-2-RF, underreporting, Diagnostic Accuracy, ROC Curve.

Introducción

En los contextos judiciales, tanto de Familia, de lo Penal como de lo Social, la minusvaloración de los síntomas de una determinada enfermedad, discapacidad o trastorno, es considerada como un tipo de conducta simulada, deliberadamente ejercida, que tiene la intención de engañar a otra persona o institución y que persigue obtener un beneficio, bien económico, psicológico o de eximir responsabilidades u obligaciones.

En la evaluación psicológica forense evaluar el grado de cooperación del examinado es determinante (Butcher, Hass, Greene y Nelson, 2015). Dado el formato que presenta el cuestionario del MMPI-2, al poder responder "verdadero/falso" a los ítems ofrece la posibilidad de responder de dos maneras: según se tenga en cuenta el contenido del ítem, o no. En el primer caso el sujeto puede tratar de manipular la respuesta para que se refleje sus propios intereses; en el segundo caso el sujeto puede contestar al azar o no entender el contenido del ítem (p. e. personas con dificultades con el idioma). Precisamente, por esta importante posibilidad en las diferentes actualizaciones y versiones (MMPI-2 y MMPI-2-RF), los investigadores se han interesado en la detección del engaño (Ben-Porath y Tellegen (2008/2011).

Algunos investigadores han centrado sus estudios en la detección de las variables de simulación en contextos, tanto clínicos (Jiménez-Gómez, Sánchez-Crespo, y Ampudia-Rueda, 2013; Jiménez, Sánchez y Tobón, 2009; Rogers, 2008) como judiciales con diversidad de técnicas y estrategias. Otros autores se han servido de la metodología de la curva ROC (*Receiver Operating Characteristic*), con el objetivo de poder detectar su precisión diagnóstica (Nicholson, Mouton, Bagby y Buis, 1997; Pelegrina, Ruiz-Soler y Wallace, 2000).

La estructura interpretativa del MMPI-2-RF se ha actualizado considerablemente en dos grandes apartados y cada uno de ellos presenta subapartados relevantes: 1) la validez del protocolo que incluye el estudio de la consistencia de la respuesta con posible presencia de patrones de exageración (*overreporting*) de síntomas (Sellbom y Bagby, 2010) o minusvaloración (*underreporting*) (Sellbom y Bagby, 2008) y 2) las diferentes escalas Sustantivas que se agrupan para estudiar específicamente la posible existencia de alteraciones somáticas/cognitivas, alteraciones emocionales, del pensamiento, comportamentales, relaciones interpersonales, intereses del sujeto evaluado o consideraciones diagnósticas y terapéuticas (Greene, 2011).

En el MMPI-2-RF, las escalas de Validez han sufrido pequeñas modificaciones en su denominación (por ejemplo, la *L (Lie)*- ahora es denominada como *Uncommon Virtues*

scale y la escala *K* como *Adjustment Validity Scale*, añadiéndole al acrónimo original de la escala una “*r*” (*revisada*) para distinguirlas de su antecesor, el MMPI-2 (por ejemplo, *L-r*, *K-r*).

El objetivo de este estudio es poder apreciar la aportación de la precisión diagnóstica de las escalas de validez que detectan la minusvaloración de síntomas para obtener las conclusiones correspondientes.

Metodología

Diseño

Se trata de un estudio experimental formado por dos grupos, control (Sincero) y experimental (Simulador), en el que participaron un total de 587 personas sin evidencia de patología alguna. El grupo control se encuentra formado por 309 participantes, elegidos al azar de un conjunto mayor de protocolos obtenidos del proceso de elaboración de los baremos españoles del MMPI-2, contesta de forma sincera y honesta el cuestionario que es denominado como "Sincero". El grupo experimental está compuesto por 278 participantes que, elegidos y elegidos igualmente de forma aleatoria de un grupo mayor, contestan al cuestionario siguiendo las instrucciones para ofrecer una imagen favorable de sí mismo con el objetivo de conseguir un determinado beneficio denominándosele como "Simulador".

Participantes

De los 309 que integran el grupo "Sincero", 163 son mujeres, con una media de edad de 32.57 años y un rango entre 19 y 64 años, y 163 son varones, con una media de edad de 33.29 años y un rango de 19 a 64, igual que las mujeres. El grupo "Simulador" lo componen 278 personas de las que 161 son mujeres, con una edad media de 26.64 años y un rango establecido entre 19 y 59; y 117 varones con una edad media de 28.08 años y un rango entre 19 y 63 años.

Instrumentos

El material empleado en este estudio es el Inventario Multifásico de personalidad de Minnesota-2-Reestructurado (MMPI-2-RF) en su adaptación española (Santamaría, 2009). Dicha adaptación, para detectar la posible minusvaloración de los síntomas (*underreporting*), dispone de las escalas de Validez *Virtudes inusuales* (*L-r*) y Validez de ajuste (*K-r*). En el estudio que se presenta se han añadido otras cuatro escalas más fruto de la investigación: la Escala Superlativa *S* de Butcher y Han (1995), la Escala de Otro engaño *Odecp* de Nichols y Greene (1991), la escala de Deseabilidad social *Wsd* de Wiggins (1959), y la escala de Deseabilidad social de Edwards (1953). A estas cuatro últimas, siguiendo las directrices del MMPI-2-RF, se les añadieron la "r" de "reestructurada".

Procedimiento

Todos los participantes respondieron al cuestionario del MMPI-2 de forma voluntaria. El grupo considerado como "Sincero" contestó al cuestionario siguiendo las instrucciones sobre sinceridad y referencia a sí mismo, establecidas en el MMPI-2 (normas estándar). Al grupo Simulador se les dio la siguiente instrucción: “*Tienen ante Vds. un cuestionario al que deben responder como verdadero o falso a las preguntas que se le plantean, pero deben mostrar en todo momento una imagen favorable, o una buena imagen, de sí mismos con el objetivo de conseguir un buen beneficio*”.

De acuerdo con las sugerencias de Greene (2011) sobre el control de fiabilidad de los datos, se eliminaron todos aquellos protocolos MMPI-2 que la suma de dobles marcas y respuestas en blanco fueran ≥ 30 ; $L \geq 80T$; $F/Fb \geq 100T$; $K \geq 70T$ así como los que alcanzaron un valor en las variables de inconsistencia $VRIN/TRIN \geq 12$ puntuaciones directas (equivalentes a $>80T$). Todos los ítems de las diversas escalas fueron obtenidos de los 567 que componen el MMPI-2 en su reestructuración del MMPI-2-RF.

Resultados

Se realizaron los cálculos referentes a la fiabilidad (α de Cronbach) aportada por cada escala, sus diferencias de medias y tamaño del efecto de Cohen (d) entre los grupos y la precisión diagnóstica a través de los índices principales de la curva ROC.

La escala que ha presentado mayor fiabilidad (α de Cronbach), es la Superlativa ($S-r = .857$; con 26 ítems), seguida por la de Otro engaño ($ODEcp-r = .827$; con 17 ítems) y por la escala de Deseabilidad social de Edwards ($ESD-r = .809$; con 24 ítems). Las escalas $L-r$ (14 ítems) y $K-r$ (14 ítems) muestran la fiabilidad más baja del grupo (0,617 y 0,730 respectivamente).

La significación estadística entre sus diferencias de medias entre los grupos Sincero y Simulador, avalada por el tamaño del efecto (" d " de Cohen), se muestran muy elevadas y significativamente diferentes entre los grupos en todas y cada una de las variables. La Tabla 1 muestra que la mayor diferencia la obtiene la $ODEcp-r$ ($d = 2.16$) y la menor la $ESD-r$ ($d = 1.21$).

Tabla 1
Diferencias de Medias y " d " de Cohen

Variables	Sincero/Simulador	N	Media	Desv. típ.	t	"d" de Cohen
ODEcp-r	Sincero	309	7,96	2,722	-26.130	-2.16
	Simulador	278	13,94	2,807		
L_r	Sincero	309	5,28	1,634	-24.303	-2.02
	Simulador	278	9,03	2,054		
K_r	Sincero	309	6,34	2,552	-17.908	-1.48
	Simulador	278	9,96	2,347		
Wsd-r	Sincero	309	6,35	2,122	-24.329	-2.01
	Simulador	278	10,74	2,238		
S-r	Sincero	309	10,92	4,229	-18.879	-1.57
	Simulador	278	17,92	4,699		
ESD-r	Sincero	309	16,20	3,804	-14.741	-1.21

Nota: *ODEcp-r*: Escala de Otro engaño (Nichols y Greene, 1991); *L-r*: Escala de Virtudes inusuales; *K-r*: Escala de Validez de ajuste; *Wsd-r*: Escala de Deseabilidad social de Wiggins (1959); *S-r*: Escala Superlativa de Butcher y Han (1995); *ESD-r*: Escala de Deseabilidad social de Edwards (1953).

Tabla 2

Precisión diagnóstica a través de la curva ROC. Escalas MMPI-2-RF de Minimización de síntomas

Escalas	Precisión diagnóstica						Punto de corte
	AUC	SE	95% CI	Sign	Se	Sp	
ODecp-r	,922	0,0114	0,897 to 0,942	<0,0001	82,37	90,61	>11
L-r	,911	0,0124	0,885-0,933	<0,0001	79,14	92,23	>7
K-r	,848	0,0162	0,816-0,876	<0,0001	79,14	79,29	>8
Wsd-r	,911	0,0123	0,885-0,933	<0,0001	85,61	85,44	>8
S-r	,859	0,0154	0,828-0,886	<0,0001	83,09	73,14	>13
ESD-r	,820	0,0174	0,787-0,851	<0,0001	81,29	70,55	>18

Note. AUC: Area bajo la curva; SE = Standard Error; CI= Intervalo de confianza; Se = Sensibilidad; Sp = Especificidad; ODecp= Escala de Otro engaño (Nichols y Greene, 1991); L-r=Virtudes inusuales; K-r = Validez de ajuste; Wsd-r = Escala de Deseabilidad social (Wiggins); S-r = Escala Superlativa (Butcher y Han); ESD-r = Escala de Deseabilidad social (Edwards).

Los resultados sobre la precisión diagnóstica a través de la curva ROC, presentados en la Tabla 2, muestra unos índices elevados en todas las escalas evaluadas, teniendo el mayor índice AUC la de Otro engaño (ODecp-r) y el de menor precisión es la Deseabilidad Social de Edward (ESD-r).

Discusión/Conclusiones

En este estudio se han añadido cuatro escalas detectoras de minimización de síntomas existentes en la literatura científica diferentes a las propuestas en la adaptación española del MMPI-2-RF (Santamaría, 2009) que podrían aportar mayor precisión diagnóstica a la detección de la minusvaloración de síntomas. Los valores sobre la fiabilidad de las escalas, las diferencias de medias entre los grupos Sincero y Simulador y el tamaño del efecto de Cohen (entre 1,21 y 2,16) aportan mayor evidencia a la precisión diagnóstica de las escalas.

Dicha precisión no sólo es apreciada por el valor del AUC (*Area Under Curve*) sino también por los índices de Sensibilidad y Especificidad. La sensibilidad se considera como la probabilidad de clasificar correctamente a un individuo cuya situación real sea definido como “positivo” (Simulador) y la Especificidad como la probabilidad de clasificar correctamente a un individuo cuya situación real sea definido como “negativo” (Sincero). Así, por ejemplo (Tabla 1), la sensibilidad de la escala ODecp-r puede detectar a una persona simuladora en el 82,7% de los casos (con el riesgo del 17,3% de detectar falsos positivos) y al 96,7% de los casos como sincera (con el riesgo de considerar como falsos negativos al 9,39%). Jiménez, Sánchez y Ampudia (2008) ya encontraron en esta variable (ODecp) con el MMPI-2, una sensibilidad del 80,58% y una especificidad del 93,16%, muy elevada. También, Jiménez et al. (2009), se preguntaron por la existencia de una escala que evaluara la deseabilidad social en el MMPI-2 entre las escalas de Wiggins y Edwards, inclinándose por Wiggins por disponer de mejor precisión mediante la curva ROC (AUC = 0,94, sensibilidad = 84,98%, y especificidad = 91,4%, con (punto de corte >18).

El estudio realizado por Sellbom y Bagby (2010), analizando las escalas L-r y K-r del MMPI-2-RF, con un grupo de 94 estudiantes que intentaba minusvalorar su sintomatología, encontró que dichas escalas son apropiadas para detectar la simulación.

Bagby, Rogers, Buis y Kalembe, (1994) ya encontraron en su investigación, con el MMPI-2, que la Escala de Fingimiento positivo (*Mp*, posteriormente reestructurada por Nichols y Greene (1991) - como *ODecp*), es más efectiva que las Escalas de Validez *L* y *K* para identificar a los que mostraban una imagen positiva de sí mismos.

Los resultados de este estudio son coincidentes con la mayor parte de los estudios que utilizan el análisis de la curva ROC, sin embargo este estudio ha aportado algo más: 1) que la escala de Otro engaño (*ODecp-r*) ha presentado la mayor precisión, superior, incluso, a la de Virtudes inusuales (*L-r*); y 2) también han resultado tener mayor precisión diagnóstica las escalas de Deseabilidad social de Wiggins (*Wsd-r*) y la Psiquiátrica (*S-r*) que la de Validez de ajuste (*K-r*).

Una de las limitaciones importantes que tiene este estudio es la de no poder investigar con los verdaderos simuladores (muy difícil detectar a personas que colaboren en estos estudios, porque, precisamente, son simuladores) y tener que disponer de simuladores instruidos.

Referencias

- Bagby, R. M., Rogers, R., Buis, T., y Kalembe, V. (1994). Malingered and defensive response styles on the MMPI-2: An examination of validity scales. *Assessment*, 1, 31-38.
- Ben-Porath, Y. S., y Tellegen, A. (2011). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2 Restructured Form): Manual for administration, scoring, and interpretation*. Minneapolis, MN: University of Minnesota Press. Trabajo original 2008.
- Butcher, J. N., y Han, K. (1995). Development of an MMPI-2 scale to assess the presentation of self in a superlative manner: The S Scale. In J. N. Butcher y Ch. Spielberger (Eds), *Advances in personality assessment* (pp. 25-50). Hillsdale, NJ, US: Lawrence Erlbaum Associates.
- Butcher, J. N., Hass, G. A., Greene, R. L., y Nelson, L. D. (2015). *Using the MMPI-2 in forensic assessment*. Washington, DC: American Psychological Association.
- Edwards, A. L. (1953). *Manual for the Edwards personal preference schedule*. New York: Psychological Corporations.
- Greene, R. L. (2011). *The MMPI-2 / MMPI-2-RF. An interpretative manual* (3rd Ed.). Boston, MA: Allyn & Bacon.
- Jiménez, F., Sánchez, G. y Ampudia, A. (2008). Utilidad de la escala *ODecp* de Nichols & Green (1991) en el MMPI-2. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 26, 75-91.
- Jiménez, F., Sánchez, G., y Tobón, C. (2009). A social desirability scale for the MMPI-2. Which of the two: Wiggins (*WSD-R*) or Edwards (*ESD*)? *The European Journal of Psychology Applied to Legal Context*, 1, 147-163.
- Jiménez-Gómez, F., Sánchez-Crespo, G., y Ampudia-Rueda, A. (2013). Is there a social desirability scale in the MMPI-2-RF? *Clínica y Salud*, 24, 161-176.
- Nichols, D. S., y Greene, F. L. (1991). *New measures for dissimulation on the MMPI/MMPI-2*. Trabajo presentado en el 26th Annual Symposium on Recent Developments in the Use of the MMPI. St. Petersburg Beach, Fl.
- Nicholson, R. A., Mouton, G. J., Bagby, R. M., y Buis, T. (1997). Utility of MMPI-2 indicators of response distortion: Receiver operating characteristic analysis. *Psychological Assessment*, 9, 471-479.
- Pelegrina, M., Ruiz-Soler, M., López, E., y Wallace, A. (2000). Análisis de variables mediante curvas ROC y modelos categóricos. *Psicothema*, 12, 427-430.

- Rogers, R. (2008). *Clinical assessment of malingering and deception* (3rd. Ed.). New York, NY: The Guilford Press.
- Santamaría, P. (2009). *Adaptación española del MMPI-2-RF*. Madrid: TEA Ediciones.
- Sellbom, M., y Bagby, R. M. (2008). Validity of the MMPI-2-RF (Rstructured Form) L-r and K-r Scales in detecting underreporting in clinical and nonclinical samples. *Psychological Assessment*, 20, 370-376.
- Sellbom, M., y Bagby, R. M. (2010). Detection of overreported psychopathology with the MMPI-2 RF form Validity Scales. *Psychological Assessment*, 22, 757-767. doi: 10.1037/a0020825.
- Wiggins, J. S. (1959). Interrelations among the MMPI measures of dissimulation under standard and social desirability instructions. *Journal of Consulting Psychology*, 23, 419-427.