

METAANÁLISISRecibido: 17 de octubre de 2016
Aceptado: 24 de noviembre de 2016
Publicado: 25 de noviembre de 2016**META-ANÁLISIS DE GENERALIZACIÓN DE LA FIABILIDAD DEL INVENTARIO DE OBSESIONES DE LEYTON VERSIÓN PARA NIÑOS AUTO-APLICADA (*)****Julio Sánchez-Meca (1), Isaac Alacid-de-Pascual (1), José Antonio López-Pina (1) y Juan de la Cruz Sánchez-Jiménez (2).**

(1) Facultad de Psicología. Campus de Espinardo. Universidad de Murcia. Murcia. España.

(2) Hospital Reina Sofía. Murcia. España

(*) Esta investigación fue financiada por el Ministerio de Economía y Competitividad, Convocatoria 2016 de I+D (Proyecto N° PSI2016-77676-P), y los Fondos FEDER, siendo el primer firmante el investigador responsable del proyecto.

RESUMEN

Fundamentos: El Inventario de Obsesiones de Leyton Versión para niños Auto-Aplicada (LOI-CV-SF) es una escala ampliamente utilizada para la detección de síntomas obsesivo-compulsivos y su interferencia en la vida diaria en niños y adolescentes de 10 a 18 años. El objetivo fue estimar la fiabilidad de las puntuaciones del test y cómo varía en función de las características de los estudios.

Método: Se realizó un metaanálisis. Una búsqueda exhaustiva permitió seleccionar 13 estudios que reportaron alguna estimación de la fiabilidad del test y, mediante la fórmula KR-21, se pudo incrementar esta base de datos hasta 43 estimaciones de la fiabilidad por consistencia interna. Tomando estos coeficientes se hizo una estimación media de la fiabilidad y un análisis de las variables que podían estar afectando a la heterogeneidad de los coeficientes de fiabilidad.

Resultados: En promedio, la fiabilidad por consistencia interna de las puntuaciones del test para la escala de síntomas fue de 0,79 (IC al 95%: 0,76 y 0,82), con valores mínimo y máximo 0,52 y 0,97, respectivamente. Se observó una alta heterogeneidad ($I^2=96\%$). El modelo predictivo identificó la desviación típica de las puntuaciones del test como la variable más influyente.

Conclusiones: Las puntuaciones del LOI-CV-SF presentan una fiabilidad media satisfactoria para propósitos de investigación, pero no para la práctica clínica. Los análisis de variables moderadoras revelaron, principalmente, que la desviación típica de las puntuaciones del test es la variable más relevante en la fiabilidad, presentando una correlación positiva con esta.

Palabras clave: Inventario de personalidad, Trastorno obsesivo-compulsivo, Meta-análisis, Fiabilidad de los resultados, Salud mental, Estudios de validación, Niños, Adolescentes.

Correspondencia
Julio Sánchez-Meca
Dpto Psicología Básica y Metodología
Facultad de Psicología
Campus de Espinardo
Universidad de Murcia
30100 Murcia
jsmeca@um.es
Tfno: 868884114
www.um.es/metaanalysis

ABSTRACT**A Reliability Generalization Meta-analysis of the Leyton Obsessional Inventory Child Version Survey Form**

Background: The Leyton Obsessional Inventory Child Version Survey Form (LOI-CV-SF) is a scale created in order to detect obsessive-compulsive symptoms and their effects on the daily lives of children and adolescents from ages 10 to 18 and for the screening of non-clinic infant population. With the purpose of estimating the reliability of the test scores, and how it varies in relation to the characteristics of the studies, a meta-analysis of reliability generalization was carried out.

Method: A meta-analysis was accomplished. An exhaustive research allowed to select 13 studies that reported some reliability estimate of the test scores and, through the KR-21 formula, this number was increased to 43 estimations of internal consistency reliability. Taking these coefficients, an estimation of the average reliability and analyses of the variables that could be affecting the heterogeneity of reliability coefficients was accomplished.

Results: On average, the reliability by internal consistency of the symptom scale scores was 0.79 (95%CI: 0.76 and 0.82), with minimum and maximum values of 0.52 and 0.97, respectively. A large heterogeneity was found ($I^2 = 96\%$). The predictive model identified the standard deviation of tests scores as the most relevant variable. The analysis of moderator variables revealed that, mainly, the standard deviation of the test scores is the most statistically related to the reliability, presenting a positive relationship with it.

Conclusions: The results of the LOI-CV-SF showed a satisfactory average reliability for research purposes, but not for clinical practice purposes.

Key words: Personality inventory, Obsessive-compulsive disorder, Meta-analysis. Reliability of results, Validation Studies as Topic, Child, Adolescents, Mental health.

Cita sugerida: Sánchez-Meca J, Alacid-de-Pascual I, López-Pina JA, de la Cruz Sánchez-Jiménez J. Meta-análisis de generalización de la fiabilidad del inventario de obsesiones de Leyton versión para niños auto-aplicada. Rev Esp Salud Pública. 2016; Vol. 90: 25 de noviembre: e1-e14.

INTRODUCCIÓN

El trastorno obsesivo-compulsivo (TOC) es uno de los trastornos mentales más comunes, con una tasa de prevalencia en población infantil en torno al 1-4%⁽¹⁾. En España se ha estimado la prevalencia del TOC en niños en un 1,8%, siendo 4,7% la prevalencia estimada de sintomatología de TOC y 5,5% la prevalencia subclínica⁽²⁾. Una de las pruebas más utilizadas para el cribado de los síntomas obsesivo-compulsivos en la infancia es el *Inventario de Obsesiones de Leyton* versión para niños auto-aplicada (LOI-CV-SF), desarrollada por Berg y colaboradores en 1988⁽³⁾. Este inventario deriva de otras versiones de la escala de Leyton^(4,5), de las que se diferencia por ser auto-aplicada, reducir el número de ítems a 20 e ir dirigida de forma exclusiva a población infantil. Las principales ventajas que los autores pretendieron obtener con esta versión fue reducir el tiempo de administración y la implicación del evaluador en la valoración de los temores, pensamientos y comportamientos del niño.

La escala, compuesta por 20 ítems, tiene dos partes. En una de ellas se da una respuesta dicotómica “sí/no” con el objetivo de comprobar la presencia de síntomas obsesivo-compulsivos. En la otra, se valora la interferencia de los síntomas en caso de que estén presentes, de 0 (no interferencia) a 3 (mucho interferencia). Como resultado de la aplicación, se dan dos puntuaciones: una de la sintomatología obsesivo-compulsiva presente (resultado de la suma de los “síes”) y otra de la interferencia de esta sintomatología en la vida del niño/adolescente (suma de las interferencias independientemente de las respuestas “sí”). Por sus características, uno de sus usos ha sido el estudio epidemiológico de población infantil general. Desde su desarrollo el LOI-CV-SF ha sido adaptado a distintos idiomas como el Polaco⁽⁶⁾, Alemán⁽⁷⁾ y Danés⁽⁸⁾ y Español⁽⁹⁾.

Diversas investigaciones analizaron las propiedades psicométricas del LOI-CV-SF^(3,9,10,11). En su estudio original, Berg y co-

laboradores⁽³⁾ encontraron cuatro factores, a los que denominaron preocupaciones obsesivas, contaminación, números/suerte y escuela. En estudios posteriores, como el realizado en España por Canals y colaboradores en 2012⁽⁹⁾, el análisis factorial dio como resultado tres factores que se denominaron orden/comprobación/contaminación, pensamientos obsesivos y supersticiones y compulsiones mentales. En un reciente estudio realizado sobre las propiedades del LOI-CV-SF, Sun y colaboradores en 2014⁽¹¹⁾ encontraron una distribución de los ítems en cuatro factores: pensamientos obsesivos, limpieza y pulcritud, números y suerte, y repetitividad y comprobación.

Los estudios psicométricos del LOI-CV-SF han mostrado, en general, buena fiabilidad para las escalas globales de síntomas y de interferencia (0,79-0,87), pero la fiabilidad de las subescalas es pobre (0,48-0,66). Además, se observa una gran variabilidad en las estimaciones de la fiabilidad a lo largo de sus aplicaciones en muestras procedentes de diferentes aplicaciones del test.

Enfoque de generalización de la fiabilidad

La teoría psicométrica establece que la fiabilidad no es una propiedad estática de un test, por lo que afirmaciones del tipo “el test tiene una fiabilidad de 0,80” son incorrectas, ya que la fiabilidad no es una propiedad inherente al test, sino de las puntuaciones del test obtenidas en una aplicación concreta del mismo^(12,13,14). La fiabilidad de las puntuaciones de un test puede variar dependiendo de la composición y las características de las muestras y del contexto de aplicación.

Como varía en cada administración de un test, los investigadores deben obtener la fiabilidad con los propios datos de la muestra. Sin embargo, es muy frecuente que los investigadores no informen de la fiabilidad del instrumento obtenida con sus propios datos, sino que la induzcan desde alguna investigación anterior (e.g., la fiabilidad obtenida en el estudio de validación original del test).

Esta desaconsejable práctica fue denominada por Vacha-Haase, Kogan y Thompson ‘inducción de la fiabilidad’ (*reliability induction*)⁽¹⁵⁾, y se basa en la concepción errónea de que es una propiedad inmutable del test. Desafortunadamente, esta práctica de inducir la fiabilidad a partir de aplicaciones previas del test está muy extendida, hasta el punto de que las tasas de inducción de la fiabilidad se sitúan en torno al 78,6%, según los resultados del estudio realizado por Sánchez-Meca et al. en 2015⁽¹⁶⁾, que obtuvo datos de 100 meta-análisis de ‘generalización de la fiabilidad’ y aglutinó más de 40.000 estudios empíricos.

Ante esta situación, desarrollar estudios que den información sobre cómo varían los coeficientes de fiabilidad obtenidos en diferentes aplicaciones de un test constituye una labor fundamental para el correcto progreso del conocimiento científico en cualquier disciplina de las Ciencias de la Salud. El mejor modo de realizar esta tarea es mediante los estudios meta-analíticos de ‘generalización de la fiabilidad’ (*reliability generalization*), término acuñado por Vacha-Haase en 1998⁽¹⁷⁾. Los estudios de generalización de la fiabilidad (GF) permiten obtener una estimación media de la fiabilidad de las puntuaciones del test, determinar si los coeficientes de fiabilidad obtenidos en las múltiples aplicaciones del test son heterogéneos entre sí y, en su caso, examinar qué características del test, de los estudios y de los participantes pueden dar cuenta de dicha heterogeneidad^(18,19,20). Desde la creación de este tipo de estudios⁽¹⁷⁾ se han publicado más de 100 estudios de GF sobre diferentes pruebas psicológicas, en particular, y sanitarias, en general. Por ejemplo, se han llevado a cabo estudios GF sobre el Inventario de Ansiedad Estado-Rasgo de Spielberger⁽²¹⁾, el Inventario de Depresión de Beck⁽²²⁾, el Inventario de Personalidad de Eysenck⁽²³⁾, el Inventario de Burnout de Maslach⁽²⁴⁾ y la Escala de Hamilton para la Depresión⁽²⁵⁾.

El LOI-CV-SF es una escala que mide la sintomatología obsesivo-compulsiva y su ni-

vel de interferencia y ha sido frecuentemente empleada en estudios como instrumento de cribado, especialmente en estudios epidemiológicos. Aunque se han desarrollado estudios sobre las propiedades psicométricas de la escala, no tenemos conocimiento de que se haya realizado un meta-análisis de GF de las puntuaciones del LOI-CV-SF.

Características tales como la media y la desviación típica de las puntuaciones de la escala, la edad, la población objetivo de los participantes (clínica *vs* no clínica) y la versión de la prueba (original *vs* adaptada) pueden afectar a la fiabilidad de las puntuaciones del test.

En consecuencia, los objetivos de esta investigación fueron estimar la fiabilidad media de las puntuaciones de la escala y así obtener un punto de referencia aproximado de su fiabilidad general y comprobar qué características de los estudios y de las muestras pueden estar provocando la heterogeneidad en sus coeficientes de fiabilidad.

MATERIAL Y MÉTODO

La realización de este meta-análisis de GF se llevó a cabo siguiendo las recomendaciones de la guía PRISMA⁽²⁶⁾, si bien las características peculiares de este tipo de meta-análisis hacen que ciertos ítems de esta guía no sean aplicables.

Búsqueda bibliográfica. Para ser incluidos en el meta-análisis de GF los estudios tenían que cumplir cuatro criterios: (a) ser un estudio empírico que hubiera aplicado la escala LOI-CV-SF a una o varias muestras de participantes, (b) estar publicado entre 1988 y febrero de 2016, (c) aportar al menos un coeficiente de fiabilidad alfa de Cronbach de la escala de síntomas calculado con los propios datos de la muestra, o facilitar la media y la desviación típica de los resultados de la escala de síntomas, de forma que fuera posible hacer una estimación de la consistencia interna mediante la aplicación de la fórmula KR-21, y (d) estar redactado en inglés o español.

Para la localización de los estudios se consultaron las bases electrónicas PsycInfo, Medline, Pudmed, Web of Science y Scopus, además de Google Académico como complemento de búsqueda de las demás bases de datos. Se utilizó la palabra clave '*Leyton obsessional*', pudiendo aparecer en cualquier parte del documento. También se consultaron revistas especializadas en TOC, monografías y libros, y se estableció contacto con investigadores expertos en la materia con el fin de localizar estudios no publicados que hubieran aplicado el LOI-CV-SF.

Del total de referencias localizadas (más de 1.000), se seleccionaron 418 que habían aplicado alguna de las versiones del LOI. El resto se descartaron, bien por ser textos escritos en lenguas distintas del inglés o el español o por tratarse de estudios teóricos. De las 418 referencias preseleccionadas, se excluyeron 334 por aplicar versiones diferentes de la LOI-CV-SF. De los 84 trabajos que habían aplicado la LOI-CV-SF, se descartaron 48 por no aportar el coeficiente alfa ni los datos suficientes para estimarlo. De los 84 artículos que aplicaron la LOI-CV-SF solo 13 aportaron alguna estimación de la fiabilidad con los propios datos de la muestra, lo que supuso una elevada tasa de inducción de la fiabilidad para esta escala del 84,5%. Afortunadamente, 23 estudios adicionales pudieron incluirse en el meta-análisis al aportar la media y la desviación típica de las puntuaciones del test en la escala de síntomas, lo que resultó en un total de 36 estudios incluidos. En la *figura 1* se presenta un diagrama de flujo que describe el proceso de selección de los estudios. La lista de referencias de estos estudios puede solicitarse a los autores.

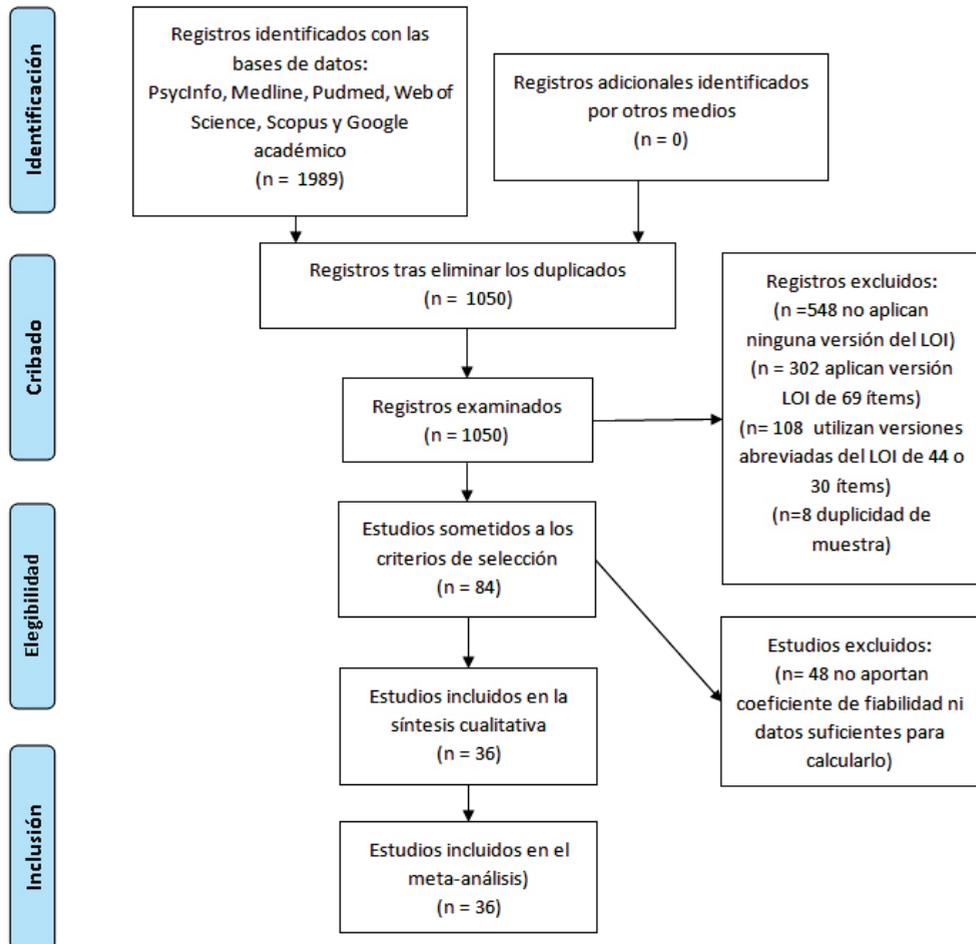
Codificación de las características de los estudios. De los 36 estudios seleccionados, algunos de ellos permitieron obtener coeficientes de fiabilidad para varias muestras, de forma que el número total de unidades de análisis (muestras) de nuestro meta-análisis fue de 43. A cada una de estas unidades de análisis se le aplicó un protocolo de registro de sus características, que incluyó: (1) tipo de estudio (psi-

cométrico vs sustantivo); (2) caso de ser psicométrico, en qué test se centró (en la escala LOI-CV-SF o en otra); (3-5) tamaño muestral, media y desviación típica de las puntuaciones del test en la escala de síntomas; (6-7) media y desviación típica de la edad de los participantes en la muestra (en años); (8) porcentaje de mujeres en la muestra; (9) porcentaje de caucásicos de la muestra; (10-11) país y continente donde se realizó el estudio; (12) año de publicación; (13) formato de aplicación (auto-aplicado vs. entrevista clínica); (14) versión del test (original en inglés vs adaptación); (15) tipo de adaptación del test (adaptación estandarizada vs libre); (16-17) formación y departamento del investigador principal (psicólogo, psiquiatra, pediatra); (18) tipo de población (no clínica, análoga, clínica); (19) instrumento diagnóstico (versión DSM, versión CIE-10), y (20-21) media y desviación típica de la historia del trastorno (en meses) en caso de ser una muestra procedente de población clínica.

Para analizar la fiabilidad del proceso de codificación de las características de los estudios, dos codificadores extrajeron de forma independiente los datos de una muestra aleatoria del 42% de los estudios meta-analizados. En general, el índice de acuerdo inter-jueces fue satisfactorio, con coeficientes kappa de Cohen que oscilaron entre 0,902 y 1 para las variables cualitativas y correlaciones intraclass entre 0,736 y 0,902 para las variables continuas.

Análisis estadístico. Se incluyeron estudios que aportaron al menos un coeficiente alfa con los propios datos o bien aportaron la media y la desviación típica de la escala de síntomas del test, de forma que fue posible obtener una estimación del coeficiente alfa mediante la aplicación de la fórmula KR-21. Con objeto de normalizar su distribución, previa la realización de los análisis estadísticos, los coeficientes alfa se transformaron mediante la fórmula propuesta por Bonett⁽²⁷⁾. Una vez realizados los análisis estadísticos, los resultados se devolvieron a la métrica del coeficiente alfa para facilitar su interpretación.

Figura 1
Diagrama de flujo del proceso de selección de los estudios



Se asumió el modelo de efectos aleatorios para los cálculos estadísticos, de forma que los coeficientes alfa se ponderaron en función de la inversa de su varianza. La heterogeneidad exhibida por las estimaciones de la fiabilidad se evaluó mediante el estadístico Q de Cochran y el índice I^2 . La influencia de variables moderadoras se examinó mediante la aplicación de análisis de meta-regresión para las variables continuas y ANOVAs ponderados para las cualitativas. Los ANOVAs se resolvieron mediante el contraste de los coeficientes medios con

el estadístico Q_B , mientras que los análisis de meta-regresión se resolvieron con el estadístico Z . La especificación del modelo se contrastó con los estadísticos Q_w y Q_E para los ANOVAs y las meta-regresiones, respectivamente. Se estimó la proporción de varianza explicada por la variable moderadora mediante el índice R^2 calculado a partir de las varianzas inter-estudios residual y total^(28,29,30). Finalmente, con objeto de identificar el conjunto de variables que mejor explican la variabilidad de los coeficientes de fiabilidad, se aplicó un modelo de meta-re-

gresión múltiple tomando como predictores las variables moderadoras que hubieran exhibido una relación estadísticamente significativa con los coeficientes de fiabilidad. Los análisis estadísticos se realizaron con el programa metafor dentro del entorno R⁽³¹⁾.

RESULTADOS

Estimación de la fiabilidad media. Del total de estudios que aplicaron la escala LOI-CV-SF, solo 8 aportaron algún coeficiente de fiabilidad alfa de Cronbach para las escalas totales de síntomas y de interferencia. Para la primera se obtuvo una fiabilidad media de 0,80 (IC al 95%: 0,78-0,82) y para la escala de interferencia la fiabilidad media fue de 0,85 (IC al 95%: 0,80-0,89). Dado que la escala de síntomas estaba formada por ítems dicotómicos, fue posible ampliar el número de coeficientes alfa hasta un total de 43 estimaciones, mediante la aplicación de la fórmula KR-21. Estos 43 coeficien-

tes alfa constituyeron la base de datos sobre la que realizamos los cálculos estadísticos. En el anexo 1 figuran las principales características de los estudios meta-analizados. Las estimaciones de fiabilidad variaron entre los valores 0,52 y 0,97, siendo la fiabilidad media igual a 0,79 (IC al 95%: 0,76 y 0,82).

Análisis de heterogeneidad

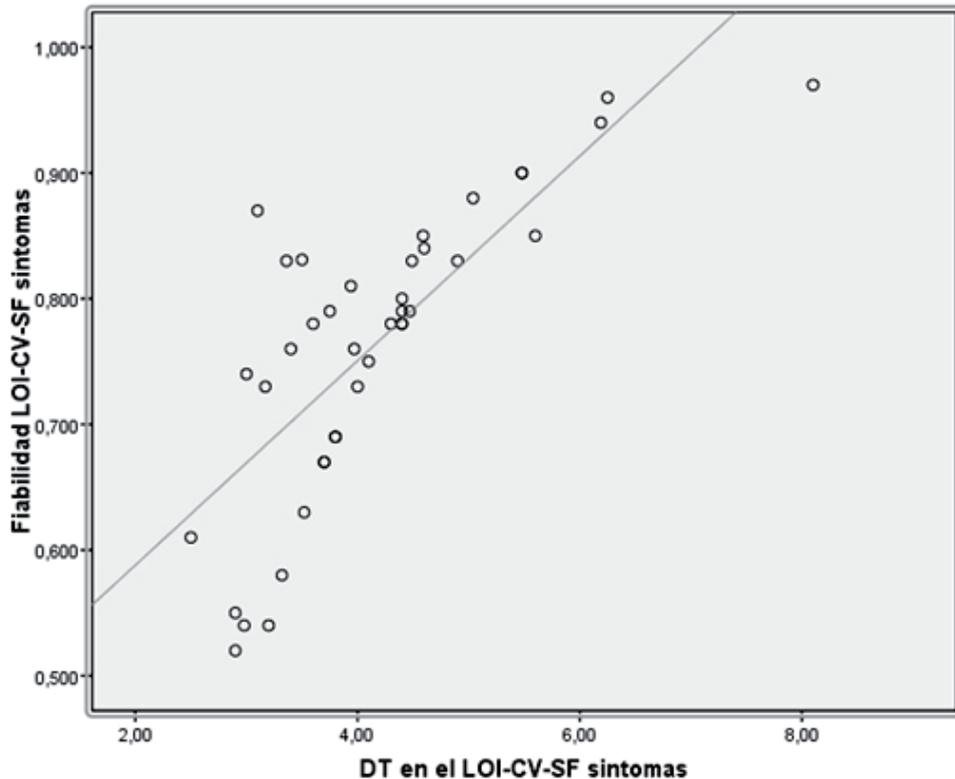
Los 43 coeficientes de fiabilidad exhibieron una elevada heterogeneidad, según se dedujo de la alta significación estadística alcanzada por el estadístico *Q* de Cochran, $Q(42)=1.055,44$, $p<0,001$, y el elevado índice *I*² obtenido, que ascendió al 96%. Estos resultados apuntaron hacia la existencia de una gran heterogeneidad entre los coeficientes de fiabilidad obtenidos en las diversas aplicaciones de la escala, por lo que se hizo necesario realizar análisis de las variables que pudieran estar asociadas a esta heterogeneidad.

Tabla 1
Resultados de los análisis de meta-regresión simples para las variables moderadoras continuas

Variable moderadora	<i>k</i>	Mínimo	Máximo	Media	<i>b_j</i>	<i>Z</i>	<i>p</i>	<i>Q_E</i>	<i>R</i> ²
Media LOI-síntomas	40	1,8	16,14	8,12	0,095	0,3737	0,70	1014,59*	0
DT LOI síntomas	40	2,5	8,1	4,15	-0,478	-7,693	<0,001	682,01*	0,30
N LOI síntomas	43	7	4551	559,6	0,001	1,28	0,200	976,41*	0
Media LOI interferencia	33	1,5	24,85	10,25	-0,041	-3,167	0,001	623,56*	0
DT LOI interferencia	33	2,3	14,10	7,8	-0,084	-3,963	<0,001	600,86*	0,07
Edad media	32	10,23	16,7	13,86	0,073	1,411	0,158	450,74*	0
DT de la edad	30	0,27	3,10	1,75	-0,312	-3,554	0,001	125,41*	0,16
% de mujeres	37	0	100	59,01	0,0005	0,1920	0,848	823,60*	0
% caucásicos	7	78,8	100	89,46	-,0068	-0,4227	0,672	18,97*	0
Historia del trastorno	7	7,5	99,6	22,32	0,0008	0,3527	0,724	4,304*	0
DT historia trastorno	7	3,10	37,3	10,39	0,0013	0,1984	0,842	4,389*	0
Año de publicación	43	1988	2014	2003	0,0051	0,5528	0,580	1020,95*	0

k=número de estudios (o coeficientes de fiabilidad). *b_j*: coeficiente de regresión de la variable moderadora. El signo de los coeficientes *b_j* debe interpretarse a la inversa, debido a que las meta-regresiones se realizaron con las transformaciones de Bonett de los coeficientes de fiabilidad, y esta transformación invierte el sentido de la variable. En consecuencia, en esta tabla un coeficiente de regresión negativo debe interpretarse como reflejando una relación positiva entre los coeficientes de fiabilidad y la variable moderadora; y vice versa. *Z*: estadístico para contrastar la significación estadística de la variable moderadora. *Q_E*: estadístico para comprobar si el modelo está bien especificado. $p<0,05$: nivel de significación asociado a la prueba *Z*. *R*²: proporción de la varianza explicada por la variable moderadora.

Figura 2
Gráfico de dispersión de los coeficientes de fiabilidad en función de la desviación típica de las puntuaciones del test en la escala de síntomas



Búsqueda de variables moderadoras

Para explorar las características de los estudios y de las muestras relacionadas con la fiabilidad de la escala LOI-CV-SF, los 43 coeficientes alfa constituyeron la variable dependiente en los análisis estadísticos, mientras que las variables moderadoras previamente codificadas actuaron como los potenciales predictores de la heterogeneidad entre los coeficientes de fiabilidad.

La **tabla 1** presenta los resultados de las meta-regresiones simples, asumiendo un modelo de efectos mixtos, aplicadas a cada variable moderadora continua sobre los coeficientes de fiabilidad. De todas las analizadas,

como era de esperar, se observó una asociación estadísticamente significativa entre el coeficiente alfa y la desviación típica de las puntuaciones del test en la escala de síntomas ($Z=7,693$, $p<0,001$), con un porcentaje de varianza explicada, nada despreciable, del 30% ($R^2=0,30$).

La **figura 2** presenta un gráfico de dispersión que ilustra la relación positiva encontrada entre la variabilidad de las puntuaciones del test y los coeficientes de fiabilidad. También se obtuvo una relación positiva y estadísticamente significativa de los coeficientes de fiabilidad con la media de las puntuaciones en la escala de interferencia ($Z=3,161$,

Tabla 2
Resultados de los ANOVAs aplicados a cada variable moderadora cualitativa

	Variable moderadora	k	Media	IC95%	Q_W	Q_B	p	R ²
Tipo de estudio	Psicométrico	6	0,79	0,73-0,84	633,66*	0,0034	0,953	0,02
	Sustantivo	37	0,78	0,76-0,81				
Foco del estudio psicom	LOI-CV-SF	5	0,78	0,75-0,81	68,85*	5,8	0,016	0,05
	Otras escalas	1	0,90	0,81-0,95				
Continente del estudio	Europa	30	0,78	0,75-0,82	894,91*	2,99	0,224	0
	América del Norte	10	0,83	0,77-0,87				
	Oceanía	4	0,73	0,56-0,84				
Versión del test	Original	17	0,84	0,81-0,87	615,24*	19,01	0,001	0
	Español	15	0,72	0,65-0,78				
	Alemán	3	0,74	0,59-0,84				
	Polaco	2	0,72	0,51-0,81				
	Danés	3	0,86	0,78-0,92				
	Otros	3	0,77	0,65-0,84				
Tipo de adaptación	Original	16	0,84	0,81-0,87	641,64*	14,55	0,001	0,20
	Estandarizada	4	0,71	0,62-0,78				
	Adaptación libre	20	0,76	0,72-0,80				
Formación investigador:	Psicología	9	0,87	0,83-0,90	1020,94*	14,29	0,002	0
	Psiquiatría	28	0,76	0,72-0,80				
	Pediatría	3	0,81	0,66-0,89				
	Otros	3	0,73	0,56-0,84				
Departamento investig.:	Psicología	7	0,88	0,83-0,92	1032,57*	12,27	0,015	0
	Psiquiatría	21	0,76	0,71-0,80				
	Pediatría	3	0,81	0,66-0,83				
	Otros	6	0,80	0,72-0,86				
	1+2	6	0,75	0,83-0,63				
Tipo de Población:	No clínica	17	0,74	0,69-0,78	1004,13*	11,21	0,003	0,04
	Clínica	25	0,83	0,79-0,86				
	Mixta	1	0,83	0,63-0,92				
Tipo población clínica:	TOC	7	0,86	0,81-0,90	66,01*	4,25	0,235	0
	Anorexia	9	0,78	0,71-0,84				
	Tourette	4	0,85	0,77-0,91				
	Otros	5	0,82	0,73-0,89				
Forma de diagnóstico:	DSM-III-R	4	0,87	0,79-0,92	71,93*	2,17	0,537	0
	DSM-IV	11	0,81	0,75-0,85				
	DSM-IV- TR	3	0,83	0,68-0,91				
	OTROS	6	0,83	0,74-0,89				

k=número de estudios (o coeficientes de fiabilidad). Q_W =estadístico para contrastar la significación estadística de la variable moderadora. Q_B =estadístico para comprobar si el modelo está bien especificado. *p<0,05.nivel de significación asociado a la prueba Q_B . R²: proporción de la varianza explicada por la variable moderadora.

$p=0,001$) y con su desviación típica ($Z=3,963$, $p<0,001$). La desviación típica de las edades de las muestras de participantes también obtuvo una relación positiva y estadísticamente significativa con las estimaciones de la fiabilidad ($Z=3,554$, $p=0,001$). La magnitud de estas tres últimas relaciones (R^2) fue baja: 0,07 para la desviación típica en la escala de interferencia, 0,16 para la desviación típica de las edades, y nula para la media de las puntuaciones en la escala de interferencia.

Con respecto a las variables moderadoras cualitativas, la **tabla 2** muestra los resultados de los ANOVAs aplicados sobre los coeficientes de fiabilidad, asumiendo un modelo de efectos mixtos. Varias alcanzaron la significación estadística, pero con muy pequeñas proporciones de varianza explicada. La variable moderadora más relevante fue el tipo de adaptación de la escala aplicada en el estudio, que obtuvo una relación estadísticamente significativa con los coeficientes de fiabilidad, $Q_B(2)=14,55$, $p=0,001$ y un porcentaje de varianza explicada del 20%. En concreto, los estudios que aplicaron la versión original de la escala exhibieron, en promedio, una mayor fiabilidad (media=0,84) que cuando se aplicó una versión adaptada de forma estandarizada (media=0,71) o de forma libre (media=0,76). En esta misma línea, se observaron diferencias estadísticamente significativas entre los coeficientes de fiabilidad medios obtenidos entre las diversas adaptaciones de la escala a diferentes idiomas, $Q_B(5)=9,01$; $p=0,001$, siendo las versiones original (media=0,84) y danesa (media=0,86) las que presentaron las mejores estimaciones de la fiabilidad media. Las 15 muestras españolas sobre las que se obtuvieron estimaciones de la fiabilidad de las puntuaciones de la escala arrojaron una fiabilidad media de 0,72 (IC 95%: 0,65-0,78). No obstante, el porcentaje de varianza explicada por esta variable fue nulo. Cuando se agruparon los estudios españoles según hubieran aplicado la escala a una muestra clínica o no clínica, los nueve coeficientes de fiabilidad obtenidos con las muestras clínicas oscilaron entre 0,67 y 0,85, con una fia-

bilidad media de 0,764 (IC 95%: 0,67-0,83), mientras que los seis coeficientes obtenidos en muestras no clínicas tomaron valores entre 0,52 y 0,79, con una fiabilidad media de 0,655 (IC 95%: 0,50-0,76). Aunque el coeficiente de fiabilidad medio para las muestras no clínicas fue inferior al de las muestras clínicas, esta diferencia no resultó ser estadísticamente significativa, $Q_B(1)=2,22$; $p=0,136$. La población de procedencia de las muestras también alcanzó la significación estadística, $Q_B(2)=110,21$; $p=0,003$, siendo más elevada la fiabilidad, en promedio, cuando la escala se aplicó a muestras clínicas (media=0,83) que a muestras comunitarias (media=0,74). No obstante, el porcentaje de varianza explicada por esta variable fue nulo. Cuando las 25 muestras clínicas se clasificaron en función del tipo de trastorno, no se observaron diferencias estadísticamente significativas, $Q_B(3)=4,25$; $p=0,235$, resultando similares los coeficientes de fiabilidad medios para las muestras de participantes con TOC (media=0,86), anorexia (media=0,78), síndrome de Tourette (media=0,85) u otros diagnósticos (media=0,82). Otras variables como la formación y/o el departamento del investigador principal del estudio (entendiendo por tal el primer firmante) mostraron también diferencias estadísticamente significativas entre los coeficientes de fiabilidad medios, pero con proporciones de varianza explicadas nulas.

Propuesta de un modelo explicativo. Aplicamos un modelo de meta-regresión múltiple tomando como predictores las variables moderadoras que presentaron una relación estadísticamente significativa con los coeficientes de fiabilidad: la desviación típica de las puntuaciones del test en la escala de síntomas, la versión del test (dicotomizada en original vs adaptada) y la población de procedencia de la muestra (clínica vs no clínica). El modelo de meta-regresión múltiple alcanzó un resultado estadísticamente significativo, $Q_R(3)=56,58$, $p<0,001$, con un 19,3% de varianza explicada, pero solo la desviación típica de las puntuaciones del test presentó una relación estadísticamente significativa una vez contro-

lado el influjo de los otros dos predictores del modelo ($Z=6,13$; $p<0,001$), mientras que ni la versión del test ($Z=1,10$, $p=0,270$) ni la población de referencia ($Z=0,91$; $p=0,362$) exhibieron una relación estadísticamente significativa con los coeficientes de fiabilidad una vez controlado el influjo de los otros predictores del modelo.

DISCUSIÓN

El propósito de esta investigación fue estimar la fiabilidad de las puntuaciones del test LOI-CV-SF mediante un meta-análisis de GF. El coeficiente de fiabilidad promedio para las escalas de síntomas y de interferencia de la LOI-CV-SF fue de 0,79 y 0,85, respectivamente. Estos valores se sitúan por encima del límite de 0,70 establecido por Nunnally y Bernstein⁽³³⁾ como fiabilidad mínima aceptable para propósitos de investigaciones exploratorias, pero la escala de síntomas exhibe una fiabilidad media que no alcanza el límite de 0,80 propuesto para considerar satisfactoria la fiabilidad de un test con propósitos de investigación en general (e.g., para investigar diferencias entre medias de grupos procedentes de diferentes poblaciones o manipulaciones experimentales). Y en cuanto al límite de 0,90 establecido para la correcta aplicación de un test en contextos clínicos de desempeño profesional, ni la escala de síntomas ni la de interferencia se acercan a este límite. En lo que respecta a la toma de decisiones clínicas, resulta razonable adoptar las recomendaciones propuestas por Cicchetti⁽³⁴⁾, según las cuales coeficientes de fiabilidad inferiores a 0,70 deben considerarse totalmente inapropiadas, entre 0,70 y 0,80 serían estimaciones moderadas de la fiabilidad, entre 0,8 y 0,9 serían buenas, y por encima de 0,90 excelentes. Así pues, se puede concluir que la escala de síntomas exhibe, en promedio, una fiabilidad moderada y la de interferencia una fiabilidad buena.

Los coeficientes de fiabilidad obtenidos en muestras españolas exhibieron una fiabilidad media satisfactoria (media=0,72) para propósitos de investigación, pero no se acercaron al límite de 0,90 mínimo requerido para la

práctica clínica. De hecho, el coeficiente de fiabilidad medio obtenido con las muestras clínicas españolas quedó ligeramente por debajo del límite 0,80 (media=0,76), indicando que la fiabilidad de las puntuaciones de esta escala para la toma de decisiones clínicas es sólo moderada.

Las puntuaciones del LOI-CV-SF arrojaron coeficientes de fiabilidad muy variables a lo largo de las 43 muestras incluidas en el meta-análisis. Varias características de los estudios presentaron una relación estadísticamente significativa con las estimaciones de consistencia interna. De todas ellas, la más importante fue la desviación típica de las puntuaciones en la escala de síntomas. Como la teoría psicométrica predice, cuanto mayor es la variabilidad de las puntuaciones de un test tanto mayor será la fiabilidad exhibida por la muestra en cuestión⁽³³⁾. En esa misma línea, se observó también una mayor fiabilidad cuanto mayor fue la desviación típica de las edades de las muestras de participantes. Del mismo modo, se observó una mayor fiabilidad en muestras clínicas que en las comunitarias. La razón de esta relación se debió a que las muestras clínicas suelen presentar mayor variabilidad entre las puntuaciones de la escala, en comparación con la exhibida por las muestras comunitarias que, en ocasiones, presentan un ‘efecto suelo’, reduciéndose la variabilidad. Así pues, el elemento fundamental que caracterizó la mayor o menor fiabilidad exhibida por las puntuaciones del test en la escala de síntomas fue la variabilidad de sus puntuaciones. De hecho, el modelo predictivo identificó a la desviación típica de las puntuaciones del test como el único predictor relevante para explicar la variabilidad de los coeficientes de fiabilidad.

Dado que la fiabilidad no es una propiedad intrínseca del test sino de las puntuaciones obtenidas del test en una aplicación concreta del mismo, es necesario realizar estudios meta-analíticos de GF que nos permitan estimar la fiabilidad media y explicar cómo varía esta en las diferentes aplicaciones de una misma escala. Aunque el LOI-CV-SF es muy utilizado

en el contexto aplicado y en estudios epidemiológicos, no son muchos los estudios que aporten algún tipo de coeficiente de fiabilidad con los datos propios de la muestra. En concreto, la tasa de inducción de la fiabilidad del LOI-CV-SF fue del 84,5% (sólo 13 de los 84 estudios aportaron algún coeficiente de fiabilidad), una tasa muy elevada, que se pudo paliar gracias al uso de la fórmula KR-21 de estimación del coeficiente de fiabilidad alfa de Cronbach⁽³²⁾.

Este meta-análisis tiene algunas limitaciones. En primer lugar, la elevada tasa de inducción de la fiabilidad exhibida por los estudios que aplicaron la LOI-CV-SF redujo el número de estudios incluíbles en el meta-análisis. En segundo lugar, no fue posible analizar la fiabilidad de las puntuaciones en las diferentes subescalas del test, debido a la ausencia de información sobre su fiabilidad en los estudios meta-analizados. En tercer lugar, la falta de información de algunas características de los estudios limitó las posibilidades de analizar su posible influjo como variables potencialmente moderadoras de los coeficientes de fiabilidad. Tal fue el caso de las medias y desviaciones típicas de las edades de las muestras de participantes así como su distribución étnica o la antigüedad en el padecimiento del trastorno en el caso de muestras clínicas.

En conclusión, podemos afirmar que las puntuaciones de la escala LOI-CV-SF presentan una fiabilidad aceptable para propósitos de investigación pero no para la toma de decisiones clínicas en un contexto de desempeño profesional. Además, la elevada tasa de inducción de la fiabilidad encontrada en los estudios que han aplicado esta escala, así como la evidenciada en otras muchas escalas e instrumentos de medida utilizados en las Ciencias de la Salud, debe poner en alerta a los investigadores y profesionales de este campo sobre la necesidad de informar de la fiabilidad de los instrumentos de medida exhibida con los datos de las propias muestras utilizadas en sus investigaciones. Los equipos editoriales de las revistas científicas deberían, así mismo, exigir el reporte de la fiabilidad calculada con los propios datos de la muestra, de todos los instrumentos de medida

utilizados en las investigaciones, y evitar así el fenómeno de la inducción de la fiabilidad. Sólo de esta forma se logrará que las investigaciones empíricas ofrezcan contribuciones relevantes para el progreso del conocimiento científico en las Ciencias de la Salud.

BIBLIOGRAFÍA

1. Noorian Z, Granero R, Ferreira E, Romero-Acosta K, Domenèch-Llaberia E. Obsessive-compulsive symptoms among Spanish adolescents: Prevalence and association with depressive and anxious symptoms. *Span J Psychol*. 2013;16:E98.
2. Canals J, Hernández-Martínez C, Cosi S, Voltas N. The epidemiology of obsessive-compulsive disorder in Spanish school children. *J Anxiety Disord*. 2012;26:746-52.
3. Berg CZ, Whitaker A, Davies M, Flament MF, Rapoport JL. The Survey Form of the Leyton Obsessional Inventory-Child Version: Norms from an epidemiological study. *J Am Acad Child Adolesc Psychiatry*. 1988;27:759-63.
4. Cooper J. The Leyton Obsessional Inventory. *Psychol Med*. 1970;1:48-64.
5. Berg CJ, Rapoport JL, Flament M. The Leyton Obsessional Inventory-Child Version. *J Am Acad Child Psychiatry*. 1986;25:84-91.
6. Bryńska A, Wolańczyk T. Epidemiology and phenomenology of obsessive-compulsive disorder in non-referred young adolescents: A Polish perspective. *Eur Child Adolesc Psychiatry*. 2005;14:319-27.
7. Buse J, August J, Bock N, Dörfel D, Rothenberger A, Roessner V. Fine motor skills and interhemispheric transfer in treatment-naive male children with Tourette syndrome. *Dev Med Child Neurol*. 2012;54:629-35.
8. Thomsen PH. Obsessive-compulsive disorder in children and adolescents. *Acta Psychiatr Scand*. 1993;88:212-7.
9. Canals JC, Hernández-Martínez C, Cosi S, Lázaro L, Toro J. The Leyton Obsessional Inventory-Child Version: Validity and reliability in Spanish non-clinical population. *Int J Clin Heal Psychol*. 2012;12:81-96.
10. Storch EA, Park JM, Lewin AB, Morgan JR, Jones AM, Murphy TK. The Leyton Obsessional Inventory-Child Version Survey Form does not demonstrate adequate psychometric properties in American youth with pediatric obsessive-compulsive disorder. *J Anxiety Disord*. 2011;25:574-8.

11. Sun J, Boschén MJ, Farrell LJ, Buys N, Li ZJ. Obsessive-compulsive symptoms in a normative Chinese sample of youth: Prevalence, symptom dimensions, and factor structure of the Leyton Obsessional Inventory - Child Version. *J Affect Disord.* 2014;164:19-27.
12. Crocker L, Algina J. Introduction to classical and modern test theory. ERIC; 1986.
13. Wilkinson, L, Task Force on Statistical Inference American Psychological Association. Statistical methods in psychology journals: Guidelines and explanations. *Am Psychol.* 1999;54:594-604.
14. Sánchez-Meca J, López-Pina JA. El enfoque meta-analítico de generalización de la fiabilidad. *Acción Psicol.* 2008;5:37-64.
15. Vacha-Haase T, Kogan LR, Thompson B. Sample compositions and variabilities in published studies versus those in test manuals: Validity of score reliability inductions. *Educ Psychol Meas.* 2000;60:509-22.
16. Sánchez-Meca J, Rubio-Aparicio M, López-Pina JA, Núñez-Núñez RM, Marín-Martínez F. El fenómeno de la inducción de la fiabilidad en Ciencias Sociales y de la Salud. Comunicación presentada en el XIV Congreso de Metodología de las Ciencias Sociales y de la Salud., Palma de Mallorca.2015 (julio).
17. Vacha-Haase T. Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educ Psychol Meas.* 1998;58:6-20.
18. Henson RK, Thompson B. Characterizing measurement error in scores across studies: Some recommendations for conducting "reliability generalization" studies. *Meas Eval Couns Dev.* 2002;35:113-26.
19. Rodríguez MC, Maeda Y. Meta-analysis of coefficient alpha. *Psychol Methods.* 2006;11:306-22.
20. Sánchez-Meca J, López-Pina JA, López-López JA. Generalización de la fiabilidad: Un enfoque meta-analítico aplicado a la fiabilidad. *Fisioterapia.* 2009;31:262-70.
21. Guillén-Riquelme A, Buela-Casal G. Metaanálisis de comparación de grupos y metaanálisis de generalización de la fiabilidad del Cuestionario State-Trait Anxiety Inventory (STAI). *Rev Esp Salud Pública.* 2014;88:101-12.
22. Yin P, Fan X. Assessing the reliability of Beck Depression Inventory scores: Reliability generalization across studies. *Educ Psychol Meas.* 2000;60:201-23.
23. Caruso JC, Witkiewitz K, Belcourt-Dittloff A, Gottlieb JD. Reliability of scores from the Eysenck Personality Questionnaire: A reliability generalization study. *Educ Psychol Meas.* 2001;61:675-89.
24. Aguayo R, Pecino CV, de la Fuente Solana EI, Fernández LML. A meta-analytic reliability generalization study of the Maslach Burnout Inventory. *Int J Clin Heal Psychol.* 2011;11:343-61.
25. López-Pina JA, Sánchez-Meca J, Rosa-Alcázar AI. The Hamilton Rating Scale for Depression: A meta-analytic reliability generalization study. *Int J Clin Heal Psychol.* 2009;9:143-59.
26. Urrutia G, Bonfill X. La Declaración PRISMA: Un paso adelante en la mejora de las publicaciones de la Revista Española de Salud Pública. *Rev Esp Salud Pública.* 2013;87:99-102.
27. Bonett DG. Sample size requirements for testing and estimating coefficient alpha. *J Educ Behav Stat.* 2002;27:335-40.
28. Botella J, Sánchez-Meca J. Meta-análisis en ciencias sociales y de la salud. Madrid: Síntesis, 2015.
29. Sánchez-Meca J, López-López JA, López-Pina JA. Some recommended statistical analytic practices when reliability generalization studies are conducted. *Brit J Math Stat Psy.* 2013;66:402-25.
30. López-López JA, Botella J, Sánchez-Meca J., Marín-Martínez F. Alternatives for mixed-effects meta-regression models in the reliability generalization approach: A simulation study. *J Educ Behav Stat.* 2013;38:443-69.
31. Viechtbauer W. Conducting meta-analysis in R with the metafor package. *J Stat Softw.* 2010;36:1-48.
32. Sánchez-Meca J, López-Pina JA, López-López JA, Marín-Martínez F, Rosa-Alcázar AI, Gómez-Conesa A. The Maudsley Obsessive-Compulsive Inventory: A reliability generalization meta-analysis. *Int J Clin Heal Psychol.* 2011;11:473-93.
33. Nunnally JC, Bernstein IH. The assessment of reliability. *Psychom Theory.* 1994;3:248-92.
34. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in Psychology. *Psychol Assessment.* 1994;6:284-90.

Anexo 1
Características de los estudios incluidos en el meta-análisis

Estudio	Tipo de estudio	Continente	Versión	Adaptación	Población	Edad Media (años)	Sexo % de mujeres	Etnia % caucásicos	Media LOI-CV síntomas	DT LOI-CV síntomas	Coefficiente de fiabilidad
Berg <i>et al</i> , 1988	Psicométrico	América	Original	Original	No clínica	--	44,80	94	6,84	3,94	0,810
Blanco <i>et al</i> , 2014	Sustantivo	Europa	Polaco	Estandarizada	Clínica	14,80	100,00	100	8,20	4,40	0,790*
Bryiska <i>et al</i> , 2005	Sustantivo	Europa	Polaco	Estandarizada	No clínica	--	53,10	--	9,67	3,52	0,630*
Buse <i>et al</i> , 2012	Sustantivo	Europa	Alemán	Libre	Mixta	11,10	--	--	2,78	3,36	0,830*
Canals <i>et al</i> , 2012	Psicométrico	Europa	Español	Estandarizada	No clínica	10,23	52,10	--	8,55	3,75	0,790
Castro-Fornieles <i>et al</i> , 2008	Sustantivo	Europa	Español	Libre	Clínica	14,70	97,70	--	10,74	5,60	0,850*
Castro-Fornieles <i>et al</i> , 2007a	Sustantivo	Europa	Español	Libre	Clínica	14,50	91,67	--	9,80	3,70	0,670*
Castro-Fornieles <i>et al</i> , 2007c	Sustantivo	Europa	Español	Libre	Clínica	14,40	98,00	--	9,80	4,40	0,780*
Castro-Fornieles <i>et al</i> , 2009a	Sustantivo	Europa	Español	Libre	Clínica	14,70	91,67	--	9,80	3,70	0,670*
Castro-Fornieles <i>et al</i> , 2010a	Sustantivo	Europa	Español	Libre	Clínica	16,02	100,00	--	10,68	4,47	0,790*
Castro-Fornieles <i>et al</i> , 2010b	Sustantivo	Europa	Español	Libre	No clínica	16,25	100,00	--	6,41	2,98	0,540*
Coles <i>et al</i> , 2010	Psicométrico	América	Original	Original	Clínica	13,72	55,17	89,70	13,35	5,48	0,900*
Crawford <i>et al</i> , 2005a	Sustantivo	Europa	Original	Original	Clínica	14,40	65,00	--	5,55	6,19	0,940*
Crawford <i>et al</i> , 2005b	Sustantivo	Europa	Original	Original	No clínica	14,25	65,00	--	4,00	6,25	0,960*
Cummings <i>et al</i> , 2002	Sustantivo	América	Original	Original	Clínica	10,40	16,70	87,50	1,80	3,10	0,870*
Gallelli <i>et al</i> , 1997	Sustantivo	América	Original	Original	Clínica	10,90	86,67	87	5,45	4,59	0,850*
King <i>et al</i> , 1995a	Psicométrico	Oceanía	Original	Original	No clínica	--	45,28	--	9,50	3,32	0,580*
King <i>et al</i> , 1995b	Sustantivo	Oceanía	Original	Original	No clínica	--	48,63	--	9,53	3,97	0,760
King <i>et al</i> , 1995c	Sustantivo	Oceanía	Original	Original	Clínica	--	90,00	--	5,90	5,04	0,880*

Anexo 1 (continuación)											
Estudio	Tipo de estudio	Continente	Versión	Adaptación	Población	Edad Media (años)	Sexo % de mujeres	Etnia % caucásicos	Media LOI-CV (síntomas)	DT LOI-CV síntomas	Coefficiente de fiabilidad
Lázaro <i>et al.</i> , 2011a	Sustantivo	Europa	Español	Libre	Clínica	15,60	44,40	--	11,30	3,80	0,690*
Lázaro <i>et al.</i> , 2011b	Sustantivo	Europa	Español	Libre	No Clínica	16,10	51,86	-	5,50	2,90	0,550*
Lázaro <i>et al.</i> , 2013a	Sustantivo	Europa	Español	Libre	Clínica	16,50	--	--	10,90	4,00	0,730*
Lázaro <i>et al.</i> , 2013b	Sustantivo	Europa	Español	Libre	No Clínica	16,70	--	--	6,20	2,90	0,520*
Libby <i>et al.</i> , 2004	Sustantivo	Europa	Original	Original	Clínica	14,80	75,00	-	--	--	0,880
Maggini <i>et al.</i> , 2001	Sustantivo	Europa	Italiano	Libre	No Clínica	--	100,00	--	9,10	3,80	0,690*
Mathews <i>et al.</i> , 2007	Sustantivo	América	Original	Original	Clínica	--	--	--	6,10	4,60	0,840*
Matthews <i>et al.</i> , 2006	Sustantivo	Europa	Original	Original	No Clínica	--	56,00	-	9,70	4,49	0,830
Noorian <i>et al.</i> , 2013	Sustantivo	Europa	Español	Libre	No Clínica	13,92	47,30	78,80	--	--	0,750
Pietrefesa <i>et al.</i> , 2010	Sustantivo	América	Original	Original	Clínica	13,86	48,86	89,28	13,35	5,48	0,900*
Pla <i>et al.</i> , 1999	Sustantivo	Europa	Español	Libre	Clínica	14,60	100,00	--	7,50	4,40	0,800*
Robertson <i>et al.</i> , 2006	Sustantivo	Europa	Original	Original	Clínica	12,00	20,80	--	9,20	4,90	0,830*
Rodríguez-Ramos <i>et al.</i> , 1998	Sustantivo	Europa	Español	Libre	Clínica	15,28	42,85	--	16,14	3,17	0,730*
Roessner <i>et al.</i> , 2009a	Sustantivo	Europa	Alemán	Libre	Clínica	11,90	,00	--	3,20	3,00	0,740*
Roessner <i>et al.</i> , 2009b	Sustantivo	Europa	Alemán	Libre	No Clínica	12,19	,00	--	3,10	2,50	0,610*
Roussos <i>et al.</i> , 2003	Psicométrico	Europa	Griego	Libre	No Clínica	--	--	--	12,30	3,50	0,831
Rueda-Jaimesa <i>et al.</i> , 2007	Psicométrico	América	Español	Estandarizada	No Clínica	14,70	54,70	--	--	--	0,757
Serrano <i>et al.</i> , 1997	Sustantivo	Europa	Español	Estandarizada	No Clínica	13,44	55,00	--	10,04	3,20	0,540*
Stewart <i>et al.</i> , 2005	Sustantivo	América	Original	Original	Clínica	11,50	--	--	8,10	4,30	0,780*
Swedo <i>et al.</i> , 1989a	Sustantivo	América	Original	Original	Clínica	--	56,53	--	8,00	4,10	0,750*
Swedo <i>et al.</i> , 1989b	Sustantivo	América	Original	Original	Clínica	--	35,71	--	4,29	3,60	0,780*
Thomsen PH, 1993	Sustantivo	Europa	Danés	Libre	No Clínica	13,80	52,71	--	9,60	4,40	0,780*
Thomsen, 1994a	Sustantivo	Europa	Danés	Libre	Clínica	13,20	20,00	--	8,60	8,10	0,970*
Thomsen, 1994b	Sustantivo	Europa	Danés	Libre	Clínica	13,20	20,00	--	4,10	3,40	0,760*

* Coeficiente de fiabilidad calculado con la fórmula KR-21.