

ARTÍCULO DE INVESTIGACIÓN

**CUESTIONARIO DE EFECTIVIDAD DE LA PSICOTERAPIA
OUTCOME QUESTIONNAIRE. VALIDACIÓN EN MUESTRA
CLÍNICA COLOMBIANA**

QUESTIONS EFFECTIVENESS OF PSYCHOTHERAPY OUTCOME QUESTIONNAIRE.
VALIDATION IN COLOMBIAN CLINICAL SAMPLE

NORA HELENA LONDOÑO¹

UNIVERSIDAD DE SAN BUENAVENTURA, MEDELLÍN - COLOMBIA

DIANA MARÍA AGUDELO

UNIVERSIDAD DE LOS ANDES, BOGOTÁ - COLOMBIA

EFRÉN MARTÍNEZ

UNIVERSIDAD DE MANIZALES, MANIZALES - COLOMBIA

DEISSY ANGUILA

UNIVERSIDAD DEL NORTE, BARRANQUILLA - COLOMBIA

DANIEL CAMILO AGUIRRE

UNIVERSIDAD DE ANTIOQUIA, MEDELLÍN - COLOMBIA

CARLOS MOGOLLÓN

UNIVERSIDAD DE SAN BUENAVENTURA, MEDELLÍN - COLOMBIA

FECHA RECEPCIÓN: 17/2/2016 • FECHA ACEPTACIÓN: 21/11/2016

Para citar este artículo: Londoño, N., Agudelo, D., Martínez, E., Anguila, D., Aguirre, D., Mogollón, C. (2017). Cuestionario de efectividad de la psicoterapia Outcome Questionnaire. Validación en muestra clínica colombiana. *Psychologia*, 11(1), 97-108.

Resumen

Establecer las propiedades estructurales y psicométricas del OQ-45.2 e identificar la sensibilidad al cambio y diferencias según género. Método: se realizó en población clínica, participaron 214 consultantes (111 hombres y 103 mujeres), Instrumento: Outcome Questionnaire (OQ-45.2). Resultados: El AFC mostró índices de bondad de ajuste bajos y valor de RMSEA de 0.071. Se recomendó realizar el AFE, y esta estructura incluyó 20 ítems y explicó el 58,8 % de la varianza total: preocupación, problemas del desempeño del rol social, insatisfacción

¹ Correspondencia: Nora Helena Londoño. Carrera 56 C 51-110, Universidad de San Buenaventura, Medellín, Colombia. Mail: nora.londono@usbmed.edu.co; nora_londono@yahoo.es

global, alcohol/drogas, dificultades en las relaciones e insatisfacción en las relaciones afectivas. Para la escala original, los coeficientes se encontraron entre 0.52 y 0.92; no se presentaron diferencias según género en ninguno de los factores pero sí sensibilidad al cambio en la terapia en todas las dimensiones. Para la estructura encontrada a través del AFE con 20 ítems, los coeficientes se encontraron entre 0.47 y 0.75. Los índices de ajustes para esta nueva estructura fueron muy favorables (CFI=0,920, TLI=0,902 y NFI=0,790) y valor de RMSEA de 0.046. Se reportaron diferencias con relación al género en las dimensiones Insatisfacción global (más elevadas en mujeres) y Alcohol/droga (puntuaciones más elevadas en hombre). La prueba reportó sensibilidad al cambio en la escala original en Estrés, Desempeño y Relaciones interpersonales negativas, y en la escala abreviada en Preocupación, Problemas del rol social e Insatisfacción en las relaciones afectivas. Conclusiones: el OQ 45.2 no presentó una estructura adecuada a través del AFC. El AFE reportó 6 factores y 20 ítems, adecuados índices de ajustes, pero no todos los factores con buen nivel de confiabilidad.

Palabras clave: Estudio de validación; salud mental; OQ-45.2; confiabilidad; psicometría.

Abstract

To establish the structural and psychometric properties of the OQ-45.2 and identify sensitivity to change and differences by gender. Method: It was done in clinical population included 214 consultants (111 men and 103 women) Instrument: Outcome Questionnaire (OQ-45.2). Results: The AFC showed low rates of kindness adjustment and RMSEA value of 0.071. It was recommended that the AFE, and this structure included 20 items, explaining 58.8% of the total variance: worry, problems performing the social role, overall dissatisfaction, alcohol / drugs, relationship difficulties and dissatisfaction with relationships. For the original scale, the coefficients were found between 0.52 and 0.92; no gender difference appeared in any of the factors but sensitivity to change in therapy in all dimensions. For the structure found throughout the AFE with 20 items, the coefficients were between 0.47 and 0.75. The rates for this new structure adjustments were very suitable (CFI = 0.920, TLI = 0.902 and NFI = 0.790) and RMSEA value of 0.046. Gender-related differences in size (higher in women) overall dissatisfaction and Alcohol / drug (higher scores in man) were reported. The test reported sensitivity to change in the original scale Stress, Performance and negative interpersonal relations, and the abbreviated scale Worry, Problems of the social role and Dissatisfaction in relationships. Conclusions: 45.2 OQ did not provide a suitable structure through the AFC. The AFE reported six factors and 20 items, appropriate indexes setup, but not all factors with a high level of reliability.

Keywords: Validation studies; mental health; OQ-45.2; reliability; psychometrics.

Introducción

El interés por evaluar la efectividad en la terapia se ha incrementado en las últimas décadas de acuerdo con los reportes investigativos, y uno de los instrumentos más citados para tal fin es el cuestionario sobre efectividad en la terapia (*Outcome Questionnaire*; OQ 45.2) de Lambert, Lunnen, Umphress, Hansen y Burlingame (1996). Este cuestionario fue diseñado para evaluar los resultados y monitorear el progreso de los pacientes en psicoterapia. Evalúa estrés o malestar sintomático (*Symptom Distress*), relaciones interpersonales (*Interpersonal Relationships*), desempeño del rol social (*Social Role Performance*) y el funcionamiento global; este último, permite discriminar correctamente entre muestras clínicas y no-clínicas (Lambert, 2015).

Algunos estudios han utilizado el cuestionario como instrumento para comparar la efectividad entre diferentes modalidades de intervención (terapias individuales, grupales, combinadas, farmacológica, etc.), en trastornos específicos como la depresión (Blais et

al., 2013; Connolly et al., 2015) o en otros trastornos (Burlingame et al., 2015; Holt et al., 2015).

El instrumento ha sido utilizado también para comparar el efecto de las intervenciones según duración (Gullo, Lo Coco, & Gelso, 2012; Painepán & Kühne, 2012), grado de alianza terapéutica y experiencia del terapeuta (Santibáñez, Román, & Vinet, 2009).

Con relación a la estructura psicométrica del instrumento, diversos estudios lo han encontrado confiable y válido (Boswell, White, Sims Harrist, & Romans, 2013; Garrido, Santelices, & Pierrehumbert, 2009; Miller, Duncan, Brown, Sparks, & Claud, 2003; Shulte, 1977). La confiabilidad a través del alfa de Cronbach reportado por Lambert et al. (2006) fueron adecuados para todos los factores; estrés o malestar sintomático 0.93, relaciones interpersonales 0.92 y desempeño del rol social 0.70.

Adicional a esto, estudios de validez se han realizado en diferentes países y en general se han confirmado propiedades psicométricas adecuadas (Lambert, 2015). En Estados Unidos, Lambert et al. (2006), obtuvieron

correlaciones satisfactorias; estrés o malestar sintomático 0.78, relaciones interpersonales 0.80 y desempeño del rol social 0.84. En Italia, los resultados mostraron propiedades psicométricas similares a la escala original y el análisis factorial confirmatorio (AFC) soportó el constructo multidimensional del instrumento, aunque los valores normativos fueron diferentes con los reportados por Lambert et al. (2006) en población americana (Chiappelli, Lo Coco, Gullo, Bensi & Prestano, 2008).

También el estudio realizado por De la Parra y Von Bergen (2006) en Chile reportó valores de correlación para la confiabilidad test re-test del 99 % de confianza. En Noruega, Amble et al. (2014), reportaron valores normativos similares a los reportados por Lambert et al. (2006), al igual que Gross et al. (2014), con muestras hebreas y arábigas. Por su parte, Machado y Fassnacht (2014), en la validación portuguesa encontraron ajuste en los datos normativos comparables con la versión original, pero diferentes al comparar muestras clínicas con no clínicas. Similares hallazgos se encontraron en el estudio de Timman, Jong y Neve-Enthoven (2015) en muestras poblacionales alemanas, y el estudio de Takara, Beecher, Okiishi, Shimokawa, Lambert y Griner (2015) en Japón, donde demostraron las adecuadas propiedades psicométricas de cuestionario.

Por otro lado, algunos estudios han encontrado resultados contradictorios. Lara, Cruz, Vacarezza, Florenzano y Trapp (2008), no encontraron relaciones entre los ítems de desempeño de rol social y los ítems de relaciones interpersonales. Flynn, Aubry, Guindon y Tardif (2002) encontraron una mejor utilidad del instrumento con factores tales como recuperación, mejora, sin cambios y deterioro, diferente a lo propuesto por Lambert et al. (1996). Según Lara et al. (2008), el OQ no mide sintomatología de trastornos de alteración grave de la conducta y de la conciencia, y sobre las adicciones el OQ no reporta el grado de conciencia ni severidad. Sin embargo, el instrumento puede ser aplicado en pacientes institucionalizados para determinar la mejoría producida por los procesos de intervención, salvo en aquellos que presentan diagnóstico por trastornos orgánicos (Correa et al., 2006).

Con relación a la sensibilidad para detectar cambio, se ha evidenciado que el cuestionario resulta ser útil y válido para ilustrar el grado de mejoría experimentada por los pacientes hospitalizados (Correa, Florenzano, & Rojas, 2006).

De acuerdo con los antecedentes, se evidencia la utilidad del instrumento para identificar la efectividad de la psicoterapia dados sus niveles adecuados de sensibilidad al cambio, y la necesidad de estudiar los valores psicométricos dada la variabilidad reportada de manera frecuente dependiendo de la cultura. Dentro de los antecedentes, no se reportan estudios de validez del instrumento en Colombia. Congruente con los vacíos investigativos y la utilización del instrumento en población clínica, el propósito del estudio fue aportar con la validación del instrumento en Colombia. La adaptación y validación de escalas casi siempre son necesarias, particularmente en estudios que han sido desarrollados en otro ámbito, y a menudo, en otras culturas (Meyer et al., 2001; Ruiz, Gómez, & Londoño, 2001). En este sentido, se justifica una investigación cuyo propósito sea adaptar y validar instrumentos, para este caso, en el ámbito de la psicología clínica en Colombia, y cuyo objetivo sea establecer las propiedades estructurales y psicométricas del OQ-45.2 a describir la sensibilidad al cambio y diferencias según género en población clínica colombiana.

Método

Muestreo

Se realizó un muestreo censado estableciendo como tiempo seis meses de aplicación de los instrumentos a todos los pacientes que solicitaran atención psicológica en diferentes centros de prácticas de universidades o consultorios particulares (primera observación). Para la segunda evaluación (segunda observación), se estableció como tiempo tres meses después de haber realizado la primera observación. Los centros de atención fueron: en Bogotá el Centro de Atención Psicológica Casa Espinosa de la Universidad de los Andes y en la Sociedad para el Avance de la Psicoterapia Centrada en el Sentido. En Medellín, Armenia e Ibagué, en los consultorios y centros de práctica de la Universidad de San Buenaventura. En Barranquilla, el Centro de Atención de la Universidad del Norte: CAIF Santa Rosa De Lima.

Muestra

Participaron 214 consultantes de servicios psicológicos (111 hombres y 103 mujeres), 41 (19 %) de los cuales fueron evaluadas por segunda vez. La edad media

de los participantes fue de 26 años (desviación 11,3; rango 14 - 65 años). Los participantes en su mayoría eran solteros (72,4 %), los niveles socioeconómicos más representativos fueron el bajo y el medio (48 % y 38 % respectivamente), el nivel académico estuvo representado en su mayoría por estudiantes (70 %), de básica secundaria y universitario (70 %), y el 35 % trabajaban. Los participantes fueron atendidos desde los diferentes enfoques de la psicología (cognitivo-conductual 45,8 %; psicoterapia analítica/dinámica 37,4 % y humanista/existencial/logoterapia 16,8 %), tanto en modalidad individual (77 %) como en doble modalidad (23 %; individual y terapia grupal, familiar, interconsulta psiquiátrica o neurológica). La tabla 1 describe los datos sociodemográficos con relación al sexo, estado civil, estrato, nivel académico, ocupación, y la atención recibida.

Tabla 1. Datos sociodemográficos y modalidad de tratamiento en participantes del estudio de validación del OQ-45.2 (Lambert et al., 1996) en población clínica colombiana.

Categoría	Frecuencia	%
Sexo		
Hombre	111	51,9
Mujer	103	48,1
Estado civil		
Soltero	155	72,4
Casado	44	20,6
Separado	13	6,1
Viudo	2	0,9
Ciudad		
Barranquilla	18	8,4
Bogotá	67	31,3
Ibagué	86	40
Armenia y Manizales	13	6,1
Medellín	28	13,1
Estrato socioeconómico		
Alto	27	12,7
Medio	81	37,9
Bajo	103	48

Nivel académico	Frecuencia	%
Primaria	22	10,3
Secundaria	82	38,3
Técnico	24	11,0
Universitario	69	32,2
Posgrado	17	7,9
Ocupación		
Ni estudia ni trabaja	1	0,4
Estudia	137	64
Trabaja	63	29,4
Estudia y trabaja	13	6,1
Tipo de terapia		
Cognitivo / conductual	98	45,8
Psicoterapia analítica/ dinámica	80	37,4
Humanista/existencial/ logoterapia	36	16,8
Modalidad		
Individual	165	77
Doble modalidad	45	23

Instrumento

El Outcome Questionnaire OQ-45.2 (Lambert et al., 1996) es un instrumento autoaplicado, creado para medir progresos en psicoterapia a través de mediciones sucesivas. La evaluación se realiza en tres áreas y contiene ítems invertidos identificados con asteriscos (*).

SD: Estrés o malestar sintomático (*Symptom Distress*). Lo conforman los ítems 2, 3, 5, 6, 8, 9, 10, 11, 13*, 15, 22, 23, 24*, 25, 27, 29, 31*, 33, 34, 35, 36, 40, 41, 42, 45.

IR: Relaciones interpersonales (*Interpersonal Relationships*): lo conforman los ítems 1*, 7, 16, 17, 18, 19, 20*, 26, 30, 37*, 43*.

SR: Desempeño del rol social (*Social Role Performance*): lo conforman los ítems 4, 12*, 14, 21*, 28, 32, 38, 39, 44.

Cuestionario sociodemográfico: se aplicó un cuestionario con la descripción general del participante: sexo, edad, estado civil, nivel académico y ocupación. Adicional a esto, el psicólogo debía responder sobre su enfoque

teórico, y si se realizaba solo terapia individual u otras modalidades como trabajo interdisciplinario, terapia grupal o de pareja.

Procedimiento

Se propuso a los psicólogos y practicantes de psicología que atendían en los diferentes centros de atención, participar en la investigación aplicando a cada paciente que iniciara tratamiento psicológico (observación 1) y luego repetir el procedimiento a los tres meses (observación 2). Para los que aceptaron participar, se les entrenó en los procedimientos establecidos: presentación del consentimiento informado y la carta de aprobación del comité de bioética expedida el 16 de julio del 2014, aprobación y aplicación de instrumentos. Se aplicó el cuestionario en la versión de papel y de manera individual. El psicólogo fue el encargado de su aplicación y la aclaración de las inquietudes que surgieron para responderlo. El tiempo aproximado de aplicación fue de 10 minutos.

El tratamiento se orientó a la intervención psicológica individual y apoyo con la familia y grupal en la mayoría de los casos. El proceso se basó en la evaluación de las condiciones problemáticas del paciente, psicoeducación sobre su problemática, los factores de riesgo y las estrategias de intervención, teniendo presente las experiencias emocionales, las interpretaciones y pensamientos, las competencias interpersonales, las habilidades en solución de problema. Se analizaron los antecedentes personales y la relación con la problemática como factores predisponentes y dependiendo del enfoque teórico del terapeuta se realizó el procedimiento correspondiente. El tipo de tratamiento fue cognitivo-conductual, psicoterapia dinámica/analítica, o humanista/existencial/logoterapia, correspondiente al enfoque del terapeuta.

Procedimiento estadístico

Se realizaron análisis para la validez de constructo basados en el Análisis Factorial Confirmatorio (AFC). Se calcularon como medidas de bondad de ajuste los índices de bondad de ajuste comparativos (CFI), índice de Tucker-Lewis (TLI), índice de ajuste normalizado (NFI); Para estos índices se sugieren valores superiores a 0.95 para indicar un buen ajuste. También se calculó el índice RMSEA (*Root Mean Square Error of Approximation*) el cual debía ser menor a 0.08 para indicar un buen

ajuste. Se propuso el Análisis Factorial Exploratorio (AFE) utilizando el paquete AMOS versión 18 si la escala no mostraba un buen ajuste, para este análisis se debía calcular el índice Kaiser Meyer Olkin (KMO) y la prueba de esfericidad de Bartlett. Los criterios de selección de los factores debían ser los siguientes:

1. Factores con valor propio mayor que uno
2. Factores que incluyeran al menos 3 ítems
3. Los ítems debían tener una carga factorial igual o superior a 0,40
4. Los ítems no podían tener cargas similares a dos o más factores.

Adicionalmente, se realizó rotación Varimax para mejorar la interpretación de los factores.

Con relación al procedimiento estadístico, se realizó el análisis de consistencia interna para cada dimensión del instrumento en su versión original y en la versión corta obtenida a través del AFE, mediante el cálculo del coeficiente alfa de Cronbach con su respectivo intervalo de confianza del 95 %.

Se analizaron diferencias de género y se realizó la construcción de los valores de referencia. Se exploró si las dimensiones del cuestionario OQ 45.2 se diferenciaban entre hombres y mujeres. Para este análisis se evaluó el supuesto de distribución normal en cada grupo utilizando la prueba de Kolmogorov-Smirnov. En caso de no cumplirse el supuesto de distribución normal se sugirió utilizar la prueba U de Mann-Whitney.

Adicionalmente, se calculó el tamaño del efecto (dnp) no paramétrico, el cual se calcula dividiendo el estadístico z de la prueba U de Mann-Whitney por la raíz cuadrada del total de la muestra. Un dnp >0.50 sugiere una diferencia importante entre los grupos. Para el control de la tasa de error tipo I, se dividió el valor convencional $\alpha=0,05$ entre 7 (7 dimensiones encontradas en el AFE) para un nivel de significación de 0.001 para determinar si las diferencias entre hombres y mujeres fueron estadísticamente significativas.

Para los valores de referencia se calcularon la media con su respectiva desviación estándar y los percentiles (1 a 10, 20, 30, 40, 50, 60, 70, 80, 90 y 91 a 99).

Para identificar la sensibilidad al cambio, se compararon los resultados entre las mediciones basales y posintervención utilizando la prueba de Suma de Rangos de Wilcoxon para muestras pareadas. Como medida de

efecto se calculó la delta de Glass, que consiste en la razón de la diferencia de medias dividida por la desviación estándar basal. Un valor superior a 0,75 se consideró clínicamente importante.

Resultados

Validez de constructo. Análisis factorial confirmatorio y Análisis factorial exploratorio

El AFC mostró índices de bondad de ajuste bajos, inferiores a 0.95 para la versión original (CFI=0,670, TLI=0,653 y NFI=0,51) y valor de RMSEA de 0.071. Dado que la estructura original del OQ-45.2 no mostró un adecuado ajuste se procedió a la realización del AFE.

La solución final del AFE (tabla 2) mostró un índice Kaiser-Meyer-Olkin de 0.813, y un resultado

significativo en la prueba de esfericidad de Bartlett (Chi cuadrado=1037,7; 190 gl, $p<0.001$); se encontró una estructura de 6 factores que explicaron el 58,8 % de la varianza total.

Esta estructura incluyó 20 ítems. Factor 1 (preocupación) incluyó 4 ítems y explicó el 14.55. El Factor 2 (problemas en el desempeño del rol social) incluyó 5 ítems y explicó el 15.4 %. El Factor 3 (insatisfacción global) incluyó 4 ítems con enunciados redactados de manera invertida y explicó el 11.7 %. El Factor 4 (alcohol/drogas) incluyó 2 ítems y explicó el 4.2 %. El Factor 5 (dificultades en las relaciones) incluyó 2 ítems con enunciados redactados de manera invertida y explicó el 3.2 %. El Factor 6 (insatisfacción en las relaciones interpersonales negativas) incluyó 3 ítems, uno de los enunciados redactado de manera invertida y explicó el 6.1 % (tabla 2).

Tabla 2. Resultados del análisis factorial exploratorio del OQ-45.2 (Lambert et al., 1996) en población clínica colombiana.

Descripción del ítem	1	2	3	4	5	6
	Preocupación	Problemas del desempeño del rol social	Insatisfacción global	Alcohol/drogas	Dificultades en las relaciones	Insatisfacción en las relaciones afectivas
q33 He sentido que algo malo va a suceder	,769					
q40	,682					
oq25	,676					
q23	,646					
q4 Me he sentido estresado en el trabajo/estudios		,721				
q2		,691				
q39		,638				
q5		,550				
oq41		,532				
q20 Me he sentido amado y querido*			,736			
oq37	*		,726			
q21	*		,702			
q13	*		,559			

q32	He tenido problemas en el trabajo/estudios por el uso de alcohol/ drogas.					,856	
q26					,805	
q1	Me he relacionado bien con los demás*					,830	
q43*					,653	
q17	He tenido una vida sexual insatisfactoria					,739	
q19					,572	
q16					,558	
		14.480	15.391	11.661	4.212	3.198	6.126
% varianza explicada		,752	.706	.690	.730	.511	.467
Alfa de Cronbach (IC95%)		(0,693-0,802)	(0,630-0,758)	(0,614-0,751)	(0,646-0,794)	(0,358-0,625)	(0,325-0,576)

Análisis de la consistencia interna

La consistencia interna de los factores que integran el OQ-42.2 en población clínica en Colombia mostró coeficientes entre 0,52 y 0,92. Estos índices fueron: estrés o malestar sintomático 0,90 (IC95 % 0,88 a 0,92), relaciones interpersonales 0,71 (IC95 % 0,65 a 0,76), desempeño del rol social 0,52 (IC95 % 0,42 a 0,61), indicador global 0,92 (IC95 % 0,90 a 0,93). La tabla 2, también muestra la consistencia interna para la estructura de 20 ítems y los 6 factores del AFE. Al analizar el índice de bondad de ajuste, los valores fueron muy favorables excepto en uno de los indicadores (se sugieren valores superiores a 0,95): CFI=0,92; TLI=0,92; y

NFI=0,790. El valor de RMSEA fue favorable (se sugiere inferior a 0,08): 0,046.

Diferencias según género

La tabla 3, presenta los resultados de las dimensiones del OQ-45.2 (Lambert et al., 1996) y de la versión abreviada de 20 ítems según género. Para las dimensiones originales del OQ-45.2, no se encontraron diferencias significativas con relación al género. En la versión abreviada de 20 ítems se encontraron diferencias significativas en dos de los factores insatisfacción global (mayores puntajes en mujeres) y alcohol/drogas (mayores puntajes en hombres) (tabla 3).

Tabla 3. Resultados según género para las dimensiones del OQ-45.2 (Lambert et al., 1996) y de la versión reducida de 20 ítems en población clínica en Colombia.

	Masculino n=94				Femenino n=106				D	Z	valor p
	Media	Desv. típ.	Mediana	RI	Media	Desv. típ.	Mediana	RI			
Estrés o malestar sintomático	8,9	15,8	39,0	25,0	41,2	16,6	42,0	25,0	0,06	-0,84	0,402
Relaciones interpersonales	7,0	7,1	17,0	10,0	17,0	6,1	17,0	10,0	0,01	-0,13	0,894
Desempeño del rol social	2,8	4,8	13,0	6,0	13,3	4,9	13,0	7,0	0,05	-0,76	0,450
OQ-45.2											
Preocupación	6,3	3,8	6,0	6,0	6,3	3,8	6,0	6,0	-,1	-,006	0,928
Problemas en el desempeño del rol social	9,3	4,1	9,0	5,0	9,6	3,7	10,0	5,0	-,6	-,040	0,555
Insatisfacción global	5,9	3,6	6,0	5,0	7,3	3,1	7,0	4,0	-3,2	-,219	0,001

Esta tabla continúa en la siguiente página —>

	Masculino n=94				Femenino n=106				D	Z	valor p
	Media	Desv. típ.	Mediana	RI	Media	Desv. típ.	Mediana	RI			
Alcohol/ drogas	1,8	2,3	0,0	4,0	0,8	1,6	0,0	1,0	-2,8	-,192	0,005
Dificultades en las relaciones	2,8	1,9	3,0	3,0	2,7	1,7	2,0	3,0	-,4	-,029	0,674
Insatisfacción en las relaciones afectivas	5,0	2,6	5,0	4,0	5,0	2,3	5,0	4,0	-,1	-,009	0,895

d: tamaño del efecto no paramétrico. Z=Estadístico z de la prueba U de Mann Whitney

Sensibilidad al cambio

El cuestionario OQ-45.2 mostró una adecuada sensibilidad al cambio en todos los factores al identificar diferencias significativas con las mediciones basales y postintervención utilizando la prueba de Wilcoxon, pero no fueron favorables al analizar los tamaños del efecto los cuales se esperaban superiores a 0.75: entre -0.47 (relaciones interpersonales) y -0.67 (estrés o

malestar psicológico) y diferencias estadísticamente significativas entre la medición basal y la medición posintervención ($p < 0.06$). En la versión abreviada de 20 ítems se encontraron diferencias entre la medición basal y posintervención en los factores relacionados con preocupación, problemas en el desempeño del rol social e insatisfacción en las relaciones afectivas (tabla 4).

Tabla 4. Descriptivos de las mediciones y sensibilidad al cambio del OQ-45.2 (Lambert et al., 1996) en población clínica colombiana.

	Basal		Post intervención		d	z	Valor p
	Mediana	RI	Mediana	RI			
OQ-45.2							
Estrés o malestar sintomático	46	32	28	16	-0,67	-3,9	,000
Relaciones interpersonales	20	11	16	8	-0,47	-2,8	,006
Desempeño del rol social	16	10	11	6	-0,58	-3,4	,001
OQ-20							
Preocupación	7,0	6,0	4,0	4,0	-0,67	-4,2	0,00
Problemas en el desempeño del rol social	11,0	5,0	6,0	6,0	-0,75	-4,7	0,00
Insatisfacción global	7,0	7,0	6,0	6,0	-0,22	-1,4	0,18
Alcohol/ drogas	1,0	4,0	1,0	3,0	-0,03	-0,2	0,84
Dificultades en las relaciones	2,0	3,0	2,0	3,0	-0,02	-0,1	0,92
Insatisfacción en las relaciones afectivas	6,0	4,0	4,0	2,0	-0,27	-1,7	0,09

RI: Rango intercuartil, d: tamaño del efecto no paramétrico; z: estadístico z con aproximación asintótica de la prueba U de Mann Whitney. Prueba de suma de rangos de Wilcoxon

Valores de psicómétricos de la muestra clínica

Los anexos 1, 2 y 3 presentan los valores psicómétricos de la muestra clínica para la versión original del OQ-45.2 y para la versión abreviada de 20 ítems producto del AFE.

Discusión

El OQ 45.2 (Lambert et al., 1996) es un cuestionario referenciado en la psicología clínica para evaluar efectividad en la psicoterapia, y fue creado para

monitorear el progreso de los pacientes en psicoterapia. Evalúa estrés o malestar sintomático (*Symptom Distress*), relaciones interpersonales (*Interpersonal Relationships*), desempeño del rol social (*Social Role Performance*) y el funcionamiento global; para el presente estudio en muestra clínica colombiana, el objetivo fue establecer las propiedades estructurales y psicométricas del OQ-45.2 e identificar la sensibilidad al cambio y diferencias según género.

De acuerdo con los antecedentes, se han evidenciado valores psicométricos favorables de sensibilidad al cambio (Blais et al., 2013; Burlingame et al., 2015; Connolly et al., 2015; Correa et al., 2006; Holt et al., 2015; Gullo et al., 2012; Painepán & Kühne, 2012; Santibáñez et al., 2009) y niveles satisfactorios de confiabilidad y validez (Boswell et al., 2013; Garrido et al., 2009; Miller et al., 2003; Shulte, 1977; Takara et al., 2015; Timman et al., 2015). Sin embargo, los resultados suelen ser contradictorios frente a la validez estructural del cuestionario (Lara et al., 2008; Flynn et al., 2002).

De acuerdo con los resultados del presente estudio, se logró confirmar la sensibilidad al cambio de acuerdo con las diferencias significativas de las medidas basales y posintervención pero no de acuerdo con las medidas del efecto. La confiabilidad se confirmó satisfactoria para dos de los tres factores, y la validez estructural no mostró ser adecuada para todos los indicadores estudiados; el AFC mostró índices de bondad de ajuste bajos, inferiores a 0.95 para la versión original (CFI=0,670, TLI=0,653 y NFI=0,51), pero valores adecuados de RMSEA, menores a 0.08 (0.071).

El factor de estrés o malestar sintomático (*Symptom Distress*) reportó el mejor nivel de confiabilidad de la prueba (Alfa de Cronbach 0.91), con excelentes valores de sensibilidad al cambio pero no con relación a las medidas de efecto. Los resultados no reportaron diferencias según género. Este factor parece ser uno de los más susceptibles al cambio en la intervención terapéutica ya referido en otros estudios (Lambert et al., 2006; Painepán & Kühne, 2012).

El factor Relaciones interpersonales (*Interpersonal Relationships*) obtuvo un nivel aceptable de confiabilidad (Alfa de Cronbach 0.71), con excelentes valores de sensibilidad al cambio pero no con relación a las medidas de efecto. Los resultados no reportaron tampoco

diferencias según género. Los estudios de Lambert et al. (2006) reportaron en este caso niveles de confiabilidad más elevados (Alfa de Cronbach 0.92).

El factor Desempeño del rol social (*Social Role Performance*) obtuvo un nivel de confiabilidad bajo (Alfa de Cronbach 0.52), con excelentes valores de sensibilidad al cambio pero no con relación a las medidas de efecto. Los resultados tampoco reportaron diferencias según género. Los estudios de Lambert et al. (2006) reportaron sin embargo, niveles de confiabilidad favorables (Alfa de Cronbach 0.70).

De acuerdo con el AFE, la estructura adecuada para la población clínica colombiana fue:

El factor 1: Preocupación. Puntuaciones elevadas describen personas que han sentido que algo malo va a suceder, algo anda mal en su mente, a su mente han acudido pensamientos perturbadores que no puede apartar y han sentido desesperanza sobre el futuro.

El factor 2: Problemas en el desempeño del rol social. Puntuaciones elevadas identifican personas que se han sentido entresados en el trabajo o el estudio, se sienten cansados rápidamente, han tenido demasiados desacuerdos en el trabajo o el estudio, se han culpado a sí mismos de algunas cosas, y han tenido problemas para conciliar el sueño o permanecer dormidos.

Factor 3: Insatisfacción global. Este factor está constituido por ítems invertidos que evalúan bienestar; por lo tanto, puntuaciones elevadas identifican personas que no se han sentido amadas ni queridas, sus relaciones interpersonales no han sido plenas ni completas, no han disfrutado el tiempo libre y no se han sentido felices.

Factor 4: Alcohol/drogas. Este factor describe personas que han percibido problemas en el trabajo o el estudio asociadas al uso de alcohol o drogas, y han sentido que las personas los han criticado por su manera de beber.

Factor 5: Dificultades en las relaciones. Este factor está constituido por ítems invertidos que evalúan satisfacción en las relaciones interpersonales; por lo tanto, puntuaciones elevadas identifican personas que no se han sentido bien con los demás, y no se han sentido satisfechas en las relaciones con otros.

Factor 6: Insatisfacción en las relaciones afectivas. Puntuaciones elevadas deben interpretarse como personas que posiblemente no han tenido una vida sexual

satisfecha, han tendido frecuentes discusiones y han estado demasiado preocupados por los problemas familiares.

Algunos estudios han encontrado de manera similar factores adicionales. En el estudio realizado por Jong, Nugter, Polak, Wagenborg, Spinhoven y Heiser (2007) se encontraron factores consistentes con el rol social, ansiedad y síntomas somáticos. El estudio de Flynn et al. (2002), los factores que mostraron mejor utilidad fueron recuperación, mejora sin cambios y deterioro.

Con relación a las diferencias según el género, en la escala abreviada se reportaron diferencias significativas en dos factores, insatisfacción global, más elevadas en mujeres, y Alcohol/droga (puntuaciones más elevadas en el hombre). Estas puntuaciones concuerdan con las diferencias reportadas en datos de prevalencia de salud mental para trastornos como depresión y ansiedad con mayor prevalencia en mujeres, y dependencia a sustancias con mayor prevalencia en hombres (APA, 2013). Estas diferencias pueden estar relacionadas con patrones de mayor sensibilidad y dependencia, validados socialmente en las mujeres y asociados a problemas internalizantes, y en los hombres por los patrones validados socialmente para el consumo de alcohol y problemas externalizantes asociados.

Respecto con los niveles de sensibilidad al cambio, para la escala original todos reportaron sensibilidad al cambio, pero sin ser significativo en tamaño del efecto, y para la escala abreviada reportado por el AFE, en tres: preocupación, problemas en el desempeño e insatisfacción en las relaciones afectivas, pero igual sin ser significativo al analizar el tamaño del efecto.

Conclusiones: el OQ 45.2 presenta indicadores de confiabilidad favorables para dos de los tres factores, sensibilidad al cambio de acuerdo, pero sin un significativo tamaño del efecto, y validez estructural NO adecuada de acuerdo con los indicadores del AFC. El AFE reportó 6 factores y 20 ítems, adecuados índices de ajustes, pero no todos los factores con buen nivel de confiabilidad.

Las limitaciones del estudio están relacionadas con el número de la muestra, se recomienda continuar avanzando con este tipo de investigaciones frente a la validez estructural del instrumento. Se debe tener en cuenta que los puntajes fueron obtenidos en población clínica, por lo tanto, no son puntos de referencia para población normal.

Referencias

- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders DSM-5.^a* ed. Washington DC.
- Amble, I., Gude, T., Stubdal, S., Oktedalen, T., Skjorten, A. M., Andersen, B. J., & Wampold, B. E. (2014). Psychometric properties of the Outcome Questionnaire-45.2: The Norwegian version in an international context. *Psychotherapy Research, 24*(4), 504-513.
- Bechstead, D. J., Hatch, A. L., Lambert, M. J., Eggett, D. L., Goates, M. K., & Vermeersch, D.A. (2003). Clinical significance of the outcome questionnaire (OQ-45.2). *The Behavior Analyst Today, 4*(1), 86-97.
- Bergen, A., & Parra, G. (2002). OQ-45.2, cuestionario para evaluación de resultados y evolución en psicoterapia: adaptación, validación e indicaciones para aplicación e interpretación. *Terapia Psicológica, 20*(2), 161-176.
- Blais, M. A., Malone, J. C., Stein, M. B., Slavin-Mulford, J., O'Keefe, S. M., Renna, M., & Sinclair, S. J. (2013). Treatment as usual (TAU) for depression: a comparison of psychotherapy, pharmacotherapy, and combined treatment at a large academic medical center. *Psychotherapy, 50*(1), 110.
- Boswell, D., White, J., Sims, W., Harrist, S., & Romans, J. (2013). Reliability and validity of the Outcome Questionnaire-45.2. *Psychological Reports: Mental and Physical Health: 112*, 1-10.
- Burlingame, G., Gleave, R., Erekson, D., Nelson, P., Olsen, J., Thayer, S., & Beecher, M. (2015). Differential effectiveness of group, individual and conjoint treatments: an archival analysis of OQ-45 change trajectories. *Psychotherapy Research, 26*(5):1-17 doi: 10.1080/10503307.2015.1044583.
- Chiappelli, M., Lo Coco, G., Gullo, S., Bensi, I., & Prestano, C. (2008). The Outcome Questionnaire 45.2. Italian validation of an instrument for the assessment of psychological treatments. *Epidemiologia e Psichiatria sociale, 17*(2), 152-161.
- Connolly, M., Kurtz, J., Thompson, D., Mack, R., Lee, J. Eisen, S., Gallop, R., & Crits-Christoph, P. (2015). The Effectiveness of clinician feedback in the treatment of depression in the community

- mental health system. *Journal of Consulting and Clinical Psychology*, 83, 748-759.
- Correa D., Florenzano, R., & Rojas, P. (2006). El uso del cuestionario OQ-45.2 como indicador de psicopatología y de mejoría en pacientes psiquiátricos hospitalizados. *Revista Chilena de Neuro-Psiquiatría*, 44(4), 258-262.
- De la Parra, G., & Von Bergen, G. (2006). Medición de los resultados en psicoterapia: uso del OQ.45.2, un instrumento validado en Chile. *Revista GU*, 2(2), 2008-222.
- Flynn R., Aubry T., Guindon, S., & Tardif, I. (2002). Validation d'une Version Française du Outcome Questionnaire et Évaluation d'un Service de Counselling en Milieu Clinique. *The Canadian Journal of Program Evaluation* 17(3), 57-74 versión On-Line. En www.evaluationcanada.ca/secure/17-3-057.pdf
- Garrido, L., Santelices, M. P., & Pierrehumbert, B. (2009). Validación chilena del cuestionario de evaluación de apego en el adulto CAMIE. *Revista Latinoamericana de Psicología*. 41(1), 81-98.
- Gross, R., Glasser, S., Elisha, D., Tishby, O., Madar, J. D., Levitan, G., & Lambert, M. J. (2014). Validation of the Hebrew and Arabic Versions of the Outcome Questionnaire (OQ-45). *The Israel Journal of Psychiatry and Related Sciences*, 52(1), 33-39.
- Gullo, S., Lo Coco, G., & Gelso, Ch. (2012). Early and later predictors of outcome in brief therapy: the role of real relationship. *Journal of Clinical Psychology*, 68, 614-619.
- Holt, H., Beutler, L., Kinpara, S., Macías, S., Haug, S., Shiloff, N., Goldblum, P., Temkin, R., & Stein, M. (2015). Evidence-based Supervision Tracking Outcome and Teaching Principles of change in Clinical supervision to bring science to integrative practice. *Psychotherapy*, 52, 185-189.
- Jong, K., Nugter, M. A., Polak, M. G., Wagenhorg, J. E., Spinhoven, P., & Heiser, W. J. (2007). The Outcome Questionnaire (OQ-45) in a Dutch population: A cross-cultural validation. *Clinical Psychology & Psychotherapy*, 14(4), 288-301.
- Lambert, M. (2015). Outcome Research: methods for improving outcome in routine care. En: O, Gelo; A. Pritz, & B. Rieken. *Psychotherapy Research: Foundations, process and outcome*. Viena: Springer Verlag.
- Lambert, M. J., Lunnen, K., Umphress, V., Hansen, N. B., & Burlingame, G. M. (1996). Administration and scoring manual for the Outcome Questionnaire (OQ – 45.1) Salt Lake: IHC Center for Behavioral Health Care Efficacy.
- Lara, C., Cruz, C., Vacarezza, A., Florenzano, R., & Trapp, A. (2008) Análisis comparativo de dos instrumentos de evaluación clínica: OQ-45 e InterRAI- Salud Mental Chile, versión On-line. En http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-92272008000300004
- Machado, P.P., & Fassnacht, D. B. (2014). The Portuguese version of the Outcome Questionnaire (OQ-45): Normative data, reliability, and clinical significance cut-offs scores. *Psychology and Psychotherapy: Theory, Research and Practice*. 88(4), 427-37.
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eximan, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological Testing and Psychological Assessment: A review of Evidence and Issues. *American Psychologist*, 53(2), 128-165.
- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The Outcome rating scale: a preliminary study of the reliability, validity, and feasibility of a Brief visual Analog Measures. *Journal of Brief Therapy*, 2(2), 91-100.
- Painepán B., & Kühne W. (2012). Efectividad según duración de la psicoterapia en un centro de atención psicológica para universitarios. Departamento de Promoción de la Salud Psicológica. Universidad de Santiago de Chile. On-Line en <http://www.kuhne.cl/documentos/efectividad-segun-duracion-de-la-psicoterapia.pdf>
- Santibáñez, P. M., Román, M. F., & Vinet, E. V. (2009). Efectividad de la psicoterapia y su relación con la alianza terapéutica. *Interdisciplinaria*, 26, versión On-Line. En http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1668-70272009000200006, recuperado el 1 de diciembre del 2015.
- Shulte, D. (1977). Dimensions of outcome measurement. En H. H. Strupp: L. M. Horowitz y M. J. Lambert (Ed.). *Measuring patient changes in mood, anxiety and personality disorders: Towards a core battery*. Washington: APA.

- Takara, R., Beecher, M. E., Okiishi, J. C., Shimokawa, K., Lambert, M. J., & Griner, D. (2015). Translation of the Outcome Questionnaire-45 (OQ) into Japanese: A cultural adaptation. *Psychotherapy Research, 11*, 1-13. doi: 10.1080/10503307.2015.1080876
- Timman, R., Jong, K., & Neve-Enthoven, N. (2015). Cut-off Scores and Clinical Change Indices for the Dutch Outcome Questionnaire (OQ-45) in a Large Sample of Normal and Several Psychotherapeutic Populations. *Clinical psychology & Psychotherapy*. doi: 10.1002/cpp.1979