

Estructura dimensional de la Escala Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) en universitarios mexicanos

Dimensional Structure of the Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM) in Mexican University Students

Introducción
Método
Resultados
Discusión
Referencias

Janett Esmeralda Sosa-Torralba *¹, Martha Patricia Romero-Mendoza²,
María Elena Medina-Mora-Icaza², Enrique Méndez-Ríos², Bertha Blum-Grynberg¹

1- Facultad de Psicología, Universidad Nacional Autónoma de México.

2- Dirección de Investigaciones Epidemiológicas y Sociales, Instituto Nacional de Psiquiatría "Ramón de la Fuente Muñiz", México.

Recibido: 17/09/2019 Revisado: 11/11/2019 Aceptado: 03/12/2019

Resumen

Este trabajo verifica la estructura dimensional del Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM; Evans et al., 2000) mediante un estudio cuantitativo-transversal en una muestra de 466 estudiantes universitarios elegidos por muestreo de conveniencia; 181 de 18 a 30 años que solicitaron atención psicológica, y 285 de 17 a 29 años que no estaban en búsqueda de atención. A través de los análisis factoriales exploratorio y confirmatorio se identificaron tres dimensiones para el instrumento con adecuados niveles de consistencia interna, lo cual replica sustancialmente los resultados previos de la versión original. Se hallaron diferencias estadísticamente significativas entre varones y mujeres y entre la población consultante y no consultante. Finalmente, se comprobó si existía funcionamiento diferencial del ítem por sexo y por población consultante. Cinco reactivos se hallaron en la primera categoría y tres en la segunda. La estructura interna del CORE-OM explora un amplio conjunto de síntomas de malestar psicológico y resulta útil para su evaluación en jóvenes universitarios.

Palabras clave: CORE-OM, malestar psicológico, universitarios, AFC, DIF

Abstract

This work verifies the dimensional structure of the Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM; Evans et al., 2000) through a quantitative-cross-sectional study in a sample of 466 university students who were selected by convenience sampling. 181 ages 18 to 30 who had requested psychological attention; and 285 ages 17 to 29 who were not seeking care. Through exploratory and confirmatory factor analyses, three dimensions for the instrument were identified, which presented adequate levels of internal consistency. This substantially replicates the previous results of the original version. Statistically significant differences were found in relation to sex, as well as of the consulting and non-consulting population. Differential item functioning (DIF) was carried out to explore these differences, and five items were found in the first category and three in the second. The internal structure of CORE-OM explores a wide range of symptoms of psychological distress and it is useful for evaluation in young university students.

Keywords: CORE-OM, psychological distress, university, CFA, DIF

*Correspondencia a: Dra. Janett Esmeralda Sosa Torralba. Facultad de Psicología, UNAM. Avenida Universidad 3004, Cubículo 2. Edificio D. Copilco Universidad, 04510, Ciudad de México. Tel: 56220555, ext. 41242. E-mail: jan_est@comunidad.unam.mx

Cómo citar este artículo: Sosa-Torralba, J. E., Romero-Mendoza, M. P., Medina-Mora-Icaza, M. E., Méndez-Ríos, E., & Blum-Grynberg, B. (2020). Estructura dimensional de la Escala Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM) en estudiantes universitarios mexicanos. *Revista Evaluar*, 20(1), 46-63. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>

Introducción

En México, la salud integral de la población estudiantil de adolescentes y jóvenes adultos ha sido reconocida como un tema preponderante en instituciones educativas (López-Bárcena & González de Cossío-Ortiz, 2005). En el caso de los jóvenes, se han detectado síntomas somáticos, de depresión, ansiedad, tristeza, soledad; conductas de riesgo, e incluso ideación suicida (Guevara-Ruiseñor, 2008; Pérez-Padilla, Ponce-Rojo, Hernández-Contreras, & Márquez-Muñoz, 2010; Rivera-Ledesma, Caballero-Suárez, Pérez-Sánchez, & Montero-López-Lena, 2013; Rosales-Pérez, Córdova-Osnaya, & Ramos-Clatempa, 2012). Ante estas situaciones, se ha hecho énfasis en identificar el malestar emocional en los jóvenes, así como en detectar las áreas de atención para dar solución a las consecuencias de este (Guevara-Ruiseñor, 2008; Pérez-Padilla et al., 2010).

El malestar emocional se identifica por sus manifestaciones en el estado de ánimo, el tono afectivo, la percepción y los pensamientos, que alteran la funcionalidad y las representaciones que conforman el mundo interno del sujeto (Espíndola-Hernández et al., 2006). Con el propósito de su evaluación se han diseñado varias escalas, inventarios e instrumentos, uno de ellos es el CORE-OM (Clinical Outcomes in Routine Evaluation-Outcome Measure; Evans et al., 2000). Fue presentado por el CORE System Group en 1998 en el Reino Unido como un medio estandarizado y viable para la evaluación de los resultados de los servicios de terapias psicológicas, dentro del marco de la práctica basada en la evidencia (Mellor-Clark, Barkham, Connel, & Evans, 1999). Evalúa áreas de síntomas, funcionamiento, bienestar subjetivo y riesgo hacia sí mismo y

hacia otros, por lo que es considerado adecuado como herramienta de detección inicial (Margison et al., 2000; Mellor-Clark et al., 1999) en servicios de atención primaria y secundaria de la salud (Barkham, Gilbert, Connell, Marshall, & Twigg, 2005; Beck, Burdett, & Lewis, 2015; Navascués, Calvo-Medel, & Bombin-Martín, 2016).

Este instrumento ha sido evaluado, traducido y aplicado a diferentes poblaciones. En el país donde fue desarrollado, se han valorado sus propiedades psicométricas y su utilidad clínica en pacientes con diferentes diagnósticos y en diferentes escenarios terapéuticos; por ejemplo, en estudiantes universitarios, pacientes con diagnóstico psiquiátrico, pacientes en psicoterapia sexual, en ambientes hospitalarios y en intervenciones con foco en la pareja (Branney & Barkham, 2006; Casey et al., 2017; Connell, Barkham, & Mellor-Clarck, 2007; Perry, Barkham, & Evans, 2013).

En algunas de las traducciones realizadas se ha visto que es un medio útil, pertinente y adecuado para la valoración del malestar psicológico, tanto en la práctica clínica como en la investigación (Dias-Sales, de Matos-Moleiro, Evans, & Gomes-Alves, 2012; Elfström et al., 2013; Juhová et al., 2018; Uji, Sakamoto, Adachi, & Kitamura, 2012; Viliū-nienė et al., 2013; Zhang et al., 2019).

La versión castellana del CORE-OM en España ha mostrado propiedades psicométricas aceptables (Trujillo-Terán, 2012), y la versión validada en Costa Rica ha sido sugerida como confiable y aplicable para medir la efectividad psicoterapéutica (Vargas-Fallas & Prada-Villalobos, 2013). Sin embargo, en comparación con los análisis de la versión original (Evans et al., 2002), los análisis exploratorios para estas versiones han presentado variaciones en la estructura de sus componentes. Por otro lado, en ambas versio-

nes se hallaron diferencias a nivel de género y de población clínica y no clínica, motivo por el cual se ha considerado necesario tomar en cuenta estas variables cuando se interpretan puntuaciones individuales en cada una de las dimensiones, en el caso de datos clínicos (Evans et al., 2002; Trujillo-Terán, 2012).

En México no se ha documentado investigación acerca de este instrumento, y dada la necesidad de contar con herramientas válidas y confiables que permitan medir el malestar psicológico en los jóvenes, el objetivo del presente estudio se centró en verificar la estructura interna del CORE-OM para identificar la complejidad del constructo, así como para obtener evidencias de validez y confiabilidad del instrumento que avalen su uso en el contexto universitario. Esta herramienta permitiría conocer las manifestaciones que experimentan los estudiantes con malestar emocional con el propósito de emprender programas de tratamiento que permitan mejorar su calidad de vida.

Método

Participantes

Se trata de un estudio cuantitativo, transversal-descriptivo (Echevarría, 2016), que incluyó a 466 estudiantes universitarios de la Universidad Nacional Autónoma de México (UNAM), seleccionados mediante muestreo por conveniencia (Otzen & Manterola, 2017). El criterio de inclusión considerado fue que pertenecieran al Campus de Ciudad Universitaria: 181 buscaron atención en centros de servicios psicológicos de la misma Universidad, los otros 285 alumnos no buscaban atención psicológica al momento de responder al instrumento. Los estudiantes no consultantes te-

nían entre 18 y 30 años ($M = 21$, $DE = 2$), 40% eran hombres, 58% mujeres y el 2% no proporcionó esa información. Aquellos que acudieron a solicitar atención tenían entre 17 y 29 años ($M = 23$, $DE = 3.8$), 61% eran mujeres, 38% hombres y 2 personas (1%) no dieron este dato.

Instrumento

CORE-OM. Se utilizó la versión en castellano que se encuentra disponible en la página web del CORE Information Management Systems disponible para hombres y mujeres (CORE IMS, 2001). Incluye 34 reactivos de autoaplicación que exploran 4 subescalas: 1) *bienestar subjetivo* (4 ítems), 2) *síntomas/problemas* (12 ítems que evalúan ansiedad, depresión, problemas físicos y trauma), 3) *funcionamiento* (12 ítems, incluyen funcionamiento general, en relaciones cercanas y en relaciones sociales), y 4) *riesgo* (6 ítems de daño hacia sí mismo y actos de agresión hacia otros). En la Tabla 1 se muestran los reactivos correspondientes a cada una de estas subescalas.

Cada ítem cuestiona la frecuencia de malestar durante los últimos siete días y se valora en una escala de 5 puntos tipo Likert según la respuesta elegida (0 = *Nunca*, 4 = *Siempre o casi siempre*). La mayoría de los reactivos reflejan reacciones negativas. Los reactivos 3, 4, 7, 12, 19, 21, 31 y 32 reflejan reacciones positivas de la persona; en consecuencia, la puntuación que se obtiene de estos ítems está invertida. De este modo, también en la escala de bienestar subjetivo mayores puntuaciones indican un mayor nivel de malestar.

Se obtienen calificaciones por subescala y de la prueba total; las puntuaciones, en ambos casos, varían de 0 a 40. La escala está diseñada de tal manera que las calificaciones más altas indi-

can mayores problemas o síntomas. Para la puntuación total del instrumento, también se considera que se sumen todas las subescalas menos la de Riesgo (T-R), por ser considerada como anexa al instrumento, y cualquiera de sus ítems que puntúe por encima de 0 es utilizado como indicador clínico de presencia de ideación suicida, autolesiones o actos de agresión a terceros.

El análisis de consistencia interna, tanto para el total de la escala como para cada una de las dimensiones, ha generado valores de entre $\alpha = .6$ y $\alpha = .9$ en la versión original y en las versiones en español (Evans et al., 2002; Trujillo-Terán, 2012; Vargas-Fallas & Prada-Villalobos, 2013).

Procedimiento

El CORE-OM fue evaluado en un análisis de contenido por ocho expertos cuya trayectoria se caracteriza por su experiencia clínica en salud mental (siete psicólogos y un psiquiatra). Mediante un cuestionario escrito cada juez valoró cualitativamente el grado de pertinencia de los ítems con respecto al constructo de malestar emocional. Los aportes cualitativos de los jueces indicaron que la escala era apropiada para su propósito de evaluación. Posteriormente, el instrumento se aplicó en dependencias de la Universidad a los estudiantes no consultantes. Asimismo, se pidió autorización en dos Centros de Atención Psicológica y se aplicó a universitarios que llegaron a solicitar el servicio. En ambos grupos se entregaba el instrumento a los sujetos, siguiendo la distinción de la versión para hombres y para mujeres, y era completado en presencia del aplicador para atender dudas o problemas que pudieran presentarse.

Análisis de los datos

Los datos se analizaron en tres fases consecutivas. Para las dos primeras fases se utilizó el programa estadístico SPSS Statistics 21.0 (IBM Corporation, 2012). En la primera fase se analizaron los estadísticos descriptivos de los ítems (media, desviación estándar, asimetría y curtosis). Con el fin de analizar la estructura interna del CORE-OM, se realizó un análisis factorial exploratorio (AFE). Ante el incumplimiento de la normalidad de algunos de los ítems se utilizó el método de extracción de mínimos cuadrados no ponderados, ya que este método no hace suposiciones acerca de la distribución y es adecuado para las variables ordinales (Forero, Maydeu-Olivares, & Gallardo-Pujol, 2009), y rotación oblicua (promax) por ser adecuada para factores que están relacionados (Campo-Arias, Herazo, & Oviedo, 2012). Para la extracción del número de factores, aparte de tener en cuenta los resultados de los estudios previos de la versión original (Evans et al., 2002; Lyne, Baret, Evans, & Barkham, 2006), se consideraron criterios teóricos y metodológicos (Campo-Arias et al., 2012): el primer criterio parte del supuesto de que un factor es importante si lo forman entre tres y cinco reactivos. El segundo fue aplicar el criterio de Kaiser (valor *eigen* superior a la unidad) y el criterio de Gorsuch o de interpretabilidad de las diferentes estructuras factoriales resultantes. El tercer criterio establecía que el modelo factorial garantizara la máxima explicación de varianza sin perder validez de contenido y que los ítems saturaran en su factor de pertenencia con valores que alcanzaran coeficientes iguales o superiores a .30. Se procedió con múltiples análisis factoriales exploratorios, entre los que se ajustó el número de factores a extraer para encontrar una estructura de factores consistente.

En la segunda fase se realizó un análisis factorial confirmatorio (AFC) a través del programa AMOS (integrado en el SPSS 21.0; [IBM Corporation, 2012](#)) para ratificar el modelo obtenido en el exploratorio. Se contrastaron tres modelos diferentes de la adecuación de las estructuras dimensionales propuestas de la versión original: modelo de dos factores, un componente general de malestar psicológico que incluye el conjunto de reactivos con reacciones negativas y positivas, y otro componente de riesgo (modelo 1; [Lyne et al., 2006](#)); modelo de tres factores, un componente de los reactivos de reacciones negativas, otro grupo de reactivos de reacciones positivas, y un último componente que agrupa los reactivos de riesgo (modelo 2; [Evans et al., 2002](#)); modelo de cuatro factores, buscando coincidencia con las subescalas del instrumento ([CORE IMS, 2001](#)).

Para el AFC, el método de estimación utilizado fue el de mínimos cuadrados no ponderados, ya que al indagar la normalidad multivariante se encontraron valores de curtosis multivariada superiores a 5.99 ($CR = 242.463$). Para la determinación del ajuste del modelo, tal como es recomendado, se evaluaron varios índices de ajuste ([Byrne, 2009](#)). Se utilizaron: el estadístico χ^2 , el índice de bondad de ajuste (GFI), el índice corregido de bondad de ajuste (AGFI), el índice de ajuste normativo (NFI) y la raíz cuadrada de los residuos al cuadrado (RMR). En el caso del GFI, AGFI y NFI se consideraron aceptables valores superiores a .90 ([Bentler, 1992](#); [Byrne, 2009](#)). Para la RMR se consideraron valores aceptables en el rango de menores o iguales que .05 ([Byrne, 2009](#)).

Se realizó un análisis de correlaciones entre las dimensiones resultantes, utilizando el coeficiente de correlación de Spearman. La consistencia interna fue analizada mediante alfa de Cronbach. Se analizaron diferencias de sexo y entre

quienes buscaron y quienes no buscaron atención psicológica mediante la prueba *t* de Student.

En la tercera fase se examinó el funcionamiento diferencial de los ítems (DIF) para analizar posibles sesgos producidos por el efecto del sexo y por el efecto de la población consultante de atención psicológica. Este estadístico trata de comparar la ejecución de un ítem entre el grupo de referencia y el grupo focal a través de distintos niveles de un determinado criterio de equiparación, asumiendo que en cada nivel los sujetos de uno y otro grupo son comparables, por lo que si no hay DIF ambos lo ejecutarán por igual. Un ítem presenta DIF cuando la probabilidad de ser resuelto correctamente por individuos con el mismo nivel de rasgo varía en función de su grupo de pertenencia (sexo, nivel socioeconómico, etc.) que genera una falta de equivalencia métrica entre las puntuaciones ([Elosua-Oliden, 2003](#)). En este trabajo se empleó el test generalizado de Mantel-Haenszel (GMH; [Mantel & Haenszel, 1959](#)). Para el análisis por sexo se consideró como grupo de referencia a las mujeres. Para el análisis por población consultante se consideró como grupo de referencia a quienes buscaron atención psicológica.

Este análisis estadístico se implementó mediante dos pruebas de significancia: en primer lugar, la prueba $M - \chi^2$ con $gl = 1$, para ítems politómicos. Un resultado $M - \chi^2 > 3.84$ detecta un DIF en el nivel $p < .05$, utilizando para el presente estudio 6.63 en el nivel $p < .01$. En segundo lugar, se estimó la significancia práctica del DIF usando los resultados obtenidos con el estimador de Liu-Agresti (LA LOR; [Penfield, 2013](#)). La magnitud del DIF para ítems politómicos mediante LA LOR se divide en los niveles de insignificante ($< .43$), moderado (entre $.43$ y $.64$) y grande ($\geq .64$; [Penfield & Algina, 2006](#)). Los valores positivos indican DIF a favor del grupo de referencia y los

valores negativos indican DIF a favor del grupo focal (Penfield, 2013). De igual manera, se aplicó la estimación estandarizada del índice común de probabilidad acumulativa de LA LOR (LOR Z), en el cual valores más grandes a 2.0 o menores a -2.0 son considerados como evidencia de DIF (Penfield, 2013). Para este estadístico se utilizó el programa informático DIFAS 5.0 (Penfield, 2009).

Consideraciones éticas

El protocolo de investigación fue aprobado por el Comité de Ética del Programa de Maestría y Doctorado en Psicología de la Facultad de Psicología de la UNAM. A los participantes se les explicó el objetivo del estudio y se requirió su consentimiento informado.

Resultados

En la Tabla 1 se presentan los estadísticos descriptivos de centralidad, dispersión y normalidad para los 34 ítems de la escala CORE-OM en la muestra total, referidos a la media, desviación estándar, asimetría y curtosis. Como se puede apreciar, las medias más bajas corresponden a los ítems pertenecientes a indicadores de riesgo de daño hacia sí mismo, el ítem 16 (*He hecho planes para acabar con mi vida*) es el de menor puntuación ($M = .19$). Respecto a la normalidad de las variables, cinco de los ítems presentan índices de asimetría y curtosis mayores al rango ± 1.5 (reactivos 9, 16, 22, 24 y 34), lo cual se considera inadecuado para el supuesto de normalidad (George & Mallery, 2001, como se citó en Pérez & Medrano, 2010).

Análisis factorial exploratorio

Puesto que todavía no ha sido referido ningún estudio factorial del CORE-OM en población mexicana, pareció conveniente realizar un análisis factorial de tipo exploratorio (AFE). La adecuación muestral de Kaiser-Meyer-Olkin (KMO) mostró un valor de .95 y el test de esfericidad de Bartlett resultó estadísticamente significativo ($\chi^2 = 7562.562$, $gl = 561$, $p < .001$); en consecuencia, la aplicación del análisis factorial resultaba pertinente.

Tras el análisis, se obtuvo una estructura dimensional de tres factores con capacidad para explicar el 48% de la varianza total. Los reactivos *He sido violento(a) físicamente hacia los demás* (ítem 6) y *He amenazado e intimidado a otra persona* (ítem 22) tuvieron cargas factoriales menores a .30 con comunalidades de .10. Los demás ítems alcanzaron pesos factoriales de entre .41 y .80, y comunalidades de entre .19 y .64. En la Tabla 2, se muestran la matriz factorial rotada, las saturaciones de los pesos factoriales y los valores de comunalidad.

Esta estructura de tres dimensiones, al igual que la distribución de los ítems, coincidió parcialmente con la propuesta de tres factores de la versión inglesa (Evans et al., 2002). La dimensión de mayor peso factorial fue donde se agruparon los ítems que reflejan reacciones negativas e indican malestar psicológico. Esta dimensión explica el 36% de la varianza total. A continuación, le sigue la dimensión que agrupó los reactivos de reacciones positivas relacionados con el bienestar psicológico, y explica el 7% de la varianza. La tercera dimensión, que hace referencia al riesgo hacia sí mismo y hacia otros, es la que menos varianza explica, con un 5%.

Tabla 1

Análisis descriptivo de los ítems de la versión en castellano del CORE-OM

Ítems	M	DE	Asimetría	Curtosis
1. Me he sentido muy solo(a) y aislado(a) (F-rc)	1.33	1.10	0.51	-0.40
2. Me he sentido tenso(a), ansioso(a) o nervioso(a) (S-a)	2.34	0.98	-0.39	-0.11
3. He sentido que tenía a alguien en quien apoyarme cuando lo he necesitado (F-rc)	1.24	1.14	0.65	-0.41
4. Me he sentido bien conmigo mismo(a) (B)	1.17	1.02	0.62	-0.24
5. Me he sentido totalmente sin energía ni entusiasmo (S-d)	1.64	1.08	0.27	-0.61
6. He sido violento(a) físicamente hacia los demás (R-o)	0.61	0.90	1.38	1.02
7. Me he sentido capaz de afrontar las cosas cuando han ido mal (F-g)	1.20	1.00	0.60	-0.18
8. He tenido molestias, dolores y otros problemas físicos (S-f)	1.59	1.09	0.19	-0.60
9. He pensado en hacerme daño a mí mismo(a) (R-s)	0.30	0.73	2.82	8.17
10. Me ha costado mucho hablar con la gente (F-rs)	1.26	1.16	0.66	-0.39
11. La tensión y la ansiedad me han impedido hacer cosas importantes (S-a)	1.24	1.17	0.62	-0.61
12. Me he sentido satisfecho(a) con las cosas que he hecho (F-g)	1.19	0.99	0.58	-0.19
13. Me han inquietado pensamientos y sentimientos no deseados (S-t)	1.61	1.17	0.21	-0.89
14. He tenido ganas de llorar (B)	1.80	1.15	0.02	-0.85
15. He sentido pánico o terror (S-a)	1.00	1.05	0.85	0.01
16. He hecho planes para acabar con mi vida (R-s)	0.19	0.62	4.03	18.11
17. Me he sentido agobiado(a) por mis problemas (B)	1.56	1.15	0.26	-0.84
18. He tenido dificultad para conciliar el sueño o permanecer dormido(a) (S-f)	1.39	1.26	0.47	-0.91
19. He sentido afecto o cariño por alguien (F-rc)	0.77	0.91	1.22	1.26
20. Me ha sido imposible dejar a un lado mis problemas (S-a)	1.53	1.15	0.47	-0.53
21. He sido capaz de hacer la mayoría de las cosas que tenía que hacer (F-g)	1.24	1.09	0.63	-0.45
22. He amenazado o intimidado a otra persona (R-o)	0.32	0.70	2.36	5.55
23. Me he sentido desesperado(a) o sin esperanza (S-d)	1.05	1.16	0.89	-0.18
24. He pensado que sería mejor que estuviera muerto(a) (R-s)	0.36	0.81	2.52	6.19
25. Me he sentido criticado(a) por los demás (F-rs)	1.46	1.12	0.54	-0.45
26. He pensado que no tengo amigos (F-rc)	0.84	1.12	1.19	0.45
27. Me he sentido infeliz (S-d)	1.16	1.17	0.75	-0.40
28. Me han angustiado imágenes o recuerdos no deseados (S-t)	1.41	1.24	0.48	-0.86
29. Me he sentido irritable cuando estaba con otras personas (F-rs)	1.52	1.08	0.19	-0.81
30. He pensado que yo tengo la culpa de mis problemas y dificultades (S-d)	1.93	1.19	1.25	-0.94
31. Me he sentido optimista sobre de mi futuro (B)	1.14	1.10	0.69	-0.39
32. He conseguido las cosas que quería (F-g)	1.38	0.97	0.41	-0.25
33. Me he sentido humillado(a) o avergonzado(a) por otras personas (F-rs)	0.86	1.03	1.19	0.76
34. Me he hecho daño físicamente o he puesto en peligro mi salud (R-s)	0.22	0.63	3.50	13.31

Nota. B = Bienestar; F-g = Funcionamiento general, F-rc = Funcionamiento en relaciones cercanas, F-rs = Funcionamiento en relaciones sociales; S-a = Síntomas/problemas ansiedad, S-d = Síntomas/problemas depresión, S-f = Síntomas/problemas físicos, S-t = Síntomas/problemas trauma; R-o = Riesgo de actos de agresión hacia otros, R-s = Riesgo de daño hacia sí mismo.

Tabla 2

Matriz de cargas factoriales, comunalidades y varianza explicada de cada uno de los factores.

	Ítems	Factores			Comunalidad
		1	2	3	
17	Me he sentido agobiado(a) por mis problemas (B)	.78			.62
28	Me han angustiada(a) imágenes o recuerdos no deseados (S)	.73			.54
23	Me he sentido desesperado(a) o sin esperanza (S)	.73			.59
11	La tensión y la ansiedad me han impedido hacer cosas importantes (S)	.72			.53
33	Me he sentido humillado(a) o avergonzado(a) por otras personas	.68			.47
13	Me han inquietado pensamientos y sentimientos no deseados (S)	.68			.47
15	He sentido pánico o terror (S)	.68			.47
1	Me he sentido muy solo(a) y aislado(a) (F)	.66			.47
14	He tenido ganas de llorar (B)	.66			.43
25	Me he sentido criticado(a) por los demás (F)	.65			.44
5	Me he sentido totalmente sin energía ni entusiasmo (S)	.64			.43
27	Me he sentido infeliz (S)	.64			.42
29	Me he sentido irritable cuando estaba con otras personas (F)	.64			.41
20	Me ha sido imposible dejar a un lado mis problemas (S)	.62			.39
26	He pensado que no tengo amigos (F)	.60			.40
30	He pensado que yo tengo la culpa de mis problemas y dificultades (S)	.57			.33
10	Me ha costado mucho hablar con la gente (F)	.56			.33
2	Me he sentido tenso(a), ansioso(a) o nervioso(a) (S)	.56			.34
18	He tenido dificultad para conciliar el sueño o permanecer dormido(a) (S)	.55			.30
8	He tenido molestias, dolores y otros problemas físicos (S)	.41			.19
12	Me he sentido satisfecho(a) con las cosas que he hecho (F)		.80		.64
31	Me he sentido optimista sobre mi futuro (F)		.78		.63
32	He conseguido las cosas que quería (F)		.78		.62
4	Me he sentido bien conmigo mismo(a) (B)		.76		.60
7	Me he sentido capaz de afrontar las cosas cuando han ido mal (F)		.69		.48
21	He sido capaz de hacer la mayoría de las cosas que tenía que hacer (F)		.57		.33
3	He sentido que tenía a alguien en quien apoyarme cuando lo he necesitado (F)		.45		.23
19	He sentido afecto o cariño por alguien (F)		.43		.21
9	He pensado en hacerme daño a mí mismo(a) (R)			.80	.64
24	He pensado que sería mejor que estuviera muerto(a) (R)			.78	.63
16	He hecho planes para acabar con mi vida (R)			.67	.47
34	Me he hecho daño físicamente o he puesto en peligro mi salud (R)			.63	.42
22	He amenazado o intimidado a otra persona (R)			.29	.10
6	He sido violento(a) físicamente hacia los demás (R)			.25	.10
	Varianza explicada	36%	7%	5%	

Nota. B = Bienestar, F = Funcionamiento, S = Síntomas/problemas, R = Riesgo.

Análisis factorial confirmatorio

Posteriormente, se realizó un análisis factorial confirmatorio (AFC) de los ítems mediante mínimos cuadrados no ponderados para ratificar el modelo que se obtuvo en el análisis exploratorio. Al retomar las propuestas de la estructura de la escala original, se realizaron comparaciones entre tres modelos de dos, tres y cuatro factores. Según puede observarse en la Tabla 3, en los modelos de dos y cuatro factores, los ajustes en el índice RMR están lejos de los criterios recomendados ($\leq .05$). El ajuste es más adecuado en el modelo 2 de tres factores, lo que confirma que es el que mejor se ajusta a los datos ($\chi^2_{(524)} = 1325.33$). Puede observarse que los valores de GFI, AGFI y NFI se sitúan por encima del valor de .90, mientras que el valor de RMR se encuentra en .05. Las cargas factoriales de los reactivos fueron superiores a .30. La representación gráfica del modelo, junto con los valores de interrelación entre las variables, se aprecia en la Figura 1.

En la Tabla 4 se recogen la media, la desviación estándar y el valor alfa de Cronbach para las dimensiones resultantes del CORE - OM, junto con las correlaciones entre estas, las diferencias de género y las diferencias entre consultantes y no consultantes. La consistencia interna fue de .93 para la primera dimensión, .86 para la segunda, .74 para la tercera y .94 para la escala total. La media para la dimensión de malestar emocional fue de 28.49 (DE = 15.05), para bienestar emocional de 9.31 (DE = 5.86) y para riesgo de 2.0 (DE = 2.92).

Tabla 3

Índices de ajuste de los modelos factoriales probados para la Escala CORE-OM.

Modelo factorial	χ^2	gl	GFI	AGFI	RMR	NFI
Modelo 1 (2 factores)	1655.20 ^a	526	.97	.97	.07	.97
Modelo 2 (3 factores)	1325.33 ^a	524	.98	.98	.05	.98
Modelo 3 (4 factores)	1838.43 ^a	521	.98	.97	.07	.97

Nota. GFI = Índice de bondad de ajuste; AGFI = Índice de bondad de ajuste corregido según los grados de libertad; RMR = Raíz cuadrada de los residuos al cuadrado; NFI = Índice de ajuste normativo. ^a: $p < .001$

La correlación r de Spearman mostró intercorrelaciones fuertes y positivas entre las tres dimensiones, con valores de entre .43 y .62 ($p \leq .01$).

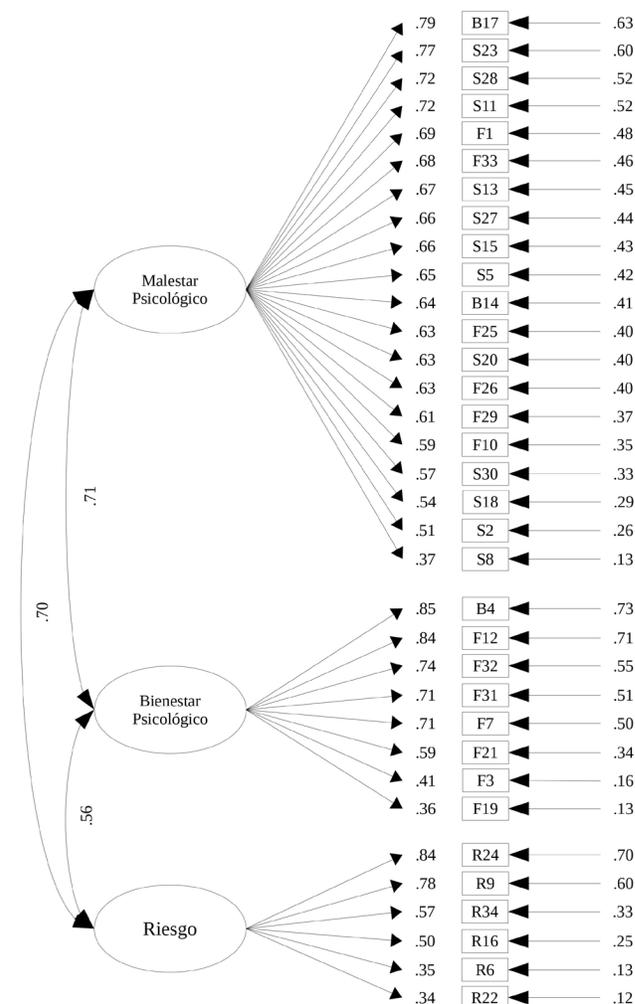


Figura 1

Modelo factorial confirmatorio de la escala CORE-OM.

Nota. B = Bienestar, F = Funcionamiento, S = Síntomas/problemas, R = Riesgo.

Tabla 4

Estadísticos descriptivos, valores de fiabilidad alfa de Cronbach, valores de correlación para las dimensiones del CORE-OM y valores *t*.

Dimensión	Sexo										Búsqueda de atención					
	Hombres					Mujeres					Buscaron			No buscaron		
M	DE	α	Bienestar emocional	Riesgo	M	DE	M	DE	<i>t</i>	M	DE	M	DE	<i>t</i>		
Malestar emocional	28.49	15.05	.93	.62**	.53**	25.54	14.74	30.43	15.00	-3.45**	36.64	15.94	23.33	11.87	-9.67**	
Bienestar emocional	9.31	5.86	.86	-	.43**	9.29	5.94	9.31	5.82	-0.32	11.99	5.91	7.60	5.15	-8.21**	
Riesgo	2.0	2.92	.74	-	-	1.98	2.96	2.00	2.92	-0.71	3.07	3.69	1.31	2.04	-5.86**	

Nota. ** $p \leq .01$

En el análisis de las diferencias de género, los resultados indican que las mujeres puntúan con una media mayor que los hombres en la dimensión de malestar emocional y que esta diferencia es estadísticamente significativa ($p \leq .01$), en tanto que los que buscaron atención puntuaron con una media significativamente mayor en todas las dimensiones con un nivel de significación de $p \leq .01$.

Análisis del funcionamiento diferencial de los ítems

En la Tabla 5 se muestran los resultados del análisis DIF para cada ítem. El análisis reveló que cinco ítems del CORE-OM mostraron un funcionamiento diferencial en función del género. Los ítems 2 (*Me he sentido tenso(a), ansioso(a) o nervioso(a)*), 14 (*He tenido ganas de llorar*) y 17 (*Me he sentido agobiado(a) por mis problemas*) favorecen al grupo focal, es decir, a los hombres. Los reactivos 19 (*He sentido afecto o cariño por alguien*) y 30 (*He pensado que yo tengo la culpa de mis problemas*) favorecieron a las mujeres (grupo de referencia), mientras que tres reactivos tuvieron funcionamiento diferencial en función de solicitar atención psicológica. Los reactivos 2 (*Me he sentido tenso(a), ansioso(a) y nervio-*

so(a)) y 5 (*Me he sentido totalmente sin energía y ni entusiasmo*) favorecieron a los consultantes (grupo de referencia) y el reactivo 4 (*Me he sentido bien conmigo mismo(a)*) a los no consultantes (grupo focal).

Discusión

Mediante el presente estudio, examinamos la estructura dimensional y otras propiedades psicométricas de la Escala de Resultados Clínicos de Evaluación Psicológica (CORE-OM) en población mexicana. Esta escala no solo valora el malestar psicológico, sino que fue desarrollada con el objetivo de evaluar los resultados de las psicoterapias. Un instrumento con adecuadas propiedades psicométricas que provea una interpretación teórica coherente en un contexto de uso bien delimitado permite su utilización más efectiva en la práctica clínica y en la investigación psicopatológica. Dada la importancia de lo anterior, se examinó la estructura interna del CORE-OM, se estimó la fiabilidad de las puntuaciones y se obtuvieron diferentes evidencias de validez a partir de una muestra de estudiantes universitarios. Este objetivo permite actualizar las propiedades métricas del CORE-OM para su uso en población de jóvenes, población perteneciente a un grupo de

Tabla 5
Resultados del funcionamiento diferencial del ítem (DIF).

	Sexo			Solicitó atención		
	$M - X^2$	LOR_{L-A}	Z_{LOR}	$M - X^2$	LOR_{L-A}	Z_{LOR}
<i>Malestar emocional</i>						
Ítem 1	0.69	-0.17	-0.78	0.01	0.02	0.09
Ítem 2	14.91	-0.83	-3.39	6.70	0.62	2.61
Ítem 5	4.59	-0.45	-2.04	10.35	0.77	3.29
Ítem 8	6.41	-0.52	-2.43	6.20	0.56	2.45
Ítem 10	0.03	0.04	0.18	1.77	0.3	1.28
Ítem 11	0.17	-0.1	-0.43	0.22	-0.11	0.49
Ítem 13	0.34	0.12	0.57	0.03	0.04	0.18
Ítem 14	56.02	-1.72	-6.96	1.24	-0.25	-1.12
Ítem 15	0.43	-0.15	-0.70	1.54	0.29	1.32
Ítem 17	13.95	-0.81	-3.59	0.05	-0.05	-0.24
Ítem 18	0.07	-0.05	-0.25	0.40	-0.14	-0.64
Ítem 20	0.29	-0.12	-0.55	0.08	0.07	0.29
Ítem 23	0.26	-0.12	-0.53	0.07	-0.06	-0.26
Ítem 25	0.53	0.15	0.71	0.45	-0.15	-0.67
Ítem 26	0.19	-0.10	-0.44	4.84	-0.57	-2.24
Ítem 27	2.82	0.39	1.72	0.50	0.17	0.78
Ítem 28	0.08	-0.06	-0.29	0.52	-0.16	-0.73
Ítem 29	1.68	-0.27	-1.31	1.63	-0.29	-1.31
Ítem 30	9.41	0.63	3.08	0.22	-0.10	-0.48
Ítem 33	4.02	0.45	1.90	2.69	-0.39	-1.58
<i>Bienestar emocional</i>						
Ítem 3	5.11	0.49	2.44	0.14	-0.08	-0.38
Ítem 4	0.09	-0.06	-0.31	8.27	-0.70	-2.88
Ítem 7	0.00	0.01	0.05	3.46	-0.46	-1.88
Ítem 12	1.09	0.24	1.11	0.38	-0.14	-0.61
Ítem 19	15.02	0.86	3.82	0.35	0.14	0.56
Ítem 21	3.62	0.41	1.87	0.29	0.12	0.56
Ítem 31	3.54	0.41	1.89	1.13	0.24	1.04
Ítem 32	4.04	0.42	1.95	2.13	0.35	1.46
<i>Riesgo</i>						
Ítem 6	0.32	0.13	0.56	0.01	-0.02	-0.08
Ítem 9	0.00	0.01	0.05	0.56	-0.29	-0.78
Ítem 16	1.81	0.58	1.33	0.48	0.33	0.84
Ítem 22	1.50	0.38	1.22	0.00	0.01	0.04
Ítem 24	0.70	0.32	0.89	0	-0.00	-0.01
Ítem 34	0.26	0.22	0.56	0.28	0.24	0.52

Nota. $M - X^2$: Chi-cuadrado de Mantel, LOR_{L-A} : Índice de probabilidad acumulativa de Liu-Agresti de Liu y Agresti, Z_{LOR} : Transformación Z de LOR_{L-A}

edad de especial vulnerabilidad en su salud mental (Auerbach et al., 2016).

Procediendo con análisis factorial exploratorio y confirmatorio, los datos muestran que la estructura de tres factores es la que mejor se ajusta para el CORE-OM y es similar a la sugerida en la versión original (Evans et al., 2002). Este modelo está justificado psicométricamente y desde una perspectiva teórica y de congruencia de significado y contenido. El primer factor hace referencia a lo que se ha definido como *malestar psicológico*, o las alteraciones psíquicas resultantes de la interpretación que la persona hace del mundo y su manera particular de resolver un conflicto externo y/o interno (Espíndola-Hernández et al., 2006). El segundo factor puede interpretarse como indicadores de *bienestar psicológico*, es decir, de la expresión del desarrollo de las capacidades y el crecimiento personal en un sentido de funcionamiento positivo (Díaz et al., 2006). Y el tercer factor hace referencia a los pensamientos, conductas o planeaciones de daño hacia sí mismo u otros, es decir, a *conductas de riesgo* (Vargas-Fallas & Prada-Villalobos, 2013).

Una debilidad psicométrica del instrumento consiste en la ambigüedad factorial en el AFE de los reactivos 6 (*He sido violento(a) físicamente hacia los demás*) y 22 (*He amenazado o intimidado a otra persona*); ambos ítems fueron elaborados para medir actos de agresión hacia otros en la subescala de riesgo. En este estudio, ambos saturaban con cargas factoriales de entre .20 y .30 tanto en el factor de malestar emocional como en el de riesgo. Estos resultados evidencian que, además de tener saturaciones pobres, estos reactivos son los que menos se relacionan con el conjunto de la dimensión de riesgo. En esta dimensión se exploran dos clasificaciones de problemas: el autodaño como un problema internalizado y el daño hacia otros como un problema externalizado. Esta

dualidad podría influir en los resultados obtenidos. Estos pesos factoriales sugieren hacer más análisis que ayuden a explicar la funcionalidad de estos ítems.

Respecto a la confiabilidad de la escala, considerando que en esta puede influir el número de reactivos del instrumento y la variabilidad de las personas a quienes se aplicó, la consistencia interna de las dimensiones y de la escala total ha sido adecuada con valores superiores a .70 (Campo-Arias & Oviedo, 2008), es decir, la escala muestra adecuada confiabilidad al aplicarla sobre esta muestra de jóvenes universitarios. Las correlaciones entre las dimensiones del CORE-OM resultaron significativas en todos los casos. Estos resultados muestran que a valores altos de una dimensión corresponden valores altos en la otra, valores de asociación que son lógicos y coherentes con el planteamiento teórico: un sujeto con malestar psicológico mostrará bajos niveles de bienestar.

En los resultados encontrados respecto a las diferencias estudiadas en las dimensiones en función del sexo, son las mujeres quienes informan un mayor malestar emocional. El análisis del DIF permite establecer si esas diferencias encontradas son parte legítima del constructo que se mide en el instrumento o, por el contrario, se deben a otras variables. Para comprender mejor la causa de sesgo en la respuesta de hombres y mujeres, cabe realizar un análisis de contenido de los ítems con DIF. Los ítems que favorecen a las mujeres plantean: *He sentido afecto o cariño por alguien*, *He pensado que yo tengo la culpa de mis problemas y dificultades*, los contenidos de estos hacen referencia a sintomatología que afecta el funcionamiento en relaciones cercanas, en el caso del primer reactivo, y a sintomatología depresiva en el segundo reactivo. Mientras que los ítems que favorecen a los hombres se refieren a: *Me he sen-*

tido tenso, ansioso o nervioso, He tenido ganas de llorar, Me he sentido agobiado por mis problemas, el primer reactivo hace referencia a sintomatología ansiosa y los otros dos a sintomatología que afecta el bienestar subjetivo. Dada la propia formulación de los ítems, donde independientemente del sexo de los jóvenes se les pregunta por su estado de malestar psicológico general, cabría esperar un comportamiento similar en hombres y mujeres. Sin embargo, debido a la infiltración de factores psicosociales de la cultura, puede suceder que las mujeres tiendan a puntuar más alto en cuestiones de externalización de síntomas de malestar emocional (Berenzon, Lara, Robles, & Medina-Mora, 2013). Por ejemplo, se ha visto que las mujeres informan experimentar prevalencias más altas de problemas de salud mental con síntomas más graves y discapacitantes (Freeman & Freeman, 2013), y en población mexicana, la sintomatología y los trastornos depresivos son más prevalentes en adolescentes (Benjet et al., 2009) y mujeres adultas (Fleiz-Bautista et al., 2012). Por el contrario, los hombres tienden a mostrar patrones internalizados de expresión emocional, siendo esta inhibida o restrictiva, lo que implica que sus emociones no son reconocidas o están enmascaradas por otras más aceptables por la condición de género (Jansz, 2000). Esto podría explicar el DIF hallado en esos reactivos. Se requiere estudiar a mayor detalle la manera en la que los hombres y las mujeres manifiestan su malestar emocional para, de esta manera, tener explicaciones más amplias sobre la funcionalidad de los ítems hallados con DIF.

Respecto a las diferencias halladas entre quienes solicitan atención y quienes no lo habían hecho hasta el momento de la aplicación del instrumento, son los consultantes quienes puntuaron más alto en todas las dimensiones, mostrando así que el malestar emocional está más presente en

aquellos que buscan atención psicológica. En relación con la evaluación del sesgo, el análisis del funcionamiento diferencial de los ítems sugiere diferencias existentes en los ítems *Me he sentido tenso(a), ansioso(a) o nervioso(a), Me he sentido totalmente sin energía ni entusiasmo y Me he sentido bien conmigo mismo(a)*. En este caso, parece lógico pensar que los jóvenes consultantes serán más propensos a presentar mayor sintomatología de ansiedad o depresión, que es lo que valoran los dos primeros reactivos, mientras que los jóvenes no consultantes puntuarán más alto en indicadores de bienestar subjetivo, que es lo que evalúa el tercer reactivo hallado con DIF.

Las diferencias encontradas tanto por sexo como por población consultante indican que el instrumento es capaz de discriminar entre grupos en los cuales se esperaba encontrar diferencias según lo publicado en los artículos de la versión original y la versión castellana (Evans et al., 2002; Trujillo-Terán, 2012), lo que confirma la generalidad del constructo bienestar como estado de la salud mental (Organización Mundial de la Salud, 2013). Así, el patrón de relación entre los indicadores del constructo fue similar, estando las medidas globales de este más relacionadas entre sí que las específicas. Los resultados de los análisis del funcionamiento diferencial ponen de manifiesto que, independientemente de que puedan existir diferencias reales en la expresión del malestar emocional entre hombres y mujeres, y entre población consultante y no consultante, es necesario realizar una revisión de los ítems que muestran un funcionamiento diferencial y complementarla con análisis cualitativos que exploren la expresión del malestar emocional en los distintos grupos comparados.

Los datos encontrados en este trabajo son preliminares y esclarecen evidencias de validez de estructura interna y confiabilidad del CORE-OM.

En la versión original, para la obtención de las características psicométricas y datos normativos de la escala, se incluyeron dos muestras: una población clínica compuesta por usuarios que se encontraban en la lista de espera o estaban ya en un proceso terapéutico en el área de salud mental, y una población no clínica, es decir, personas que no cumplían con las dos condiciones anteriores. En ambas muestras se encontraron tres factores; sin embargo, en la muestra no clínica los reactivos se agruparon de forma más clara con respecto a un componente de reactivos de reacciones negativas, otro de reacciones positivas y un último de reactivos que reflejan riesgo (Evans et al., 2002). En la validación española, se tomaron los resultados de ambas muestras en el análisis de la estructura del instrumento y se obtuvo un componente donde se situaron la mayoría de los reactivos de riesgo, y otros dos componentes en los cuales no se identificaron agrupaciones relacionadas con alguna escala específica o con el tipo de reacciones positivas o negativas (Trujillo-Terán, 2012). La validación realizada en Costa Rica mostró ocho factores en la muestra clínica y nueve en la muestra no clínica (Vargas-Fallas & Prada-Villalobos, 2013). Ante las diferencias en la estabilidad factorial y los resultados de este estudio, es importante considerar que la estructura factorial no determina por qué se genera la agrupación de los reactivos, a lo cual se suma que las diferencias encontradas podrían deberse al sesgo de información que se asocia a su aplicación en culturas diferentes (Ramada-Rodilla, Serra-Pujadas, & Delclós-Clanchet, 2013). Por ello, son convenientes estudios que indaguen al respecto y permitan obtener mayores elementos explicativos de las diferencias encontradas.

Los resultados permiten concluir que el CORE-OM es un instrumento de medida confiable que permite explorar un amplio conjunto de síntomas de malestar psicológico en jóvenes uni-

versitarios. En consonancia con la importancia de valorar la sintomatología que los estudiantes experimentan, este instrumento puede resultar útil para adecuar programas de intervención de acuerdo con sus síntomas y problemáticas, además de valorar su impacto.

Estos resultados deben ser interpretados a la luz de las siguientes limitaciones: en primer lugar, los resultados encontrados pueden estar influidos por las características del muestreo por conveniencia (jóvenes universitarios), lo cual no permite generalizarlos a otras poblaciones de interés; incluso sería útil ampliar y variar la muestra y realizar otros estudios para obtener mayores evidencias de validez y funcionalidad y para aportar más información sobre la estabilidad de la estructura del instrumento. Entre otras evidencias, se considera incluir la relación del instrumento con otras mediciones del mismo constructo (validez convergente), la relación con mediciones de constructos diferentes con el objetivo de mostrar que efectivamente puede realizarse una distinción entre ellos (evidencia discriminante) y la capacidad del instrumento para predecir distintos resultados relevantes (evidencia de predicción de criterio). En segundo lugar, están los problemas inherentes a este tipo de estudios basados en autoinformes, con la posible falta de comprensión de ciertos ítems o la poca capacidad de introspección por parte de los estudiantes, por lo que puede resultar útil emplear otras medidas, por ejemplo, entrevistas, para analizar con mayor profundidad la percepción que tienen los jóvenes sobre el malestar emocional.

Referencias

Auerbach, R., Alonso, J., Axinn, W., Cuijpers, P., Ebert, D.,

- Greif, J., ... Bruffaerts, R. (2016). Mental disorders among college students in the World Health Organization World Mental Health Surveys. *Psychological Medicine*, 46(14), 2955-2970. doi: [10.1017/S0033291716001665](https://doi.org/10.1017/S0033291716001665)
- Barkham, M., Gilbert, N., Connell, J., Marshall, C., & Twigg, E. (2005). Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *The British Journal of Psychiatry*, 186(3), 239-246. doi: [10.1192/bjp.186.3.239](https://doi.org/10.1192/bjp.186.3.239)
- Beck, A., Burdett, M., & Lewis, H. (2015). The association between waiting for psychological therapy and therapy outcomes as measured by the CORE-OM. *British Journal of Clinical Psychology*, 54(2), 233-248. doi: [10.1111/bjc.12072](https://doi.org/10.1111/bjc.12072)
- Benjet, C., Borges, G., Medina-Mora, M. E., Méndez, E., Fleiz, C., Rojas, E., & Cruz, C. (2009). Diferencias de sexo en la prevalencia y severidad de trastornos psiquiátricos en adolescentes de la Ciudad de México. *Salud Mental*, 32(2), 155-163. Recuperado de https://revistasaludmental.mx/index.php/salud_mental/index
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological Bulletin*, 112(3), 400-404. doi: [10.1037/0033-2909.112.3.400](https://doi.org/10.1037/0033-2909.112.3.400)
- Berenzon, S., Lara, M. A., Robles, R., & Medina-Mora, M. E. (2013). Depresión: Estado del conocimiento y la necesidad de políticas públicas y planes de acción en México. *Salud Pública de México*, 55(1), 74-80. doi: [10.1590/s0036-36342013000100011](https://doi.org/10.1590/s0036-36342013000100011)
- Branney, P., & Barkham, M. (2006). Core outcomes in psychosexual therapy: A feasibility study of the CORE-OM. *Sexual and Relationship Therapy*, 21(1), 15-26. doi: [10.1080/14681990500281414](https://doi.org/10.1080/14681990500281414)
- Byrne, B. (2009). *Structural equation modeling with AMOS: Basic concepts, applications and programming* (2ª ed.). New York, NY: Routledge.
- Campo-Arias, A., & Oviedo, H. (2008). Propiedades psicométricas de una escala: La consistencia interna. *Revista de Salud Pública*, 10(5), 831-839. doi: [10.1590/s0124-00642008000500015](https://doi.org/10.1590/s0124-00642008000500015)
- Campo-Arias, A., Herazo, E., & Oviedo, H. (2012). Análisis de factores: Fundamentos para la evaluación de instrumentos de medición en salud mental. *Revista Colombiana de Psiquiatría*, 41(3), 659-671. doi: [10.1016/s0034-7450\(14\)60036-6](https://doi.org/10.1016/s0034-7450(14)60036-6)
- Casey, P., Cowan, P., Cowan, C., Draper, L., Mwamba, N., & Hewison, D. (2017). Parents as partners: A U.K. trial of a U.S. couples-based parenting intervention for at-risk low-income families. *Family Process*, 56(3), 589-606. doi: [10.1111/famp.12289](https://doi.org/10.1111/famp.12289)
- Connell, J., Barkham, M., & Mellor-Clark, J. (2007). CORE-OM mental health norms of students attending university counselling services benchmarked against an age-matched primary care sample. *British Journal of Guidance & Counselling*, 35(1), 41-57. doi: [10.1080/03069880601106781](https://doi.org/10.1080/03069880601106781)
- CORE IMS. (2001). Reino Unido: In CORE Information Management Systems. Recuperado el 2 de febrero de 2014 de <https://www.coreims.co.uk>
- Dias-Sales, C. M., de Matos-Moleiro, C. M., Evans, C., & Gomes-Alves, P. C. (2012). Versão portuguesa do CORE-OM: Tradução, adaptação e estudo preliminar das suas propriedades psicométricas. *Revista de Psiquiatria Clínica*, 39(2), 54-59. doi: [10.1590/S0101-60832012000200003](https://doi.org/10.1590/S0101-60832012000200003)
- Díaz, D., Rodríguez-Carvajal, R., Blanco, A., Moreno-Jiménez, B., Gallardo, I., Valle, C., & van Dierendonck, D. (2006). Adaptación española de las escalas de bienestar psicológico de Ryff. *Psicothema*, 18(3), 572-577. Recuperado de <https://www.psicothema.com>
- Echevarría, H. D. (2016). *Los diseños de investigación*

- cuantitativa en psicología y educación*. Río Cuarto: UniRío. Recuperado de <https://www.unrc.edu.ar>
- Elfström, M., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M., & Carlsson, S. (2013). Validation of the Swedish version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology and Psychotherapy*, 20(5), 447-455. doi: [10.1002/cpp.1788](https://doi.org/10.1002/cpp.1788)
- Elosua-Oliden, P. (2003). Sobre la validez de los tests. *Psicothema*, 15(2), 315-321. Recuperado de <https://www.psicothema.com>
- Espíndola-Hernández, J. G., Morales-Carmona, F., Díaz, E., Pimentel, D., Meza, P., Henales, C., ... Ibarra, C. (2006). Malestar psicológico: Algunas de sus manifestaciones clínicas en la paciente gineco-obstétrica hospitalizada. *Perinatología y Reproducción Humana*, 20(4), 112-122. Recuperado de <https://www.medigraphic.com>
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *The British Journal of Psychiatry*, 180(1), 51-60. doi: [10.1192/bjp.180.1.51](https://doi.org/10.1192/bjp.180.1.51)
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health*, 9(3), 247-255. doi: [10.1080/jmh.9.3.247.255](https://doi.org/10.1080/jmh.9.3.247.255)
- Fleiz-Bautista, C., Villatoro-Velázquez, J., Medina-Mora-Icaza, M. E., Moreno-López, M., Gutiérrez-López, M. L., & Oliva-Robles, N. (2012). Sociodemographic and personal factors related to depressive symptomatology in the Mexican population aged 12 to 65. *Revista Brasileira de Psiquiatria*, 34(4), 395-404. doi: [10.1016/j.rbp.2012.03.004](https://doi.org/10.1016/j.rbp.2012.03.004)
- Forero, C., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(4), 625-641. doi: [10.1080/10705510903203573](https://doi.org/10.1080/10705510903203573)
- Freeman, D., & Freeman, J. (2013). *The stressed sex: Uncovering the truth about men, women, and mental health*. Oxford, Reino Unido: Oxford University.
- Guevara-Ruiseñor, E. S. (2008). Vida sexual y malestar emocional. Las contradicciones de la modernidad. *REMO*, 6(14), 2-11. Recuperado de <https://remo.ws>
- IBM Corporation. (2012). IBM SPSS Statistics for Windows (Version 21.0) [software de cómputo]. Armonk, NY: IBM.
- Jansz, J. (2000). Masculine identity and restrictive emotionality. En A. Fischer (Ed.), *Gender and Emotion. Social Psychological Perspectives*, pp. 166-186. Cambridge, Reino Unido: Cambridge University. doi: [10.1017/cbo9780511628191.009](https://doi.org/10.1017/cbo9780511628191.009)
- Juhová, D., Řiháček, T., Cígler, H., Dubovská, E., Saic, M., Černý, M., ... Evans, C. (2018). Česká adaptace dotazníku CORE-OM: Vybrané psychometrické charakteristiky. *Československá Psychologie*, 62(1), 59-74.
- López-Bárcena, J. J., & González de Cossío-Ortiz, M. G. (2005). Modelo universitario de salud integral, en la Universidad Nacional Autónoma de México. *FACMED*, 48(6), 224-231. Recuperado de <https://www.revistas.unam.mx>
- Lyne, J., Barret, P., Evans, C., & Barkham, M. (2006). Dimensions of variation on the CORE-OM. *The British Journal of Clinical Psychology*, 45(2), 185-203. Recuperado de <https://onlinelibrary.wiley.com/journal/20448260>
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22(4), 719-

748. doi: [10.1093/jnci/22.4.719](https://doi.org/10.1093/jnci/22.4.719)

- Margison, F. R., Barkham, M., Evans, C., McGrath, G., Mellor-Clark, J., Audin, K., & Connell, J. (2000). Measurement and psychotherapy. Evidence-based practice and practice-based evidence. *The British Journal of Psychiatry*, *177*(2), 123-130. doi: [10.1192/bjp.177.2.123](https://doi.org/10.1192/bjp.177.2.123)
- Mellor-Clark, J., Barkham, M., Connell, J., & Evans, C. (1999). Practice-based evidence and standardized evaluation: Informing the design of the CORE system. *European Journal of Psychotherapy and Counselling*, *2*(3), 357-374. doi: [10.1080/13642539908400818](https://doi.org/10.1080/13642539908400818)
- Navascués, A., Calvo-Medel, D., & Bombin-Martín, A. (2016). Efectos del bienestar subjetivo y psicológico en los resultados terapéuticos de un hospital de día. *Acción Psicológica*, *13*(2), 143-156. doi: [10.5944/ap.13.2.15818](https://doi.org/10.5944/ap.13.2.15818)
- Otzen, T., & Manterola, C. (2017). Técnicas de muestreo sobre una población de estudio. *International Journal of Morphology*, *35*(1), 227-232. doi: [10.4067/S0717-95022017000100037](https://doi.org/10.4067/S0717-95022017000100037)
- Organización Mundial de la Salud. (2013). Salud mental: Un estado de bienestar. Recuperado el 1 de diciembre de 2013 de <https://origin.who.int>
- Penfield, R. (2009). DIFAS 5. Differential Item Functioning Analysis System. Computer Program Exchange (Versión 1.00) [software de cómputo]. Recuperado de <https://soe.uncg.edu/academics/departments/erm/erm-software>
- Penfield, R. (2013). *DIFAS 5.0. Differential Item Functioning Analysis System. User's Manual*. Recuperado de <https://soe.uncg.edu>
- Penfield, R. D., & Algina, J. (2006). A generalized DIF effect variance estimator for measuring unsigned differential test functioning in mixed format tests. *Journal of Educational Measurement*, *43*(4), 295-312. doi: [10.1111/j.1745-3984.2006.00018.x](https://doi.org/10.1111/j.1745-3984.2006.00018.x)
- Pérez, E., & Medrano, L. (2010). Análisis Factorial Exploratorio: Bases conceptuales y metodológicas. *Revista Argentina de Ciencias del Comportamiento*, *2*(1), 58-66. Recuperado de <https://www.revistas.unc.edu.ar>
- Pérez-Padilla, M. L., Ponce-Rojo, A., Hernández-Contreras, J., & Márquez-Muñoz, B. A. (2010). Salud mental y bienestar psicológico en los estudiantes universitarios de primer ingreso de la Región Altos Norte de Jalisco. *Revista de Educación y Desarrollo*, *14*, 31-37. Recuperado de http://www.cucs.udg.mx/revistas/edu_desarrollo
- Perry, T., Barkham, M., & Evans, C. (2013). The CORE-OM and CORE-OM (SV) in secure settings: A template analysis of the experiences of male patients and their staff. *The British Journal of Forensic Practice*, *15*(1), 32-43. doi: [10.1108/14636641311299068](https://doi.org/10.1108/14636641311299068)
- Ramada-Rodilla, J. M., Serra-Pujadas, C., & Delclós-Clanchet, G. L. (2013). Adaptación cultural y validación de cuestionarios de salud: Revisión y recomendaciones metodológicas. *Salud Pública de México*, *55*(1), 57-66. doi: [10.1590/s0036-36342013000100009](https://doi.org/10.1590/s0036-36342013000100009)
- Rivera-Ledesma, A., Caballero-Suárez, N., Pérez-Sánchez, I., & Montero-López-Lena, M. (2013). SCL-90: Distrés psicológico, género y conductas de riesgo. *Universitas Psychologica*, *12*(1), 105-118. doi: [10.11144/javeriana.upsy12-1.dpgc](https://doi.org/10.11144/javeriana.upsy12-1.dpgc)
- Rosales-Pérez, J. C., Córdova-Osnaya, M., & Ramos-Clatempa, R. (2012). Ideación suicida en estudiantes mexicanos: Un modelo de relación múltiple con variables de identificación personal. *Psicología y Salud*, *22*(1), 63-74. Recuperado de <https://psicologiay-salud.uv.mx/index.php/psicysalud/index>
- Trujillo-Terán, L. A. (2012). *Análisis psicométrico preliminar de la versión en español del CORE-OM: Clinical Outcomes in Routine Evaluation-Outcome Measure* (Tesis de maestría). Universidad de Barcelona, España. Recuperado de <https://babel.banrepcultural.org>

- Uji, M., Sakamoto, A., Adachi, K., & Kitamura, T. (2012). Psychometric properties of the Japanese version of the Clinical Outcomes in Routine Evaluation-Outcome Measure. *Comprehensive Psychiatry*, 53(5), 600-608. doi: [10.1016/j.comppsy.2011.09.006](https://doi.org/10.1016/j.comppsy.2011.09.006)
- Vargas-Fallas, K., & Prada-Villalobos, S. (2013). Validación del instrumento: Resultados Clínicos en Evaluación de Rutina (CORE). *Revista Humanitas*, 10(10), 33-44. Recuperado de <https://www.unirioja.es>
- Viliū-nienė, R., Evans, C., Hilbig, J., Pakalniškienė, V., Danilevičiūtė, V., Laurinaitis, E., & Navickas, A. (2013). Translating the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM) into Lithuanian. *Nordic Journal of Psychiatry*, 67(5), 305-311. doi: [10.3109/08039488.2012.745599](https://doi.org/10.3109/08039488.2012.745599)
- Zhang, Y., Hu, J., Evans, C., Jin, L., Wu, M., Wang, C., ... Chen, G. (2019). Psychometric properties of the Chinese version of the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM). *British Journal of Guidance & Counselling*, 1-11. doi: [10.1080/03069885.2019.1682120](https://doi.org/10.1080/03069885.2019.1682120)
-