

LA VERSIÓN ESPAÑOLA DEL CORE-OM: Clinical Outcomes in Routine Evaluation - Outcome Measure

Guillem Feixas^{1,5}, Chris Evans², Adriana Trujillo¹, Luis Ángel Saúl³, Lluís Botella⁴, Sergi Corbella⁴, Esther González⁴, Arturo Bados¹, Eugeni Garcia-Grau¹ y M. Angeles López-González³

¹ Universitat de Barcelona

² Nottinghamshire Personality Disorder & Development Network, NHS, Reino Unido

³ UNED

⁴ Universitat Ramon Llull

⁵ Institut de Recerca Cervell, Cognició i Conducta (IR3C)

Cómo referenciar este artículo/How to reference this article:

Feixas, G., Evans, C., Trujillo, A., Saúl, L. A., Botella, L., Corbella, S., González, E., Bados, A., Garcia-Grau, E. y López-González, M. A. (2012). La versión española del CORE-OM: Clinical Outcomes in Routine Evaluation - Outcome Measure. *Revista de Psicoterapia*, 23(89), 109-135.

Resumen

Se presenta la versión en español del Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM), un instrumento creado por el Core System Group para la evaluación del cambio terapéutico. El CORE-OM es un cuestionario que evalúa el malestar psicológico a partir de cuatro dimensiones: Bienestar subjetivo, Problemas/Síntomas, Funcionamiento general y Riesgo. Se describe el cuestionario y se exponen los estudios psicométricos realizados, los cuales indican que el instrumento posee un nivel de validez y fiabilidad adecuadas, así como una excelente aceptación y sensibilidad al cambio terapéutico. Con la supervisión y guía de miembros del equipo creador del CORE-OM, se realizó el proceso de traducción de la versión original en inglés. Colaboraron 12 personas competentes en ambas lenguas de diferentes lugares de España; posteriormente 64 castellano-parlantes de distintas condiciones y orígenes lingüísticos participaron en la revisión del instrumento. Así se obtuvo una versión definitiva en español del CORE-OM, de la que se derivaron las versiones más breves resultantes (CORE-SFA, CORE-SFB, CORE-10 y CORE-5) y todas ellas en versión femenina y masculina. Ahora que el CORE-OM está disponible para todos los psicoterapeutas de habla hispana (www.ub.edu/terdep/core), se hace necesario continuar con el estudio de validación con el objetivo de disponer de las propiedades psicométricas del instrumento en su versión en español.

Palabras clave: CORE-OM (Clinical Outcomes in Routine Evaluation-Outcome Measure), traducción, psicología transcultural, investigación de resultados en psicoterapia, práctica basada en la evidencia, evidencia basada en la práctica.

Abstract

We present the Spanish version of the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM), an instrument created by the Core System Group for the evaluation of therapeutic change. The CORE-OM is a self-report questionnaire that assesses psychological distress with domains of Subjective well-being, problems/symptoms, general functioning and risk. We describe its psychometric properties studies which suggest a good level of validity and reliability, as well as excellent acceptability and sensitivity to therapeutic change. With the supervision and guidance of team members from the CORE-OM developers the original English version was translated into Spanish. 12 people, competent in both languages from different parts of Spain, participated in this process and later 64 Castilian-speaking from different demographic and linguistic backgrounds participated in the review of the instrument. This provided a final version in Spanish of the CORE-OM from which the resulting shorter versions (CORE-SFA, SFB-CORE, CORE and CORE-10-5, all in male and female versions) were derived. Now that the CORE-OM is available to all Spanish-speaking psycho-therapists (www.ub.edu/terdep/core), it is necessary to continue the validation study with the objective of having the psychometric properties of the instrument for the Spanish version.

Keywords: CORE-OM (Clinical Outcomes in Routine Evaluation-Outcome Measure), translation, transcultural psychology, psychotherapy outcome research, evidence-based practice, practice-based evidence.

La evaluación del cambio producido durante el proceso terapéutico resulta especialmente importante no sólo para el estudio de la eficacia (véase por ej., Bados, Garcia-Grau y Fusté, 2002; Botella y Feixas, 1994; Corbella y Botella, 2004) sino también como retroalimentación para el terapeuta (y para el propio cliente). Observar el progreso terapéutico no sólo a través de los informes verbales y la comunicación no-verbal sino también a partir de un cuestionario de auto-informe puede aportar una fuente adicional y valiosa de información. El compromiso ético de los psicoterapeutas con la optimización de sus servicios y la mejora de la salud mental de sus pacientes hace que sea necesario disponer de evidencias contrastadas sobre los efectos terapéuticos de los tratamientos psicológicos. Estos resultados que evidencian el efecto positivo (en la mayoría de los casos) de la psicoterapia sirven para justificar la rentabilidad del gasto que genera su inclusión en los servicios psicoterapéuticos públicos y privados. Botella y Feixas (1994) consideran también que el avance de la psicoterapia como disciplina científica depende, en parte, del contraste empírico de las modalidades de intervención propuestas desde las diferentes perspectivas teóricas.

La utilización rutinaria del *Clinical Outcomes in Routine Evaluation-Outcome Measure* (CORE-OM) se enmarca dentro del paradigma de la *práctica basada en la evidencia*, que ha promovido la construcción de herramientas para el asesoramiento y la terapia psicológica orientadas a conocer con mayor claridad qué funciona y para quién. Por otro lado, el enfoque emergente de la *evidencia basada en la práctica* (Barkham, Hardy y Mellor-Clark, 2010; Margison et al. 2000) desarrollado en años recientes, basado en estudios naturalísticos no controlados pero con muestras grandes, requiere también la utilización de medidas de resultado, pues su objetivo se orienta a mostrar qué procedimientos funcionan y son eficaces en la mejora de la calidad de la atención al paciente en situaciones prácticas de la vida real. En la actualidad ambos paradigmas son complementarios, se retroalimentan y fomentan el conocimiento sobre la rigurosidad y eficacia de la práctica psicológica y los servicios existentes para tal fin (Barkham, Mellor-Clark, Connell, y Cahill, 2006).

En consecuencia, el esfuerzo en pro de la evaluación sistemática se orienta a identificar el potencial de una intervención dada, siendo importante la existencia de un mecanismo complementario con el cual monitorizar el cumplimiento de los criterios de evidencia; por tanto, se hace necesaria la estandarización de herramientas, que como el CORE-OM, son útiles para medir la efectividad y la eficacia de los servicios de terapias psicológicas (Evans et al., 2000; Mellor Clark, Barkham, Conell y Evans, 2007; Evans, in press).

Beutler (2001) aboga porque parte de los recursos utilizados en la evaluación de los efectos terapéuticos de los tratamientos psicológicos provengan de la aplicación de algún tipo de instrumento de medida. Además del CORE-OM, es amplia la variedad de cuestionarios, inventarios y escalas que se han utilizado en las investigaciones para comprobar la mejora sintomática de los pacientes y evaluar los

resultados finales; sin embargo, sin una comunicación y organización adecuada de los datos puede subsistir un estado de desorden, con profesionales independientes o entidades de salud mental eligiendo cada uno unas medidas de respuesta y resultado sin idea de cómo integrar la información entre clínicos, lugares de tratamiento o dimensiones del problema (Froyd, Lambert y Froyd, 1996). Por otro lado, la mayoría de los instrumentos sintomáticos que se utilizan para evaluar el cambio son específicos de un trastorno o problema clínico dado, lo que hace más difícil la comparación o agregación de datos.

Es por ello que propuestas como el *Sistema CORE*, desarrollado por el Core System Group (Evans et al. 1998) que funciona a partir de la aplicación del CORE-OM y sus diversas versiones, adquieren especial importancia al integrar la evaluación de pacientes con sintomatología muy diversa y el análisis de los datos recogidos tanto a nivel individual como institucional, demostrando que los procedimientos funcionan y son eficaces para mejorar la calidad de la atención al paciente en los contextos en los que se practica la psicoterapia en la vida real (Barkham, Mellor-Clark, Connell, y Cahill, 2006).

Gracias al desarrollo e implementación de este instrumento en diversos países, centros de terapia y de investigación, se espera aumentar la frecuencia de uso y los periodos de tiempo en los que se recopila información. Con ello se pretende ampliar el conjunto de datos útiles en estudios comparativos sobre eficacia en diferentes poblaciones y en diferentes niveles de prestación del servicio de atención en salud mental (Evans et al. 1998).

Como cuestionario de medida de respuesta al tratamiento, esta prueba puede aplicarse y ser útil según lo planteado por el Departamento de Salud del Reino Unido (The DoH Strategic Review, 1996) tanto como mecanismo de apoyo para la evaluación de la eficacia y la eficiencia, como en la provisión de servicios psicológicos en cuatro áreas específicas: (a) la evaluación de la atención individual del paciente, (b) la evaluación de servicios para determinados grupos de pacientes, (c) como estrategia de retroalimentación para los proveedores de instrumentos psicológicos y, (d) en investigación.

Como evidencia de ello, el CORE-OM y sus diversas versiones han sido utilizados en contextos muy variados: (1) en servicios de atención primaria y secundaria (Barkham, Gilbert, Connell, Marshall, y Twigg, 2005); (2) en la validación del instrumento en población de edad avanzada, en la que la medida ha demostrado mantener su sensibilidad en comparación con los datos de referencia de población adulta en edad de trabajar (Barkham, Culverwell, Spindler, y Twigg, 2005); (3) como medida de respuesta respecto a síntomas y funcionamiento en el estudio del efecto de los antidepresivos serotoninérgicos y noradrenérgicos en el procesamiento emocional en pacientes deprimidos (Tranter et al. 2009); (4) en servicios de atención primaria y asesoría estudiantil (Connell, Barkham y Mellor-Clark, 2007); (5) en la evaluación del riesgo de los pacientes (Bedford, Lukic, y Tibbles, 2011); (6) en la implementación de terapias cognitivo-conductuales

computarizadas (Cavanagh, Seccombe y Lidbetter, 2011); y (7) para valorar la percepción de terapeutas y clientes respecto a la utilización de medidas de evaluación rutinaria (Unsworth, Cowie, y Green, 2011) entre muchas otras (para una ampliación de las revisiones realizadas es posible visitar la página web del *CORE Information Management Systems*: www.coreims.co.uk).

Así mismo, el CORE-OM ha demostrado su utilidad en la evaluación de diversos servicios de terapia psicológica (Leach et al. 2005). En España, se ha utilizado el sistema de datos CORE (CORE System Data) en más de 500 usuarios a lo largo de seis años, con el objetivo de evaluar el progreso terapéutico de forma continuada en un centro de psicoterapia universitario con muy buenos resultados. El instrumento y, en general, el sistema ha sido útil y se ha integrado dentro del servicio con diversos objetivos: para valorar la efectividad terapéutica, como referencia en la construcción y desarrollo de planes terapéuticos, para evaluar la gestión del servicio, para estudiar la relación entre finalización inesperada de la terapia y la alianza terapéutica, y como recurso en la supervisión de casos de los psicólogos en formación que llevaban a cabo sus prácticas en este centro (Botella, 2006).

En este artículo se pretende describir el proceso de traducción y adaptación del CORE-OM en español, facilitar el acceso al instrumento y dar a conocer sus posibilidades de aplicación con el propósito de consolidarlo en el futuro como una herramienta ecuaníme y efectiva. De esta forma se intenta promover su uso generalizado tanto en la evaluación de personas individuales, grupos de pacientes, instituciones y servicios de atención en salud mental como los estudios en distintas áreas con población de habla hispana.

Descripción del cuestionario

El CORE-OM es un instrumento de libre acceso (www.coreims.co.uk) que puede ser utilizado con la única condición de no modificar su contenido ni el formato en el que se presenta. Fue creado por el *Core System Group* (1998) como parte de un protocolo de evaluación de los servicios de asistencia psicológica. Desarrollado en la Universidad de Leeds junto con otros centros del Reino Unido, el CORE-OM es un cuestionario de autoinforme que contiene 34 ítems que caben en una hoja Din A4 en doble cara, es breve y rápido de contestar, lo que facilita su aplicación (Beutler, 2001). Actualmente es el instrumento de evaluación de resultados más empleado en el Reino Unido, especialmente en los servicios públicos.

Tal como lo explican Evans et al. (2002) y Lyne, Barrett, Evans y Barkham (2006), el CORE-OM evalúa el estado del cliente a partir de una serie de dimensiones definidas como (véase tabla 1):

- *Bienestar subjetivo*, compuesta por 4 ítems;
- *Problemas/Síntomas*, con 12 ítems que valoran ansiedad, depresión, trauma y síntomas físicos;

- *Funcionamiento general*, que incluye 12 ítems que evalúan relaciones íntimas, sociales y otros aspectos del funcionamiento cotidiano;
- *Riesgo*, escala compuesta por 6 ítems que son utilizados como indicadores clínicos de intentos de suicidio, autolesiones o actos de agresión a terceros.

El valor de cada ítem puede estar entre 0-4 y puede puntuarse manualmente o corregirse mediante lectora óptica. En cuanto al formato, se alternan los renglones en fondo claro y oscuro para cada uno de los ítems, intentando con ello reducir el riesgo de perder la línea de respuesta al contestar el cuestionario (Evans et al. 2000).

El CORE-OM incluye ítems de alta y baja intensidad con el objetivo de incrementar el nivel de discriminación, además el 25% de los ítems se encuentran redactados “positivamente”, esto es, reflejan reacciones positivas de la persona (ítems 4 y 17 de la dimensión de bienestar subjetivo, ítems 3, 7, 12, 19, 21 y 32 de la dimensión de funcionamiento cotidiano). El instrumento está diseñado de forma que a mayor puntuación mayor nivel de problemas o síntomas. Teniendo esto en cuenta, la puntuación de la escala *bienestar subjetivo* resulta en cierto modo contradictoria (puesto que una mayor puntuación indica mayor *malestar*), pero los autores optaron por mantener el nombre de la escala, aunque la puntuación que se obtenga esté invertida, por criterios de consistencia con las demás puntuaciones. En todo caso, es especialmente importante tener este punto en cuenta si se quiere interpretar esta escala por separado (Evans et al. 1998).

Tabla 1
Desglose por dimensiones del CORE-OM (Evans et al. 1998)

Ítem N°	Dimensiones	Ítems*	Intensidad del ítem
4	Bienestar Subjetivo (W)	Me he sentido bien conmigo mismo**	Baja
14	Bienestar Subjetivo (W)	He tenido ganas de llorar	Alta
17	Bienestar Subjetivo (W)	Me he sentido agobiado por mis problemas	Baja
31	Bienestar Subjetivo (W)	Me he sentido optimista sobre mi futuro**	Alta
2	Problemas-Síntomas Ansiedad (P)	Me he sentido tenso, ansioso o nervioso	Baja
11	Problemas-Síntomas Ansiedad (P)	La tensión y la ansiedad me han impedido hacer cosas importantes	Alta
15	Problemas-Síntomas Ansiedad (P)	He sentido pánico o terror	Alta
20	Problemas-Síntomas Ansiedad (P)	Me ha sido imposible dejar a un lado mis problemas	Baja
5	Problemas-Síntomas Depresión (P)	Me he sentido totalmente sin energía ni entusiasmo	Alta

Ítem N°	Dimensiones	Ítems*	Intensidad del ítem
23	Problemas-Síntomas Depresión (P)	Me he sentido desesperado o sin esperanza	Alta
27	Problemas-Síntomas Depresión (P)	Me he sentido infeliz	Baja
30	Problemas-Síntomas Depresión (P)	He pensado que yo tengo la culpa de mis problemas y dificultades	Baja
8	Problemas-Síntomas físicos (P)	He tenido molestias, dolores y otros problemas físicos	Baja
18	Problemas-Síntomas físicos (P)	He tenido dificultad para conciliar el sueño o permanecer dormido	Baja
13	Problemas-Síntomas Trauma (P)	Me han inquietado pensamientos y sentimientos no deseados	Alta
28	Problemas-Síntomas Trauma (P)	Me han angustiado imágenes o recuerdos no deseados	Alta
7	Funcionamiento General (F)	Me he sentido capaz de afrontar las cosas cuando han ido mal**	Alta
12	Funcionamiento General (F)	Me he sentido satisfecho con las cosas que he hecho**	Baja
21	Funcionamiento General (F)	He sido capaz de hacer la mayoría de las cosas que tenía que hacer**	Baja
32	Funcionamiento General (F)	He conseguido las cosas que quería**	Alta
1	Funcionamiento Relaciones Íntimas (F)	Me he sentido muy solo y aislado	Alta
3	Funcionamiento Relaciones Íntimas (F)	He sentido que tenía a alguien en quien apoyarme cuando lo he necesitado**	Baja
19	Funcionamiento Relaciones Íntimas (F)	He sentido afecto o cariño por alguien**	Baja
26	Funcionamiento Relaciones Íntimas (F)	He pensado que no tengo amigos	Alta
10	Funcionamiento Relaciones Sociales (F)	Me ha costado mucho hablar con la gente	Alta
25	Funcionamiento Relaciones Sociales (F)	Me he sentido criticado por los demás	Baja
29	Funcionamiento Relaciones Sociales (F)	Me he sentido irritable cuando estaba con otras personas	Baja
33	Funcionamiento Relaciones Sociales (F)	Me he sentido humillado o avergonzado por otras personas	Alta
9	Riesgo Daño a sí mismo (R)	He pensado en hacerme daño a mí mismo	Baja

Ítem N°	Dimensiones	Ítems*	Intensidad del ítem
16	Riesgo Daño a sí mismo (R)	He hecho planes para acabar con mi vida	Alta
24	Riesgo Daño a sí mismo (R)	He pensado que sería mejor que estuviera muerto	Baja
34	Riesgo Daño a sí mismo (R)	Me he hecho daño físicamente o he puesto en peligro mi salud	Alta
6	Riesgo Daño a otros (R)	He sido violento físicamente hacia los demás	Alta
22	Riesgo Daño a otros (R)	He amenazado o intimidado a otra persona	Alta

Notas.

(*) Los ítems se han descrito en español de acuerdo a la traducción realizada en este estudio de adaptación;

(**) ítems redactados en forma positiva.

Dado su diseño, el CORE-OM intenta captar lo más esencial del malestar psicológico de los clientes, por lo que es adecuado como herramienta de evaluación inicial y como medida de resultado. Al igual que la mayoría de los instrumentos de auto-informe no se puede utilizar para obtener un diagnóstico de un trastorno específico. Sus creadores sugirieron inicialmente la utilización del cuestionario desde el inicio de la terapia o en la evaluación pre-tratamiento, ya que proporciona información sobre la gravedad de los síntomas, en la finalización y también en estadios intermedios en casos de terapia de larga duración. Consideran que es una medida útil para realizar el seguimiento del proceso terapéutico y que incluso se pueden aplicar rutinariamente versiones más breves (ver más adelante), en cada sesión, a lo largo de la psicoterapia (Evans et al. 2000; Evans et al. 2002), aunque Evans (2012) en este caso advierte que no todas las decisiones terapéuticas deben basarse en las puntuaciones obtenidas.

El sistema CORE

El CORE-OM forma parte de un sistema de gestión de la información denominado Sistema CORE (*CORE System*) desarrollado por un grupo multidisciplinario de profesionales e investigadores y su contenido se explica gracias a la colaboración de los directivos y encargados de los diferentes servicios públicos del sistema de salud británico (Mellor-Clark, Barkham, Connell y Evans, 1999). En este sistema se incluye información sobre la evaluación inicial (línea base) y la evaluación progresiva de los resultados terapéuticos, incluyendo datos sobre la demografía de los pacientes, la casuística y los antecedentes en el servicio (Mellor-Clark y Barkham, 2000). Se compone de tres herramientas principales, que

se utilizan a partir de los datos que proveen los usuarios durante la evaluación pretratamiento y postratamiento utilizando el CORE-OM. Los profesionales, a su vez, completan el *Formato de Evaluación de la Terapia CORE* antes de iniciar la terapia y el *Formato Final de Terapia* una vez el proceso ha culminado.

Como complemento de la evaluación realizada con el CORE-OM, el *Formato de Evaluación de la Terapia CORE* recoge datos básicos a nivel contextual que ayudan a mejorar la evaluación y el desarrollo del proceso terapéutico ofrecido por el servicio (Mellor-Clark et al. 1999). Este formato indaga sobre aspectos como el apoyo prestado al cliente, la atención previa y/o simultánea a la terapia y el seguimiento de la medicación; así como un sistema de categorización para las dificultades que se presentan, su impacto sobre el funcionamiento cotidiano y otros riesgos asociados. En relación con la calidad del servicio, se recopilan datos que permiten la obtención de un perfil sobre la accesibilidad y la adecuación de los servicios ofrecidos; para ello se incluyen datos demográficos de los clientes, tiempos de espera e idoneidad de la remisión (Barkham et al. 2006).

Por último, el *Formato Final de Terapia* incluye algunos datos adicionales que permiten obtener los descriptores básicos del tratamiento que, junto a la interpretación de las puntuaciones del CORE-OM, ayudan a contextualizar los resultados de la terapia así como obtener información sobre el desarrollo del servicio ofrecido. La información recogida incluye duración del tratamiento, tipo de intervención, modalidad y frecuencia. Para mejorar la calidad del servicio se recopilan datos sobre la eficacia y la eficiencia en la prestación del servicio. Se incluye información sobre el problema y el riesgo asociado, beneficios de la terapia, tasas de asistencia a las sesiones y finalización del tratamiento (planeado o no planeado) (Barkham et al. 2006).

El total de los datos recogidos se pueden manejar mediante el CORE-Net, un software estandarizado diseñado para ayudar a los servicios a analizar e informar sobre sus datos cuando sea necesario. Esta herramienta se adapta según el enfoque de análisis, las necesidades particulares y los informes requeridos (Mellor-Clark, Curtis Jenkins, Evans, Mothersole y McInnes, 2006).

Familia de medidas CORE

La estructura del *Sistema CORE* ofrece una amplia gama de opciones para los profesionales, siendo posible elegir entre diferentes niveles de evaluación según los objetivos particulares en cada caso (Barkham et al. 2006).

El conjunto de medidas parte del CORE-OM como instrumento principal, presentando una serie de versiones adicionales que han sido diseñadas con diferentes propósitos. Para la fase de valoración inicial y la posterior evaluación del resultado, se recomienda la utilización del CORE-OM. Para la administración repetida (sesión por sesión) existen dos formas breves de 18 ítems (CORE-SFA y CORE-SFB; Barkham et al. 2001) que han demostrado un alto nivel de validez

convergente en relación con el BDI-II (Cahill et al. 2006). Las versiones breves han sido diseñadas especialmente para la investigación, cuando interese alternar las formas A y B en cada sesión con el objetivo de reducir los efectos de la memorización de los ítems. Así mismo, se encuentra disponible una versión breve de 10 ítems (CORE-10; Barkham et al. 2012) diseñada como una herramienta de cribado y medida de resultado de rápida aplicación para usar en cada sesión; incluso existe una versión aún más breve que consta de 5 ítems (CORE-5). Estos formatos facilitan su uso en los entornos de evaluación rutinaria, en los que una única medida breve y corta puede ser apropiada para monitorizar el progreso en la práctica clínica habitual (Barkham et al. 2006). Otras posibilidades de evaluación incluyen la aplicación del cuestionario a población no-clínica (GP-CORE) con 14 ítems derivados del CORE-OM (Sinclair, Barkham, Evans, Connell, y Audin, 2005), así como la versión para adolescentes (YP-CORE) (Twigg et al. 2009) mucho más breve y que contiene ítems redactados de forma distinta. Finalmente, se encuentra en desarrollo una versión dirigida a personas con dificultades de aprendizaje (CORE-LD), que incluye además de ítems simplificados del CORE-OM nuevos ítems destinados a abordar los principales problemas a los que se enfrentan. El YP-CORE y el CORE-LD aun no están disponibles en español.

Desarrollo de la versión original del CORE-OM

El origen del CORE-OM se remonta al capítulo realizado por Waskow (1975), titulado “*Selección de una batería básica*”, producto de la conferencia científica de la American Psychological Association (APA) en 1970, en la que se discutieron temas relacionados con las medidas de cambio en psicoterapia. Posteriormente, a mediados de la década de 1990, se retomó el tema de las medidas de cambio durante una reunión científica realizada por la APA (Strupp, Horowitz y Lambert, 1997) lo que terminó por impulsar las iniciativas existentes en el Reino Unido para la elaboración de un instrumento de medida de resultado que pudiera ser ampliamente adoptado por profesionales e investigadores (Barkham et al. 1998).

En los años siguientes, *The Mental Health Foundation* optó por apoyar el desarrollo de un instrumento básico para la medición del resultado y el cambio en psicoterapia. Adicionalmente se incluyeron datos sobre el contexto en el que se producía la terapia o asesoramiento psicológico, y todo ello se consolidó en el proyecto del *Sistema CORE*, como un medio plausible para el almacenamiento y análisis de información de diferentes fuentes, que pueden a la vez ser combinados para obtener un nivel de evidencia relevante para los profesionales e investigadores (Barkham et al. 2006).

El CORE-OM fue creado como respuesta a la existencia y utilización de medidas sin el adecuado rigor psicométrico. Evidencia de ello fue la identificación de una amplia variabilidad de los instrumentos utilizados que fueron informados en una encuesta aplicada a los representantes, administradores y proveedores de

servicios de terapias psicológicas en el Sistema Nacional de Salud del Reino Unido a la que respondieron 26 de 78 jefes ejecutivos de Administración y Autoridades del Servicio de Salud Familiar (33%) y 220 de sus 1.190 proveedores (18%) (Evans et al. 2000).

La encuesta incluía datos sobre la información que se solicitaba antes del inicio del tratamiento. Se incluyeron aspectos como perfiles del problema, características del paciente y tratamientos previos. Una vez finalizado el tratamiento, la información sobre el resultado de la terapia provenía de informes clínicos y la revisión de las medidas iniciales, en caso de haber utilizado algún instrumento específico; tan solo dos (8%) de los servicios encuestados dijeron recibir información proveniente de evaluaciones rutinarias como una medida del cambio producido durante la terapia (Evans et al. 2000).

En general, la información recogida por los servicios clínicos se clasificó en cuatro categorías principales: síntomas, funcionamiento, historia personal/médica y riesgo. En el ochenta por ciento de los casos se utilizaron medidas nomotéticas, dos terceras partes de estas solo eran utilizadas en un único servicio o consulta y 77 instrumentos eran utilizados por todos los servicios encuestados. Aunque eran pocos los servicios que realizaban evaluaciones al final del tratamiento, las principales áreas abordadas fueron: síntomas, funcionamiento, satisfacción con el servicio y bienestar subjetivo. El 59% de los servicios utilizaron algún tipo de medida. En total se registró la utilización de 57 instrumentos diferentes para la fase de seguimiento y el 63% de ellos fueron aplicados en un solo lugar. Solo el 16% de los servicios utilizaron las mismas medidas antes y después del tratamiento. En este punto, tanto profesionales como proveedores de instrumentos psicológicos consideraron deseable la utilización de una medida básica de resultado (Evans et al. 2000).

Al indagar sobre los requisitos que los encuestados consideraban importantes para una medida básica de resultado, se establecieron tres categorías: contenido, proceso y utilidad. Respecto al contenido, el cuestionario debería ser corto y legible, teóricamente imparcial, capaz de detectar el cambio clínico, válido y fiable, sensible a las necesidades del cliente y susceptible de establecer una relación entre el *input* del cliente y el *output* del servicio. En cuanto al proceso sería deseable que el instrumento no sea intrusivo, de fácil administración y calificación, con información útil tanto de la población clínica como de la no clínica, y fácil de interpretar. Finalmente, se esperaba que a nivel de utilidad contribuyera a la evaluación, mejorase el abordaje de los casos, facilitase la comparación entre servicios, fuera asequible y promoviera la planificación y el desarrollo del curso de la terapia (Evans et al. 2000).

Entre los servicios que promovían la utilización de medidas en algún punto del proceso terapéutico, se identificaron aquellos instrumentos que gozaban de mayor popularidad: *Beck Depression Inventory* (Beck, Ward, Mendelson, Mock y Erbaugh,

1961), *Hospital Anxiety and Depression Scales* (Zigmond y Snaith, 1983), *Symptom Check List 90 y 90-R* (Derogatis, Lipman y Covi, 1973; Derogatis, 1983; Derogatis y Melisaratos, 1983) y su versión abreviada el *Brief Symptom Inventory*, *General Health Questionnaire* (Goldberg, 1978; Goldberg y Williams, 1988), *Rosenberg Self-Esteem Inventory* (Rosenberg, 1965) y el *Inventory of Interpersonal Problems* (IIP, Horowitz, Rosenberg, Baer, Ureno y Villasenor, 1988; IIP-32, Barkham, Startup y Hardy, 1996). Todos estos instrumentos se orientan principalmente a la evaluación de los síntomas y/o problemas específicos e indican su presencia y gravedad, por lo que su objetivo no es medir el resultado global y el cambio producido durante la terapia psicológica, sino la valoración de un conjunto de síntomas. Esta constatación confirmó el valor de la construcción y desarrollo de una medida de malestar general que pudiera proporcionar información sobre el cambio terapéutico en un sentido más global (Evans et al. 2000).

El siguiente paso fue constituir un grupo que incluía psicólogos clínicos, psiquiatras e investigadores no clínicos de orientaciones teóricas distintas (psicoanalíticas o dinámicas, interpersonales, humanísticas, cognitivas, cognitivo-conductuales y sistémicas). Posteriormente, se aumentó la representatividad con la inclusión de investigadores no clínicos, terapeutas de comunidad terapéutica y no profesionales. Finalmente, un total de 45 personas contestaron 638 ítems obtenidos de los instrumentos mencionados que se utilizaban con mayor frecuencia como medidas de referencia, además del *Beck Anxiety Inventory* (Beck, Epstein, Brown y Steer, 1988), el *Borderline Syndrome Index* (Conte, Plutchik, Karasu y Jerrett, 1980), el *Irritability, Depression and Anxiety Inventory* (Snaith, Constantopoulos, Jardine y McGuffin, 1978) y el *Personality Diagnostic Questionnaire* (PDQ-R, Hyler y Rieder, 1987) que se incluyeron con el objetivo de abarcar áreas poco representadas (Evans et al. 2000). Los ítems presentados fueron clasificados según su contenido y, a partir de los resultados de fiabilidad entre jueces, se agruparon y eliminaron aquellos que fueron considerados como problemáticos. Evans et al. (2000) invirtieron varias semanas formulando un conjunto final de 40 ítems orientados a la evaluación de áreas como la autoestima, problemas/síntomas, funcionamiento cotidiano y riesgo. Se pretendía que los ítems abarcaran tanto experiencias intrapersonales como interpersonales, positivas y negativas y que, por tanto, no se centraran exclusivamente en problemas.

La propuesta inicial de ítems, su distribución en escalas y el texto que introduce la instrucción general fueron probados en un amplio grupo de terapeutas, investigadores y muestra comunitaria, a quienes también se les solicitó una retroalimentación cualitativa de la prueba. De la misma forma, se consultó con terapeutas, servicios multiculturales de salud mental y usuarios pertenecientes a diversos grupos étnicos con el objetivo de realizar los cambios necesarios respecto a lenguaje y diferencias relacionadas con la cultura (Evans et al. 2000).

Considerando los requisitos que los encuestados señalaron como importantes

para una medida básica de resultado (contenido, proceso y utilidad), las puntuaciones obtenidas de los ítems se concluyó que las medidas existentes por lo general eran demasiado largas o se centraban en áreas demasiado particulares. A estos aspectos se suma que la mayoría de instrumentos desarrollados para este fin son costosos puesto que requieren del pago de los derechos de autor (*copyright*) para su utilización (Evans et al. 2000).

Como resultado de estos ajustes y de las consideraciones de los encuestados se obtuvo un modelo de cuestionario compuesto por 34 ítems presentados en una hoja a doble cara de tamaño A4. Inicialmente, en la cabecera, se piden algunos datos demográficos básicos para identificar al respondiente, posteriormente se presentan los ítems que deben ser puntuados mediante una marca en la casilla de respuesta con los mismos cinco niveles de respuesta para todos los ítems (Evans et al. 2000).

Características psicométricas y datos normativos originales

Las personas incluidas en el estudio de los datos normativos (Evans et al. 1998) pertenecían a dos grupos: una población no clínica compuesta por estudiantes universitarios y una muestra de conveniencia del personal, familiares y amigos relacionados con el equipo de investigación ($n = 1.106$), y una población clínica compuesta por los usuarios que se encontraban en lista de espera o dentro de un proceso terapéutico en diversas opciones de servicio en el área de salud mental en el Reino Unido ($n = 890$).

Los datos de población no clínica procedían de tres muestras. La primera fue de 691 estudiantes encuestados de una universidad británica, de los que 304 (44%) eran mujeres, 381 (55%), hombres y 6 no completaron los datos respecto al género (1%). El rango de edad fue de 17 a 43 años con cuartiles en 19, 20 y 23. La segunda muestra comprendía 55 estudiantes de otra universidad que participaron en el estudio de fiabilidad test-retest. De estos, 8 (15%) eran hombres, 46 (84%) mujeres, y una persona no declaró su género. Sus edades estaban entre los 20 y 45 años, con cuartiles en 20, 21 y 22 años, lo que quiere decir que el 75% tenían 22 años o menos (Evans et al. 1998).

La última muestra no clínica fue una muestra de conveniencia ($n = 360$) compuesta por terapeutas, investigadores, colegas, amigos y familiares. Estaba formada por 251 (70%) mujeres y 109 (30%), hombres. El rango de edad de esta muestra fue de 14 a 45 años y cuartiles en 18, 20 y 23 (Evans et al. 1998).

Por otra parte, los datos clínicos fueron recogidos principalmente a través del Sistema Nacional de Salud del Reino Unido, además de un servicio de asesoramiento estudiantil y un servicio de apoyo para el personal. La orientación teórica de los servicios fue variable; un reducido número se definieron como puramente comportamentales o cognitivo-conductuales, pero la mayoría se catalogaban como servicios eclécticos y había algunos con una fuerte adscripción psicodinámica.

Los datos clínicos fueron recogidos durante el primer contacto con cada individuo (fase de pre-tratamiento) o durante la primera sesión de tratamiento. El

rango de cuestionarios aplicados en cada servicio estuvo entre 10 y 196 (media 42). El género fue registrado en 874 de los 890 casos (98%): 530 (61%) fueron mujeres y 344 (39%), hombres. En cuanto a la edad, se registraron 850 datos que oscilaron entre los 16 y 78 años, con una media de 36, una mediana de 34 y cuartiles en 26, 34 y 45 (Evans et al. 1998).

Evans et al. (2002) aportan los resultados originales respecto a la utilidad de la prueba, fiabilidad interna, test-retest, validez convergente y discriminante (con un tamaño del efecto grande en la comparación de muestras en el CORE-OM y un tamaño del efecto pequeño en las variables sociodemográficas), correlaciones entre las puntuaciones por dimensiones y un análisis factorial exploratorio. Se llevaron a cabo pruebas no-paramétricas dado que las distribuciones diferían significativamente respecto a la curva normal. Todas las pruebas inferenciales de diferencias fueron bilaterales con un alfa de 0,05. En esta investigación se siguieron los planteamientos de Gardner y Altman (1986) quienes recomiendan en muestras de gran tamaño con una potencia estadística alta informar del tamaño del efecto y los intervalos de confianza.

Aceptabilidad

El principal requisito para cualquier instrumento es que el usuario logre contestarlo. En total, el 91% de los participantes de la muestra no clínica y el 80% de los pertenecientes a la muestra clínica completaron la prueba, es decir, contestaron todos los ítems. La diferencia fue estadísticamente significativa ($p < 0.0005$). Los ítems omitidos fueron pocos, lo que permitió realizar el pro-rateo de las puntuaciones correspondientes. 1084 (98%) de los participantes de la población no clínica y 863 (97%) de los pertenecientes a la muestra clínica respondieron los ítems suficientes para permitir la puntuación de los cuestionarios (Evans et al. 2002).

El ítem que se dejó sin contestar con mayor frecuencia en ambas muestras (2,5% en el grupo no-clínico y 3,8% en la muestra clínica) fue el número 19 (“He sentido afecto o cariño por alguien”). La tasa de omisión general fue de 1,7%. Se observó una heterogeneidad en la omisión de los ítems, que sugiere que los ítems finales fueron omitidos con mayor frecuencia, pero no existe ninguna conexión con alguna dimensión específica (Evans et al. 2002).

Consistencia interna

La fiabilidad interna fue identificada mediante el coeficiente alfa (Cronbach, 1951), el cual indica la proporción de varianza que es resultado de la covariación de los ítems. Todas las dimensiones mostraron un alfa de entre 0,75 y 0,90, lo que implica que la fiabilidad interna es apropiada. La dimensión de Problemas/síntomas mostró una menor fiabilidad en la muestra clínica con respecto a la no clínica. Aun así, esta diferencia del 2% (88% vs. 90%) en la proporción de la covarianza no se consideró problemática (Evans et al. 2002).

Estabilidad test-retest

Cuando se identifican cambios marcados en la puntuación de la prueba en periodos cortos de tiempo, se plantean problemas en la estabilidad del instrumento. La correlación obtenida sirve de indicador de hasta qué punto son estables las mediciones realizadas durante la primera aplicación del test. De 55 estudiantes que fueron contactados, 43 devolvieron los cuestionarios completados en ambas ocasiones. Las correlaciones test-retest fueron altas (0,87-0,91) en las diferentes dimensiones excepto en la escala de riesgo que fue más baja con un 0,64, lo cual no sorprende en vista de la corta duración y la dependencia situacional propia de estos ítems. Al comparar las medianas de las puntuaciones test y retest se encontraron disminuciones pequeñas pero significativas en Problemas/síntomas y Funcionamiento (Evans et al. 2002).

Validez convergente

El diseño del CORE-OM tiene como objetivo señalar diferencias entre los pacientes y proporcionar información sobre cambios durante la terapia según las dimensiones que lo componen. La falta de correlación respecto a otras medidas clínicas podría sugerir la invalidez de la prueba. En este caso, las correlaciones más altas se presentan en medidas conceptualmente cercanas, evidenciando la existencia de validez convergente y que por ende, las puntuaciones no solo reflejan conjuntos comunes de respuesta, sino también diversas dimensiones, todas relacionadas con el malestar psicológico de la persona que contesta la prueba. Los instrumentos que mostraron correlaciones altas con el CORE-OM fueron: *Beck Anxiety Inventory* (BAI, Beck et al. 1988), *Beck Depression Inventory* (BDI, Beck et al. 1961, 1996), *Symptom Checklist-90-R* (SCL-90-R, Derogatis, 1983), *Brief Symptom Inventory* (BSI, Derogatis y Melisaratos, 1983), *Inventory of Interpersonal Problems* (IIP-32, Barkham et al. 1996), y *General Health Questionnaire*, versión de 28 ítems (GHQ, Goldberg y Hillier, 1979). También se utilizó la nueva versión del Inventario para la Depresión de Beck (*Beck Depression Inventory*, BDI-II, Beck et al. 1996), pero la muestra resultó insuficiente para una buena precisión (Evans et al. 2002).

Diferencias entre las muestras clínica y no clínica

El requisito principal de cualquier medida clínica de malestar psicológico es que debe reflejar las diferencias existentes entre las poblaciones clínicas para las cuales ha sido diseñado y las muestras no clínicas de comparación. Los datos evidencian diferencias importantes entre la población clínica y la no clínica, siendo significativas en todas las dimensiones ($p < 0.0005$) (Evans et al. 1998). Las diferencias fueron estimadas con precisión, tal como demuestran los reducidos intervalos de confianza hallados, y fueron superiores a un punto (en una escala de 0 a 4) en todas las dimensiones con excepción de la de riesgo (Evans et al. 2002).

Diferencias de género

Las diferencias a nivel de género pueden afectar un cuestionario si hay sesgo en la respuesta. El diseño del CORE-OM buscaba minimizar el sesgo debido al género teniendo presente que igualmente éste podría presentarse. Los resultados muestran diferencias moderadas, pero estadísticamente significativas respecto al género en la muestra no clínica en todas las dimensiones excepto en *Funcionamiento Cotidiano*. Las diferencias en la muestra clínica fueron pequeñas pero estadísticamente significativas en *Bienestar Subjetivo* y, con menor diferencia, en *Riesgo*. En general, los resultados sugieren que la variable género debe tenerse en cuenta a la hora de interpretar las puntuaciones individuales de los datos clínicos, pero son irrelevantes en comparación con los efectos del estatus clínico frente al no clínico (Barkham, et al. 1998).

Correlaciones entre las puntuaciones de las diversas dimensiones

Dada la interrelación entre las dimensiones clínicas, se espera que las puntuaciones correlacionen positivamente. Las correlaciones analizadas muestran que los ítems de riesgo presentan bajas correlaciones con las otras dimensiones, con mayores valores en la muestra no clínica que en la muestra clínica. Las otras tres escalas mostraron altas correlaciones entre ellas (Barkham, et al. 1998).

Análisis de componentes principales

Los creadores del CORE-OM pretendían que los 34 ítems cubrieran cuatro dimensiones (ver secciones anteriores). Sin embargo, ya era esperable que se diera una alta correlación entre ellas por lo que no se plantearon como factores o componentes independientes. Aun así, Barkham, et al. (1998) realizaron un análisis de componentes principales de forma separada para las muestras clínica y no clínica. Para la muestra no clínica, el resultado evidencia la amplia proporción de varianza en el primer componente (38%) respecto a los tres componentes analizados. El patrón de la matriz después de la rotación oblicua muestra una clara separación de los ítems en un grupo de presencia de problemas, un grupo compuesto en gran parte de los ítems de riesgo y un grupo correspondiente a reacciones positivas de la persona.

Para la muestra clínica, de nuevo el patrón de la matriz sugiere tres componentes: uno de presencia de problemas, otro relacionado con riesgo y otro de reacciones positivas. Sin embargo, según Barkham, et al. (1998) los resultados encontrados corresponden a diferencias en pequeños detalles respecto a la muestra no clínica. Estos resultados corresponden a lo postulado por Lyne et al. (2006) quienes, después de explorar diversos modelos de agrupación, describen al CORE-OM como un cuestionario de estructura compleja, que puede comprenderse mejor a partir de una escala de riesgo (que incluye las dos subescalas de daño a sí mismo y a otros) y una escala general de malestar psicológico, que a su vez puede ser

analizada a partir de los ítems redactados positivamente y aquellos que fueron redactados en forma negativa, aceptando también las posibles diferencias en la agrupación de los ítems para la muestra no clínica.

Sensibilidad al cambio

Para probar las posibles diferencias en relación a la naturaleza de los problemas y las diferencias en el número típico de sesiones ofrecidas, el grado de cambio observado fue considerado en tres contextos distintos: consultas de atención primaria, asesoría estudiantil y un grupo “clínico” que incluyó psicoterapia proporcionada mediante el Servicio Nacional de Salud del Reino Unido y/o servicios de asesoría psicológica. Los resultados muestran mejoras sustanciales y estadísticamente significativas después de terminar la intervención psicológica en los tres contextos estudiados (Barkham, et al. 1998).

Cambio fiable y clínicamente significativo

El cambio fiable y clínicamente significativo complementa y amplía los análisis realizados mediante comparaciones de medias grupales. Los datos de referencia presentados proporcionan los puntos de corte que pueden utilizar los profesionales interesados en determinar la importancia clínica de las puntuaciones de cambio en un paciente dado (Evans et al. 1998; Jacobson, y Truax, 1991). A nivel de fiabilidad, este índice hace referencia al grado de cambio en una medida que se podría esperar por efecto del azar o por el error de medida; mientras que el cambio clínicamente significativo corresponde a una mejora suficiente para que el cliente se traslade de una puntuación más representativa de la población clínica a la de la población no clínica.

Utilizando valores del coeficiente alfa (0,94) fue posible calcular los criterios de cambio fiable, además de identificar el cambio en diversas categorías. Las tres posibles categorías de *cambio fiable* son: no fiable (tan pequeño que podría explicarse por el azar o el error de medida), mejora fiable y deterioro fiable (Evans et al. 2002).

Las cuatro categorías de *cambio clínicamente significativo* son: permanecer en el rango clínico, permanecer en el rango no clínico, cambiar del rango clínico al no clínico (mejora clínicamente significativa) y cambio del rango no clínico al rango clínico (deterioro clínicamente significativo). Los resultados del análisis de los datos de los tres servicios en los que se estudió la sensibilidad al cambio evidenciaron que la mayoría de los pacientes presentaron un cambio clínicamente significativo en el CORE-OM. Sin embargo, como es habitual, algunos pacientes puntuaron demasiado bajo al entrar en terapia como para mostrar una mejora clínicamente significativa, mientras que algunos puntuaron alto al inicio de la terapia y mejoraron de forma fiable pero no tanto como para que pudieran superar el punto de corte para entrar en el rango de mejora clínicamente significativa (Evans et al. 2002). Finalmente, unos pocos (7%) mostraron deterioro, un porcentaje

relativamente usual, pero que merecería un estudio específico caso a caso para comprenderlo.

En resumen, los resultados del estudio original mostraron que el CORE-OM es ampliamente aceptable, fiable y válido. En efecto, la fiabilidad interna y la estabilidad test-retest fueron altas y se encontró también una excelente validez convergente respecto a los instrumentos más populares con los cuales fue comparado. Así mismo, las puntuaciones evidencian diferencias significativas entre la muestra clínica y la no clínica para un amplio rango de datos. Respecto al género se observa un efecto pequeño pero estadísticamente significativo. La distribución clínica y no clínica y la fiabilidad proveen los parámetros pertinentes para calcular si se produce un cambio fiable y clínicamente significativo. Por lo que respecta a la estructura, el análisis de componentes principales no valida la existencia independiente de las escalas del CORE-OM sino que sugiere una alta relación entre sus ítems y una agrupación en un único factor general de malestar psicológico, del que se puede diferenciar la escala de riesgo que sigue otro patrón.

Corrección e interpretación

Las cuatro dimensiones del CORE-OM pueden ser identificadas por la letra contigua a la columna etiquetada como “uso profesional” en el extremo derecho de cada ítem. Por tanto, para obtener una puntuación total de alguna de las dimensiones, por ejemplo *Bienestar Subjetivo*, primero se escriben los valores de las respuestas en los cuadros asignados y luego se suman las puntuaciones de las cuatro casillas marcadas con una “W” y se escribe el resultado en la casilla marcada “W” en la parte baja del instrumento. Las puntuaciones medias para cada dimensión se calculan dividiendo la puntuación total por el número de respuestas completadas según los ítems de cada dimensión; en la escala de *Bienestar Subjetivo* por ejemplo, la puntuación normalmente se divide por cuatro. Si en esta dimensión se ha omitido un ítem, entonces la puntuación deberá ser dividida por tres y así con el resto de puntuaciones y dimensiones (Evans et al. 1998).

Ítems que evalúan riesgo

Estos ítems evalúan la presencia de ideación suicida, autolesiones o actos de agresión a terceros (Evans et al. 1998). Cuando cualquiera de los ítems de la escala de riesgo, marcados con una “R”, puntúe por encima de 0, el clínico debe prestar atención y contemplar la necesidad de abordar el tema con el cliente para valorar si es necesario tomar medidas al respecto. Para calcular la puntuación media total con excepción de los ítems de la escala de riesgo (“Todo menos R”), primero se debe sumar la puntuación total sin contar los ítems de la escala de riesgo y luego dividir este resultado por el número de respuestas a los ítems respondidos en las casillas ‘W’, ‘P’ o ‘F’ (normalmente 28).

Puntuación clínica y punto de corte

A partir de las dificultades en la utilización de números con decimales planteadas por algunos profesionales (normalmente las medias obtenidas incluyen decimales) y de la necesidad de asignarles significado clínico, se optó por multiplicar la puntuación media resultante por 10 y llamar a este resultado puntuación clínica. Ninguna de estas operaciones altera las propiedades psicométricas del instrumento, aunque la desviación estándar también necesita ser multiplicada por 10 (Barkham et al. 2006).

Así mismo se ha determinado la puntuación 10 como punto de corte (puntuación directa 1.0) entre población clínica y no clínica. Dentro del rango de puntuaciones no clínicas, se han establecido dos categorías: saludable y leve. El límite superior para los puntajes de la primera categoría es 5, asumiendo que una persona en un momento determinado puede aumentar su puntuación total y aún así continuar en el rango de *saludable*. En el mismo sentido, una persona podría puntuar *leve* como resultado del aumento del estrés o de situaciones particulares pero su puntuación se mantendría dentro del rango de puntuaciones de la población general. Se ha establecido la puntuación de 10 como el límite superior para el nivel *leve*, 15 para el nivel *medio*, 20 para el nivel *moderado* y 25 para un nivel de *moderado a severo*. Una puntuación por encima de 25 indicaría un nivel de *malestar severo* (Barkham et al. 2006).

Los datos provenientes de la combinación de una muestra por conveniencia y otra conformada por estudiantes del Reino Unido proporcionan un punto de corte de 11,9 para hombres y 11,29 para mujeres (1,19 y 1,29 en puntuaciones directas), aproximando a 12 y 13 respectivamente (Evans et al. 2002). Sin embargo, Connell et al. (2007) proponen aplicar la puntuación de 10 que separa población clínica de la población general tanto a hombres como a mujeres en aras de redondear y facilitar su aplicación rutinaria, a la vez que se evitan cálculos separados según el género.

Puntos de corte para el cambio fiable

Cuando se analiza el cambio del pre-tratamiento al postratamiento, los profesionales necesitan saber si el cambio conseguido supera el nivel de azar o error de medición. Para el CORE-OM se utiliza un índice de cambio fiable de 0,48 (aproximado a 0,50 con un puntaje clínico de 5). En este sentido, para estar seguro de que una persona ha realizado un cambio fiable, se esperan cambios superiores a 5 en la puntuación clínica (0,5 si se utiliza la puntuación promedio de ítems). Así pues, si una persona ha puntuado en el extremo superior, por ejemplo, moderado (puntuación = 19) y cumple el criterio de cambio fiable (mejora de 5 puntos por lo menos), entonces cambiará del nivel “moderado” al nivel inferior “leve”. En conclusión, con excepción de las puntuaciones del nivel severo, un cambio fiable se reflejará en una cambio (reducción) del nivel de gravedad (Barkham et al. 2006). Este punto de corte de cambio significativo junto con el punto de corte entre la

población clínica y no clínica permite valorar si el cambio realizado por un paciente es clínicamente significativo.

Adaptaciones en diversos países e idiomas

A partir del 2000 empezaron a aparecer versiones adaptadas a nuevos idiomas, primero en la India y poco después en Noruega. Hasta el momento se han llevado a cabo diferentes procesos de traducción y adaptación con resultados satisfactorios respecto a las propiedades psicométricas del cuestionario. En la versión eslovaca, por ejemplo, se llevaron a cabo tres estudios diferentes. En el primero se profundizó sobre la consistencia interna y la validez de constructo, además de las comparaciones respecto a las puntuaciones en muestras clínicas y no clínicas (Gampe, Biescad, Balúnová-Labanicová, Timulák, y Evans, 2007).

En el segundo estudio, se presentaron los datos descriptivos de las comparaciones entre muestras realizadas previamente; y adicionalmente, se llevó a cabo la prueba de fiabilidad test-retest y se determinaron los criterios de cambio estadístico y clínicamente significativo. Finalmente, en el último estudio se realizaron ajustes de la versión final en eslovaco, en colaboración con los autores de la versión original inglesa. El CORE-OM evidenció propiedades psicométricas favorables, con una alta consistencia interna, adecuada fiabilidad test-retest y buena discriminación entre las muestras clínica y no clínica (Gampe et al. 2007).

Del mismo modo, la versión italiana fue probada con 263 personas pertenecientes a la población no clínica y 647 de la muestra clínica. La traducción evidenció una adecuada aceptabilidad, consistencia interna y validez convergente en ambas muestras. Las diferencias fueron significativas en las dos poblaciones, los criterios de cambio y el resto de datos obtenidos fueron similares a los de la versión original, con excepción de criterio de cambio clínicamente significativo para el género masculino, que tuvo un punto de corte moderadamente diferente (Palmieri et al. 2009).

En Portugal, se han publicado resultados preliminares llevados a cabo en una muestra de población general de 111 personas, mostrando que la traducción realizada es fiable y presenta una adecuada consistencia interna (Dias Sales, Matos Moleiro, Evans, y Gomes Alves, 2011). Incluso en Japón se construyó una versión del CORE-OM que fue valorada en una población total de 1684 personas, con resultados similares respecto a la consistencia interna, fiabilidad test-retest, validez convergente, sensibilidad y discriminación entre las muestras clínicas y no clínicas. Por tanto, los autores proponen la utilización del cuestionario tanto en contextos clínicos como de investigación (Uji, Sakamoto, Adachi, y Kitamura, 2012).

Entre las publicaciones más recientes se encuentra la validación de la versión sueca del CORE-OM. El estudio incluyó una muestra clínica de 619 personas provenientes de servicios de atención primaria, mientras que la muestra no clínica estuvo compuesta por 229 estudiantes universitarios. Los resultados mostraron una excelente aceptación, alta consistencia interna y fiabilidad test-retest, y una adecuada validez convergente. Se identificaron claras diferencias entre las muestras clínicas y no clínicas, evidenciando sensibilidad al cambio en contextos de atención primaria. En todos los casos, los datos fueron similares a los presentados para la versión original (Elfström et al. 2012).

Actualmente se han realizado traducciones en 20 idiomas diferentes incluyendo el galés, español (presentada en este artículo), portugués, alemán, holandés, griego, italiano, danés, islandés, sueco, polaco, finlandés, lituano, eslovaco, japonés, turco, croata, albanés y el gujarati. Además de ello se encuentran en proceso de traducción y/o adaptación versiones en: Lengua de Signos Británica (BSL) portugués de Brasil, español de Argentina, rumano, búlgaro, árabe, francés, checo, Bangla/Bengalí/Sylheti, isixhosa, catalán, persa, kurdo, canarés, hindi, urdu, tamil, sami, el gaélico escocés, irlandés (gaélico irlandés), Mandarín (chino y otros idiomas). Para más información es posible visitar la página web del *CORE Information Management System* (<http://www.coreims.co.uk>).

Método

Para la traducción y adaptación del CORE-OM al español se siguieron los pasos establecidos por los grupos internacionales que se reflejan en un documento interno del *Core System Trust* (CST) en el que se indica la conveniencia de que un miembro del grupo fundador del instrumento participe en el proceso junto a los responsables locales. En general, dicho proceso es homologable a las pautas de la International Test Commission (2010), pero se enfatiza la importancia de traducir los ítems de acuerdo a su significado contextualizado en la cultura y el entorno en el que se va a aplicar y que sea comprensible para el rango más variado posible de potenciales usuarios. Se intenta evitar así una excesiva confianza en la traducción inversa por el riesgo a producir traducciones demasiado literales.

Participantes

Para la traducción de la versión original en inglés del CORE-OM, se contó con la colaboración de 12 personas de diferentes lugares de España, seleccionadas por tener un alto nivel de inglés, de los cuales 6 eran profesionales en psicología y 4 personas del común. Para la revisión del instrumento se incluyó un total de 64 personas de entre 16 y 76 años, de distintas condiciones y orígenes lingüísticos (véase tabla 2), plenamente competentes en español; de ellas 12 eran profesionales de la psicología y 52 eran personas legas.

Tabla 2
Descripción de las personas que revisaron y/o comentaron la versión revisada del CORE-OM en español

Origen	52 Personas legas			12 Profesionales de la Psicología		
	Edad	Sexo		Edad	Sexo	
		Mujer	Hombre		Mujer	Hombre
lingüístico						
Alicante	25	1	0	30	1	0
Colombia				29	1	0
Chile	46	1	0	43-48	1	1
Barcelona	16-76	6	6			
Galicia	41-54	0	2	46	1	0
Grecia	27	1	0			
Italia				32	1	0
Las Palmas	21	1	0			
Madrid				40	0	1
Málaga	26-54	2	1			
México				26-27	1	1
Polonia	36	1	0			
Salamanca				28	1	0
Sevilla	38-40	2	0	40	1	0
Tarragona	29	1	0			
Valladolid	22	1	0			
Zaragoza	18-28	17	9	53	0	1

Instrumentos

Clinical Outcomes in Routine Evaluation-Outcome Measure, Versión española (CORE-OM; Evans et al. 1998). El CORE-OM es un cuestionario de autoinforme que contiene 34 ítems que evalúa el estado del cliente a partir de una serie de dimensiones definidas como: *Bienestar subjetivo*, compuesta por 4 ítems; *Problemas/Síntomas*, con 12 ítems que valoran ansiedad, depresión, trauma y síntomas físicos; *Funcionamiento general*, que incluyen 12 ítems que evalúan relaciones de intimidad, relaciones sociales y nivel de funcionamiento cotidiano y la escala de *Riesgo*, compuesta por 6 ítems que son utilizados como indicadores clínicos de intentos de suicidio, autolesiones o actos de agresión a terceros (Lyne et al. 2006).

El instrumento está diseñado con base en una dimensión de problemas, por lo que a mayor puntuación mayor nivel de problemas o síntomas. La puntuación directa puede variar desde 0 hasta 136, este resultado se divide por el número de ítems respondidos obteniendo el resultado general del test. Es posible mediante la

misma operación obtener puntuaciones para cada una de las dimensiones. La puntuación de la escala *bienestar subjetivo* resulta ligeramente contradictoria (puesto que a mayor puntuación mayor *malestar*), pero los autores optaron por mantener el nombre de la escala, aunque la puntuación que se obtenga esté invertida, por criterios de consistencia con las demás puntuaciones (Evans et al. 1998).

Procedimiento

En primer lugar, se pidió a 12 personas seleccionadas por tener un buen nivel de inglés, residentes en distintos puntos de España, que realizarán una traducción al español de los ítems del CORE-OM. De los 10 que respondieron a la petición proporcionando una traducción, 6 eran expertos en psicología y 4 legos.

Con este material se convocó una sesión de trabajo en la que asistieron 2 de estos expertos y dos de los legos junto con uno de los miembros del CST (Chris Evans), que actuó como consultor/supervisor. En esta sesión se debatió cada ítem teniendo en cuenta las traducciones disponibles, discutiendo y escogiendo por consenso la mejor opción. De todo ello resultó una primera versión preliminar que fue revisada por tres expertos en psicología con más de 20 años de experiencia en el ámbito clínico quienes realizaron algunas modificaciones que fueron discutidas por correo electrónico con el miembro citado del CST. Esta versión revisada fue sometida a un escrutinio amplio por parte de un conjunto de 64 personas (los participantes arriba descritos) a quienes se les pidió que la leyeran atentamente y juzgaran si les resultaban comprensibles los ítems. Se les animó también a que hicieran todos los comentarios que consideraran oportunos a propósito de la forma en que estaban redactados.

Como es de suponer, con un conjunto tan variado de personas surgieron un cierto número de comentarios y observaciones que fueron debatidos por parte de los tres expertos citados, y las dudas se consultaron con el miembro del CST, hasta que se llegó a una versión revisada. Esta versión se entregó a una traductora completamente bilingüe inglés-español, con experiencia en traducción y licenciada en psicología, para que realizase la traducción inversa, sin tener acceso a la versión original. Ni los tres expertos españoles ni el miembro del CST consideraron necesario realizar ninguna modificación posterior a la versión que había resultado del proceso anterior, y se procedió a su aprobación definitiva por parte del CST.

Resultados

Como resultado del procedimiento anterior fue posible obtener la versión definitiva en español del CORE-OM. Los formularios incluyen el cuestionario para hombre y para mujer junto con las versiones más breves resultantes (CORE-SFA, CORE-SFB, CORE-10 y CORE-5). Todos ellos se pueden encontrar ahora en una página web de la Universitat de Barcelona destinada a tal efecto (<http://www.ub.edu/terdep/core>). De acuerdo con la política propia del CST, es posible descargar tales

formularios de forma gratuita en la forma exacta como se presentan, pero no se autoriza ningún tipo de modificación ni de los ítems ni de la forma de presentarlos.

Discusión

Mediante el proceso de traducción realizada y los lineamientos establecidos por el documento interno del *Core System Trust* (CST) fue posible obtener la versión en español del CORE-OM. Ahora se hace necesario continuar con el proceso de adaptación y validación del instrumento aplicándolo en un conjunto representativo de personas a partir de las cuales sea posible obtener información sobre las propiedades psicométricas del instrumento en su versión en español. Una vez sea posible contar con la versión estandarizada de este cuestionario (estudio en curso) se podrán plantear y desarrollar estudios en los que se incluyan las formas derivadas del CORE-OM, estudios transculturales, aplicaciones en diversas poblaciones y centros de atención en salud mental, y en consecuencia la oportunidad de integrar y analizar los datos mediante las posibilidades ofrecidas por el Sistema Core. También sería interesante promover versiones del CORE-OM en otras lenguas como el catalán, euskera y gallego para aquellos usuarios que las tengan como lengua materna.

En suma, la continuidad de este estudio someterá a prueba la versión española del CORE-OM como una herramienta válida para la evaluación del malestar psicológico—con la correspondiente retroalimentación para el paciente y el terapeuta— y la obtención de evidencias sobre la efectividad y eficacia de las modalidades de intervención propuestas por centros e instituciones que ofrecen servicios de terapia psicológica (Mellor Clark et al. 2007). Por tanto, contribuirá a la consolidación del movimiento de terapia basada en la evidencia en la fase de desarrollo de herramientas para la evaluación de los efectos terapéuticos, la monitorización de los criterios propuestos por el modelo y el fortalecimiento del compromiso ético en la comprobación de la mejora sintomática y la evaluación de los resultados al finalizar el proceso terapéutico.

Nota:

Este estudio forma parte del proyecto “Terapia cognitiva para la depresión unipolar: eficacia de una intervención centrada en dilemas” (ref. PSI2011-23246; véase en Internet: www.ub.edu/terdep) de la Universitat de Barcelona, financiado en parte por el Ministerio de Economía y Competitividad.

Referencias bibliográficas

- Bados, A., García-Grau, E. y Fusté, A. (2002). Eficacia y utilidad clínica de la terapia psicológica. *Revista Internacional de Psicología Clínica y de la Salud/International Journal of Clinical and Health Psychology*, 2, 477-502.
- Barkham, M., Bewick, B.M., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Unsworth, G., Richards, D., y Evans, C. (2012). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research*, publicado online. doi:10.1080/14733145.2012.729069
- Barkham, M., Culverwell, A., Spindler, K. y Twigg, E. (2005). The CORE-OM in an older adult population: Psychometric status, acceptability, and feasibility. *Aging and Mental Health*, 9, 235-245.
- Barkham, M., Evans, C., Margison, F., McGrath, G., Mellor-Clark, J., Milne, D. y Connell, J. (1998). The rationale for developing and implementing core batteries in service settings and psychotherapy outcome research. *Journal of Mental Health*, 7, 35-47.
- Barkham, M., Gilbert, N., Connell, J., Marshall, C. y 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. *British Journal of Psychiatry*, 186, 239-246.
- Barkham, M., Hardy, G. E. y Mellor-Clark, J. (2010). *Developing and delivering practice-based evidence: a guide for the psychological therapies*. Chichester, UK; Malden, MA: Wiley-Blackwell.
- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K. y McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69: 184-196.
- Barkham, M., Mellor-Clark, J., Connell, J. y Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling and Psychotherapy Research*, 6 (1) 3-15.
- Barkham, M., Startup, M. y Hardy, G. E. (1996). The IIP-32: development a short version of the Inventory of Interpersonal Problems. *British Journal of Clinical Psychology*, 35, 21-35.
- Beck, A. T., Epstein, N., Brown, G. y Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 56, 893-897.
- Beck, A. T., Steer, R. A. y Brown, G. K. (1996). *Manual for the Beck Depression Inventory-Second edition (BDI-II)*. Texas, USA: Psychological Corporation.
- Beck, A. T. Ward, C. H., Mendelson, M., Mock, J. y Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Bedford, A., Lukic, G. y Tibbles, J. (2011). Evaluation of risk by patients' and with clinicians' ratings: A CORE-OM and CORE-A investigation. *Clinical Psychology and Psychotherapy*, 18, 244-249.
- Beutler, L. E. (2001). Comparisons of quality assurance systems: From outcome assessment to clinical utility. *Journal of Consulting and Clinical Psychology*, 69, 197-204.
- Botella, L. (2006). Routine evaluation in psychotherapy service: The use of CORE System data. *European Journal of Psychotherapy and Counselling*, 8, 235-241.
- Botella, L. y Feixas, G. (1994). Eficacia de la psicoterapia: Investigaciones de resultados. En M. Garrido y J. García (Eds.), *Psicoterapia: Modelos contemporáneos y aplicaciones*. Valencia: Promolibro.
- Cahill, J., Barkham, M., Stiles, W. B., Twigg, E., Hardy, G. E., Rees, A. y Evans, C. (2006). Convergent validity of the CORE measures with measures of depression for clients in cognitive therapy for depression. *Journal of Counseling Psychology*, 53, 253-259
- Cavanagh, K., Seccombe, N. y Lidbetter, N. (2011). The Implementation of Computerized Cognitive Behavioural Therapies in a Service User-Led, Third Sector Self Help Clinic. *Behavioural and Cognitive Psychotherapy*, 39, 427-442.
- Connell, J., Barkham, M. y Mellor-Clark, J. (2007). Mental health CORE-OM norms of students attending university counselling services benchmarked against an age-matched primary care sample. *British Journal of Guidance and Counselling*, 35, 41-56.
- Connell, J., Barkham, M., Stiles, W.B., Twigg, E., Singleton, N., Evans, C. y Miles, J. V. (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *British Journal of Psychiatry*, 190, 69-74.
- Conte, H. R., Plutchik, R., Karasu, T. B. y Jerrett, I. (1980). A self-report borderline scale. *Journal of Nervous and Mental Diseases*, 168, 428-435.

- Corbella, S. y Botella, L. (2004). *Investigación en Psicoterapia: Proceso, Resultado y Factores Comunes*. Madrid: Visión Net.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of test. *Psychometrika*, 16, 297-334.
- Derogatis, L. R. (1983). *SCL-90-R: Administration, scoring and procedures: Manual II*. Towson, MD: Clinical Psychometric Research.
- Derogatis, L. R., Lipman, R. S., y Covi, L. (1973) SCL-90: An out-patient psychiatric rating scale-preliminary report. *Psychopharmacological Bulletin*, 9, 13-17.
- Derogatis, L. R. y Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, 13, 595-605.
- Dias Sales, M. C., de Matos Moleiro, C. M., Evans, C. y Gomes Alves, P. C. (2011). The Portuguese version of the CORE-OM: translation, adaptation, and preliminary study of psychometric properties. *Revista de Psiquiatria Clínica*, 39, 54-59.
- Elfström, M. L., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M. y Carlsson, S. G. (2012). Validation of the Swedish Version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychological and Psychotherapy*, publicado online, doi: 10.1002/cpp.1788.
- Evans, C. (in press). The CORE-OM (Clinical Outcomes in Routine Evaluation) and its derivatives. *Science and Practice*.
- Evans, C. (2012). Cautionary notes on power steering for psychotherapy. *Canadian Psychology/Psychologie Canadienne*, 53, 131-139. doi:10.1037/a0027951
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J. y Audin, J. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51-60.
- Evans, C., Connell, J., Barkham, M., Mellor-Clark, J., Margison, F., McGrath, G. y Audin, K. (1998). *The CORE outcome measure: user's manual (version 2.1)*. Leeds: Core System Group
- Evans, C. E., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J. y McGrath, G. (2000). CORE: Clinical outcome in routine evaluation. *Journal of Mental Health*, 9, 247-255.
- Froyd, J. E., Lambert, M. J. y Froyd, J. D. (1996). A review of practices of psychotherapy outcome measurement. *Journal of Mental Health*, 5, 11-15.
- Gampe, K., Biescad, M., Balúnová-Labanicová, L., Timulák, L. y Evans, C. (2007). Slovak Adaptation of CORE-OM (Clinical Outcomes in Routine Evaluation-Outcome Measure). *Ěes. a Slovak Psychiatry*, 103, 4-13
- Gardner, M. y Altman, D. G. (1986). Confidence intervals rather than P values: estimation rather than hypothesis testing. *British Medical Journal*, 292, 746-750
- Goldberg, D. (1978). *Manual of the General Health Questionnaire*. Oxford: NFER-NELSON.
- Goldberg, D. P. y Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine*, 9, 139-145.
- Goldberg, D y Williams, P. (1988). *A user's guide to the General Health Questionnaire*. Oxford: NFER-NELSON.
- Hyder, S. E. y Rieder, R. C. (1987). *PDQ-R Questionnaire*. New York: New York State Psychiatric Institute.
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureno, G. y Villasenor, V. S. (1988). Inventory of the interpersonal problems: psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, 56, 885-892.
- International Test Commission (2010). *International Test Commission Guidelines for Translating and Adapting Tests*. [Página web]. Recuperado de <http://www.intestcom.org/upload/sitefiles/40.pdf>
- Jacobson, N. S. y Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.
- Leach, C., Lucock, M., Barkham, M., Noble, R., Clarke, L. y Iveson, S. (2005). Assessing risk and emotional disturbance using the CORE-OM and HoNOS outcome measures. *Psychiatric Bulletin*, 29, 419-422.
- Lyne, K., Barrett, P., Evans, C. y Barkham, M. (2006). Dimensions of variation on the CORE-OM. *British Journal of Clinical Psychology*, 45, 185-203.
- Margison, F.R., McGrath, G., Barkham, M., Mellor-Clark, J., Audin, K., Connell, J. y Evans, C. (2000). Measurement and psychotherapy: Evidence-based practice and practice-based evidence. *British Journal of Psychiatry*, 177 123-130.
- Mellor-Clark, J. y Barkham, M. (2000). Quality evaluation: methods, measures and meaning. En C. Feltham y I. Horton (Eds.), *Handbook of counselling and psychotherapy* (pp. 255-270). London, UK: Sage Publications.
- Mellor-Clark, J., Barkham, M., Connell, J. y Evans, C. (1999). Practice-based evidence and the need for a standardised evaluation system: Informing the design of the CORE System. *European Journal of Psychotherapy, Counselling and Health*, 3, 357-374.

- Mellor-Clark, J., Barkham, M., Connell, J. y Evans, C. (2007). Practice-based evidence and standardized evaluation: Informing the design of the CORE system. *European Journal of Psychotherapy and Counselling*, 2, 357-374.
- Mellor-Clark, J., Curtis Jenkins, A., Evans, R., Mothersole, G. y McInnes, M. (2006). Resourcing a CORE Network to develop a National Research Database to help enhance psychological therapy and counselling service provision. *Counselling and Psychotherapy Research*, 6, 16-22.
- Palmieri, G., Evans, C., Hansen, V., Brancaleoni, G., Ferrari, S., Porcelli, P., Reitano, F. y Rigatelli, M. (2009). Validation of the Italian version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology and Psychotherapy*, 16, 444-449.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Sinclair, A., Barkham, M., Evans, C., Connell, J. y Audin, K. (2005). Rationale and development of a general population well-being measure: Psychometric status of the GP-CORE in a student sample. *British Journal of Guidance and Counselling*, 33, 153-174.
- Snaith, R. P., Constantopoulos, A. A., Jardine, M. Y. y McGuffin, P. (1978). A clinical scale for the self-assessment of the irritability. *British Journal of Psychiatry*, 132, 164-171.
- Strupp, H. H., Horowitz, L. M. y Lambert, M. J. E. (1997). *Measuring patient changes in mood, anxiety, and personality disorders: Toward a core battery*. Washington DC: American Psychological Association.
- The DoH (Departamento de Salud del Reino Unido), (1996). *NHS Psychotherapy Services in England: Review of Strategic Policy*. Londres, UK: HMSO
- Tranter, R., Bell, D., Gutting, P., Harmer, C., Healy, D. y Anderson, I.M. (2009). The effect of serotonergic and noradrenergic antidepressants on face emotion processing in depressed patients. *Journal of Affective Disorders*, 118, 87-93.
- Twigg, E., Barkham, M., Bewick, B. M., Mulhern, B., Connell, J. y Cooper, M. (2009). The YP-CORE: Development and validation of a young person's version of the CORE-OM. *Counselling and Psychotherapy Research*, 9, 160-168.
- Uji, M., Sakamoto, A., Adachi, K. y Kitamura, T. (2012). Psychometric properties of the Japanese version of the Clinical Outcomes in Routine Evaluation-Outcome Measure. *Comprehensive Psychiatry*, 53, 600-608.
- Unsworth, G., Cowie, H. y Green, A. (2011). Therapists' and clients' perceptions of routine outcome measurement in the NHS: A qualitative study. *Counselling and Psychotherapy Research*, 12, 71-80.
- Waskow, I. E. (1975). Selection of a core battery. En M. B. Parloff (Ed.), *Psychotherapy change measures* (pp. 245-269). Washington, DC: U.S. Government Printing Office.
- Zigmond, A. S. y Snaith, R. P. (1983) The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67, 361-370.