

EL TRATAMIENTO PSICOLÓGICO DEL TRASTORNO OBSESIVO-COMPULSIVO EN EUROPA: UN ESTUDIO META-ANALÍTICO*

Julio Sánchez Meca, Antonia Gómez Conesa
y F. Xavier Méndez Carrillo
Universidad de Murcia (España)

Resumen

En este trabajo se presentan los resultados de un meta-análisis sobre la eficacia de las intervenciones psicológicas, solas o en combinación con psicofármacos, en el trastorno obsesivo-compulsivo. Se incluyeron 23 artículos europeos que cumplieron con los criterios de selección, dando lugar a 43 estudios independientes, sobre los que se calculó la diferencia media tipificada entre las medias del pretest y del postest. El tamaño del efecto medio global, $d_+ = 1,443$, reflejó una clara eficacia para disminuir las obsesiones/compulsiones, así como los síntomas de depresión, ansiedad y ajuste social, aunque éstos en menor medida. Los tratamientos más eficaces consistieron en la combinación de técnicas de exposición con prevención de respuesta o de reestructuración cognitiva con antidepresivos ($d_+ = 2,044$ y $d_+ = 2,953$, respectivamente), tales como la clomipramina o la fluvoxamina. Se propone un modelo predictivo de la eficacia, en función de los tipos de tratamientos y de la calidad metodológica de los estudios. Finalmente, se discuten las implicaciones prácticas y clínicas de los resultados, así como para la investigación futura en este campo.

PALABRAS CLAVE: Trastorno obsesivo-compulsivo, tratamientos psicológicos, terapia conductual, tratamiento farmacológico, metaanálisis.

Abstract

This article discusses the results of a meta-analysis of the effectiveness of psychological treatment, by itself or in combination with drugs, of the obsessive-compulsive disorder. Twenty-three European articles meeting the selection criteria were included, offering a total of 43 independent studies. Standardized mean difference was calculated between the pretest and posttest means. The global mean effect size, $d_+ = 1.443$, showed a clear efficacy for reducing obsessions and

Correspondencia: Julio Sánchez Meca, Dpto. Psicología Básica y Metodología, Facultad de Psicología, Campus de Espinardo, Apdo. 4021, 30100 Murcia (España). E-mail: jsmeca@um.es

* Esta investigación ha sido financiada por el Programa Sectorial de Promoción General del Conocimiento del Ministerio de Educación y Cultura (Proyecto Nº PB98-0401).

compulsions as well as symptoms of depression, anxiety, and social adjustment, although the latter was reduced to a lesser extent. The most effective treatments consisted of combining exposition and response prevention techniques or cognitive restructuring with antidepressants ($d_+ = 2.044$ and $d_+ = 2.953$, respectively), such as clomipramine or fluvoxamine. A predictive model of the efficacy is proposed as a function of the different treatments and the methodological quality of studies. Finally, the practical, clinical, and research implications of the results are discussed.

KEY WORDS: *Obsessive-compulsive disorder, psychological treatments, behavior therapy, pharmacologic treatment, meta-analysis.*

Introducción

Aunque hace algún tiempo se consideraba un trastorno poco frecuente, el trastorno obsesivo-compulsivo (TOC) es actualmente el cuarto trastorno psiquiátrico más común después de las fobias, el abuso de sustancias tóxicas y la depresión mayor (Reiger, Narrow y Raye, 1990). Se han reportado tasas de prevalencia del TOC que se sitúan entre el 1% y el 4% (Douglass *et al.*, 1995; Karno y Golding, 1991; Rasmussen y Eisen, 1990; Weisman, Bland, Canino y Greenwald, 1994) y se estima que éstas van en aumento (Stoll, Tohen y Baldessrini, 1992). Se trata de un trastorno que suele iniciarse en las primeras edades adultas, si bien los sujetos afectados suelen demorar unos siete años la búsqueda de ayuda profesional (Yaryura-Tobias, 1997), lo cual dificulta su tratamiento.

Según los criterios diagnósticos del DSM-IV (American Psychiatric Association, 1994), el TOC se caracteriza por la presencia de obsesiones recurrentes (pensamientos, imágenes o ideas intrusivas y persistentes) vividas por el sujeto con gran malestar, que provocan conductas compulsivas dirigidas a reducir la ansiedad generada por las obsesiones. De entre las obsesiones y compulsiones más comunes destacan las ideas de contaminación o suciedad, con compulsiones de lavado y limpieza, ideas obsesivas sobre posibles accidentes (dejarse el gas encendido, la puerta abierta, etc.), con compulsiones de comprobación e ideas de daños físicos (miedo a hacer daño a otros o a sí mismo), con compulsiones de reaseguración (Cruzado, 1997; Vallejo, 2001). Un pequeño porcentaje de los casos, en torno al 13%, presenta obsesiones puras, sin compulsiones (Ball, Baer y Otto, 1996). La gravedad de este trastorno se agudiza por la frecuente presencia de psicopatologías asociadas, en especial la depresión.

El TOC era considerado el trastorno de ansiedad más resistente al tratamiento, especialmente su variante de «obsesiones puras» sin compulsiones (Caballo y Mateos, 2000; Stanley y Turner, 1995). Pero tras 20 años de investigación desarrollada en paralelo, y sin a penas interconexión desafortunadamente, desde el enfoque cognitivo-conductual, por una parte, y el psicofarmacológico, por otra, el panorama ha cambiado. Desde el enfoque cognitivo-conductual las técnicas más investigadas han sido las de exposición con prevención de respuesta, EPR (Foa, Steketee y Milby, 1980) y más recientemente se ha puesto a prueba la reestructuración cognitiva (RC).

Desde el modelo biomédico se ha probado la eficacia de antidepresivos serotoninérgicos, en especial de la clomipramina, y más recientemente de inhibidores selectivos de la recaptación de la serotonina (ISRS) tales como la fluvoxamina, la fluoxetina y la sertralina (Van Balkom *et al.*, 1994).

Tanto las revisiones y estudios empíricos (Caballo y Mateos, 2000; Cruzado, 1997; Rosa, Olivares y Sánchez, 1999; Vallejo, 2001) como los estudios meta-analíticos (Abramowitz, 1996, 1997, 1998; Christensen, Hadzi-Pavlovic, Andrews y Mattick, 1987; Van Balkom *et al.*, 1994) parecen estar de acuerdo en que las técnicas de EPR son, en general, el tratamiento de elección, si bien los antidepresivos serotoninérgicos han demostrado también su eficacia (Cox, Swinson, Morrison y Lee, 1993; Piccinelli, Pini, Bellantuono y Wilkinson, 1995). Sin embargo, sólo recientemente se han publicado unos pocos estudios sobre la eficacia de otros tratamientos psicológicos, como la RC, así como de la combinación de tratamientos psicológicos con farmacológicos.

El objetivo de nuestro trabajo fue, pues, realizar un estudio meta-analítico de la eficacia diferencial de los tratamientos psicológicos, solos o en combinación con tratamientos psicofarmacológicos, sobre el TOC con objeto de integrar las últimas investigaciones. Nuestro meta-análisis no pretendía ser una mera réplica de los ya publicados. En primer lugar, limitamos el período de búsqueda desde el año 1980 para hacerlo coincidir con la aparición de los criterios diagnósticos del DSM y, al mismo tiempo, solaparnos lo menos posible con los meta-análisis previos. En segundo lugar, limitamos los estudios al contexto europeo con objeto de lograr una mayor homogeneidad en las características de los tratamientos y los sujetos. Así, al basarnos en nuevas evidencias podemos poner a prueba la robustez y la consistencia de los resultados de los meta-análisis anteriores sobre el tratamiento del TOC. Y en tercer lugar, aplicamos técnicas de meta-análisis más potentes que ponderan por la inversa de la varianza de cada estudio y examinamos la consistencia de los resultados desde dos modelos estadísticos: los modelos de efectos fijos y de efectos aleatorios, algo que ningún meta-análisis previo sobre el TOC ha hecho.

Método

Selección de los estudios

Para que un estudio fuera incluido en el meta-análisis tenía que cumplir con los siguientes criterios: (a) estar publicado entre 1980-2001; (b) haber sido realizado en algún país europeo; (c) haber aplicado procedimientos diagnósticos reconocidos internacionalmente (DSM y/o CIE, en cualquiera de sus versiones), con objeto de asegurar el carácter clínico de la población adulta estudiada; (d) aplicar algún tratamiento psicológico, con o sin el uso concomitante de psicofármacos; (e) incluir medidas pretest y posttest, y algún seguimiento opcionalmente; (f) ser un diseño de grupo/s, excluyéndose los diseños de caso único; (g) aportar datos estadísticos suficientes para el cálculo del tamaño del efecto, y (h) estar escrito en inglés, castellano o francés.

Se utilizaron fuentes formales e informales como procedimientos de búsqueda de los estudios. En primer lugar, se realizaron búsquedas computerizadas en las bases PsycLit, MedLine, PsycDoc y CSIC —CINDOC—, comprendiendo los años 1980-2001, con la palabra-clave «obsessive-compulsive» buscada en el título, en el abstract y en las palabras-clave, y excluyendo las direcciones de contacto de los principales países no europeos (USA, Canadá, Australia, etc.). En segundo lugar, se escribieron cartas a investigadores de reconocido prestigio en el ámbito europeo solicitándoles el envío de trabajos, publicados o no. En tercer lugar, se revisaron las referencias de los meta-análisis ya publicados sobre el tema (Abramowitz, 1996, 1997, 1998; Ball *et al.*, 1996; Christensen *et al.*, 1987; Cox *et al.*, 1993; Modigh, 1987; Piccinelli *et al.*, 1995; Ruhmland y Margraf, 2001; Van Balkom *et al.*, 1994). Finalmente, se revisaron las referencias de los estudios recuperados.

El proceso de búsqueda supuso la consulta de más de 1300 referencias, que nos permitió identificar y recuperar un total de 23 artículos que cumplieron con los criterios de selección, lo que supuso un total de 43 estudios independientes, 40 de los cuales fueron grupos de tratamiento y los 3 restantes grupos de control, siendo 1,9 el número medio de grupos por artículo. Todos los artículos seleccionados estaban escritos en inglés y procedían de cuatro países europeos: Holanda (48,8%), Gran Bretaña (35%), Francia (11,6%) y Alemania (4,6%). En total, el meta-análisis recogió datos de 923 participantes en el pretest (890 tratados y 33 de control) que, con una mortalidad media en torno al 15,6%, se redujeron a 779 sujetos en el postest (753 tratados y 26 de control) y a 340 sujetos en el seguimiento más largo (todos sujetos tratados). Los tamaños muestrales de los grupos fueron, en promedio, de 20 sujetos en el pretest, 15 en el postest y 12 en el seguimiento, con un rango que fue de 3 a 50 sujetos en el postest.

El grado de solapamiento con los meta-análisis previos no es elevado. Los porcentajes de artículos de nuestro meta-análisis que ya fueron incluidos en los previos fueron: 52% (Abramowitz, 1996), 17% (Abramowitz, 1997), 13% (Christensen *et al.*, 1987), 8,7% (Cox *et al.*, 1993) y 52% (van Balkom *et al.*, 1994). Los meta-análisis de Abramowitz (1996) y Van Balkom *et al.* (1994) son los que mayor grado de solapamiento tienen con el nuestro (52%). No obstante, el primero de ellos se centró exclusivamente en la eficacia de las técnicas EPR, mientras que el nuestro incluyó otras técnicas psicológicas, solas o en combinación con psicofármacos. Y el segundo no está limitado a Europa como el nuestro, sino que incluye también estudios norteamericanos. Por otra parte, 8 de los 23 artículos (el 34,8%) de nuestro meta-análisis no han sido incluidos en ningún meta-análisis previo. En consecuencia, nuestro meta-análisis no es una réplica de ninguno de los ya publicados y, por las razones expuestas, aporta nuevas evidencias en este campo.

Codificación de los estudios

Con objeto de examinar las características de los estudios que pudieran dar cuenta de la heterogeneidad de los resultados, se aplicó sobre cada estudio un protocolo de codificación que incluyó variables moderadoras de distinta naturaleza, dis-

tinguiendo entre variables sustantivas (de tratamiento, de participantes y de contexto), variables metodológicas y variables extrínsecas (Lipsey, 1994)¹.

Para caracterizar las intervenciones psicológicas se codificaron: (a) la técnica de intervención (EPR, RC y otras técnicas psicológicas) solas o en combinación; (b) el hecho de que se controlara o no el consumo de psicofármacos para el TOC; (c) la homogeneidad del tratamiento; (d) la inclusión de tareas para casa; (e) la utilización de programa de seguimiento una vez finalizada la intervención; (f) el uso de agentes externos; (g) el modo de intervención (directo o indirecto); (h) la provisión de consentimiento informado a los pacientes; (i) el tipo de entrenamiento (grupal, individual o mixto); (j) el soporte de entrenamiento (oral, escrito o mixto); (k) el tipo de soporte oral (directo, diferido o mixto); (l) la duración de la intervención (nº de semanas); (m) la intensidad de la intervención (nº de horas por semana que recibió cada participante) y (n) la magnitud de la intervención (nº total de horas recibidas por cada participante). Dentro de las técnicas de exposición (todas ellas incluían prevención de respuesta) se clasificaron éstas según: (ñ) el tipo de exposición (*in vivo*, en imaginación o mixto) y (o) el modo de la exposición (guiada por terapeuta, auto-exposición, auto-exposición asistida o mixto). Con respecto a las características de los terapeutas se codificaron las siguientes variables: (p) el número de terapeutas; (q) el hecho de que coincidieran con los investigadores; (r) su formación (psicólogo, psiquiatra, enfermero o mixto); (s) la experiencia de los terapeutas (alta, media, baja) y (t) el género de los terapeutas (hombres, mujeres, mixto). Respecto de los tratamientos psicofarmacológicos se codificaron: (u) el tipo de antidepressivo (fluvoxamina o clomipramina); (v) el control del consumo de otros fármacos; (w) la duración del tratamiento (nº de semanas) y (x) la dosis media (en mlgr/día).

En cuanto a las características de las muestras de participantes se codificaron: (a) la edad media de la muestra (en años); (b) el género de la muestra (% de hombres); (c) la historia del TOC (nº medio de años de padecimiento del problema en la muestra); (d) el hecho de haber estado anteriormente bajo tratamiento, y (e) la presencia de comorbilidad en la muestra (básicamente, depresión). En lo que respecta a las características contextuales se codificaron: (a) el país y (b) el régimen de la intervención, distinguiendo entre los estudios en los que los participantes fueron ingresados en un hospital o clínica y los que no lo fueron.

Dentro de las características metodológicas se incluyeron: (a) el tipo de diseño (pretest-postest de un solo grupo, pretest-postest con grupos no equivalentes y pretest-postest con grupos equivalentes); (b) la asignación aleatoria a los grupos de intervención; (c) el tipo de grupo de control, cuando lo hubo, distinguiendo entre control activo (placebo de atención o píldora placebo) y control inactivo o puro (lista de espera); (d) el modo de reclutamiento de los participantes (a instancias del experimentador, a petición propia o mixto); (e) el seguimiento más largo (en meses); (f) la mortalidad experimental en el postest; (g) la mortalidad experimental en el seguimiento y (h) la calidad metodológica del diseño, medida en una escala de 0 a 9 puntos, consistente en la suma de las puntuaciones obtenidas en nueve ítems de calidad

1 Puede solicitarse al primer autor el manual de codificación de las variables moderadoras, así como toda la base de datos meta-analítica.

(asignación aleatoria, tipo de diseño, tamaño muestral, mortalidad experimental, inclusión de algún seguimiento, ceguera experimental, reporte de todas las medidas pretest-postest, uso de escalas normalizadas y homogeneidad de la intervención). Por último, se registraron las siguientes variables extrínsecas: (a) la formación del primer autor (psicólogo, psiquiatra o ambos); (b) la fecha de publicación y (c) la fuente de publicación (publicado vs. no publicado).

El proceso de codificación de las variables moderadoras implicó la toma de decisiones en algunos casos complejas por la falta de información en los estudios. Con objeto de examinar la adecuación del proceso, se realizó un análisis de la fiabilidad de la codificación (Orwin, 1994). Para ello, dos investigadores independientes codificaron un subconjunto de los estudios (el 65%), alcanzándose en general índices de fiabilidad satisfactorios (proporciones de acuerdo por encima de 0,80 para las variables cualitativas y coeficientes de correlación de Pearson por encima de 0,90 para las variables cuantitativas)². Los desacuerdos entre los codificadores se resolvieron por consenso y el manual de codificación se corrigió cuando la causa de las inconsistencias se debió a un defecto del mismo.

Cálculo del tamaño del efecto

Se utilizó como índice del tamaño del efecto la diferencia media tipificada (Hedges y Olkin, 1985). Pero, dada la práctica ausencia de grupos de control en los estudios recuperados, no fue posible definirla como la diferencia entre las medias de los grupos de tratamiento y control. En su lugar, se optó por definir como unidad de análisis el grupo y se calculó la diferencia media tipificada entre las medias del pretest y del postest, dividida por la desviación típica del pretest; es decir:

$$d = c(m) \frac{\bar{y}_{Pre} - \bar{y}_{Pos}}{S_{Pre}},$$

siendo $c(m) = 1 - 3/(4n - 5)$. Valores positivos indicaron una mejora del grupo, mientras que valores negativos supusieron un empeoramiento de los participantes. El índice d se obtuvo a partir de las medias y desviaciones típicas reportadas en los estudios, o bien a partir de los resultados de las pruebas estadísticas aplicadas (Glass, McGaw y Smith, 1981; Sánchez y Ato, 1989). Cuando el estudio aportó alguna medida de seguimiento también se calculó el índice d comparando la media del pretest con la media del seguimiento más largo.

2 Las variables cualitativas que no alcanzaron el criterio del 80% de acuerdo fueron: la coincidencia de los investigadores con los terapeutas ($p = 0,75$), la experiencia de los terapeutas ($p = 0,78$), la ceguera experimental ($p = 0,64$) y el modo de reclutamiento de los pacientes ($p = 0,18$). Las variables cuantitativas que no alcanzaron una correlación de 0,90 fueron: la intensidad de la intervención ($r = 0,767$), la mortalidad experimental en el postest ($r = 0,689$) y la puntuación global de calidad metodológica del estudio ($r = 0,771$). En la mayoría de los casos, las inconsistencias entre los codificadores se debieron a la escasa información contenida en los estudios y a errores de aplicación del manual de codificación, más que a deficiencias en la definición de las variables moderadoras.

Los estudios presentaron de forma rutinaria múltiples variables de resultado. Para cada una de ellas se calculó el índice d y, a continuación, se promediaron en función de los siguientes criterios de clasificación: (a) por el tipo de constructo psicológico medido (síntomas obsesivos y compulsivos, ansiedad, depresión, ajuste social y otros) y (b) por el modo de registro (autoinformes e informes del terapeuta o de un asesor independiente). En consecuencia, de cada estudio pudieron obtenerse hasta un total de 18 índices d en el postest y otros 18 índices en el seguimiento que, al ser dependientes entre sí, se analizaron por separado.

Análisis estadístico

Se aplicaron los análisis estadísticos propuestos por Hedges y Olkin (1985; Cooper y Hedges, 1994; Sánchez y Ato, 1989), según los cuales el peso específico que debe ejercer cada estudio en los análisis tiene que ser proporcional a su precisión. El factor de ponderación de cada estudio se define, pues, como la inversa de la varianza del índice d : $w = 1/Var(d)$, siendo la varianza de d (Morris, 2000):

$$Var(d) = \left(\frac{n-1}{n(n-3)} \right) (1 + nd^2) - \frac{d^2}{[c(m)]^2}.$$

Partiendo de este esquema de ponderación, se calculó el tamaño del efecto medio global para todos los estudios en función del tipo de constructo medido y del tipo de registro, así como intervalos de confianza y pruebas de homogeneidad. Estos promedios se obtuvieron también para las diferentes técnicas de intervención y sus combinaciones. Así mismo, se analizó el efecto de las diferentes variables moderadoras mediante análisis de varianza y análisis de regresión por mínimos cuadrados ponderados. Con objeto de examinar la consistencia de los análisis, se aplicaron dos modelos de análisis: el modelo de efectos fijos y el de efectos aleatorios³. Todos los análisis se realizaron con los programas SPSS 11.0 (2001) y MetaWin 2.0 (Rosenberg, Adams y Gurevitch, 2000).

3 Los modelos de efectos fijos y aleatorios difieren en la concepción de la población de estudios de partida. En el modelo de efectos fijos se asume que los estudios incluidos en el meta-análisis están estimando a un mismo, y único, tamaño del efecto paramétrico, d , por lo que la única variabilidad asumida es la debida a error de muestreo aleatorio, $Var(d)$, es decir, al hecho de que los estudios utilizan muestras de sujetos diferentes. En el modelo de efectos aleatorios se asume que los estudios estiman a una distribución de tamaños del efecto paramétricos en la población, que sigue una ley normal $N(m_d; s_d^2)$ por lo que, además de la variabilidad debida al error de muestreo, $Var(d)$, hay que contemplar también la variabilidad inter-estudios, s_d^2 . En consecuencia, mientras que en el modelo de efectos fijos el factor de ponderación de cada índice d viene determinado exclusivamente por la varianza intra-estudio, $w = 1/Var(d)$, en el de efectos aleatorios el factor de ponderación se define como $w = 1/[Var(d) + s_d^2]$. Las consecuencias de asumir uno u otro modelo afectan al grado de generalización de los resultados del meta-análisis. En el modelo de efectos fijos, la generalización se limita a la población de estudios de características similares a los incluidos en el meta-análisis. En el modelo de efectos aleatorios, por el contrario, los resultados pueden generalizarse a una población mayor de posibles estudios (Hedges y Vevea, 1998).

Resultados

De los 43 estudios meta-analizados, tres fueron grupos de control (dos activos: Lovell, Marks, Noshirvani y O'Sullivan, 1994 y Marks, *et al.*, 2000; y uno inactivo: Van Balkom *et al.*, 1998), uno incluyó sólo tratamiento psicofarmacológico (Cottraux *et al.*, 1990) y los 39 estudios restantes aplicaron algún tipo de intervención psicológica, siendo la mayoría de ellos estudios que aplicaron técnicas de EPR solas (21 estudios), RC sola (5 estudios), otras técnicas psicológicas (un estudio: Emmelkamp y van der Heyden, 1980), o bien EPR en combinación con otras técnicas (11 estudios) y un estudio que combinó RC con psicofármaco (Van Balkom *et al.*, 1998). Las intervenciones con tratamiento psicológico se caracterizan por ser, en general, homogéneas (72,5%), incluir tareas para casa (90%), no llevar programa de seguimiento (80%), no utilizar agentes externos (90%), ser intervenciones directas (85%), con entrenamiento individual (95%), soporte oral (82,5%) y directo (94%), con una duración mediana de 9 semanas (rango: 3,7-24 semanas), a razón de 1,5 horas de intervención por semana (rango: 0,4-18 horas) y con una intensidad total de unas 16 horas por participante (rango: 5,2-389 horas). Casi la mitad de los terapeutas fueron psicólogos (47,5%), siendo menor la presencia de psiquiatras (15%) y enfermeros (12,5%). Los 7 estudios que incluyeron tratamiento psicofarmacológico lo hicieron con los antidepresivos fluvoxamina (5 estudios) y clomipramina (2 estudios), con una duración media del tratamiento de unas 16 semanas (rango: 8-24 semanas) y una dosis mediana de 216 mlgr/día (rango: 183-300 mlgr/día). Muy pocos estudios aplicaron las intervenciones en régimen de ingreso en algún hospital o clínica (sólo el 16,3%).

Las muestras de participantes analizadas se caracterizan por pertenecer a población clínica adulta, siendo el DSM en sus distintas versiones el criterio diagnóstico más utilizado (62,8%)⁴. Las muestras presentaron una edad mediana de 34 años (rango: 24-41,5 años) y 12 años de padecimiento del trastorno (rango: 4-18 años), siendo el 38,4% varones (rango: 0%-100%). Muchos estudios no detallaron con claridad la presencia de comorbilidad con otros trastornos, pudiendo establecer de forma clara la presencia de comorbilidad (básicamente depresión) en sólo el 27,9% de los estudios y la ausencia de trastornos comórbidos en el 34,9% de los casos. Así mismo, muy pocos estudios clarifican el hecho de que los participantes hayan recibido tratamiento previo (el 11,6%).

El tipo de diseño más utilizado fue el experimental (74,4%), implicando asignación aleatoria de los participantes a los grupos de intervención. La mayoría de los estudios reclutaron a los participantes a petición propia (51,2%), o bien combinando esta estrategia con la de «a instancias del experimentador» (44,2%). La mortalidad experimental estuvo en torno al 16,5% (rango: 0%-41,7%) y el 22,7% (rango: 0%-55%) para el postest y el seguimiento, respectivamente. En promedio los períodos de seguimiento fueron de 6 meses (rango: 1-19 meses). Finalmente, la

4 Aunque 14 de los 43 estudios no especificaron el instrumento diagnóstico aplicado, éstos fueron incluidos en el meta-análisis porque aportaban indicios suficientes para considerar a los participantes como pertenecientes a población clínica.

puntuación global de calidad del estudio se situó en el valor 6,5 sobre una escala de 0 a 9 (rango: 3,5-8).

Con respecto a las variables extrínsecas, cabe mencionar que todos los estudios incluidos en el meta-análisis se obtuvieron de artículos publicados en revistas, no siendo posible la recuperación de trabajos no publicados, a pesar de los esfuerzos por conseguirlos. En consecuencia, será preciso realizar un análisis del sesgo de publicación. En más de la mitad de los estudios (55,8%) el primer firmante del trabajo es psiquiatra, seguido de la formación de psicólogo (37,2%). Por último, la mayoría de los estudios meta-analizados son posteriores a 1989 (67,4%).

La ausencia de estudios no publicados en el meta-análisis podía suponer una amenaza contra la validez de nuestras conclusiones, por lo que se calculó el índice de tolerancia a los resultados nulos (Orwin, 1983). Tomando los síntomas obsesivos y compulsivos como la medida de resultado principal, encontramos que para anular los resultados de eficacia obtenidos con los 23 artículos incluidos en nuestro meta-análisis deberían existir unos 120 artículos europeos sobre el tema no recuperados por nuestro equipo. Dada la exhaustiva búsqueda bibliográfica que realizamos, y sobre una base racional, no cabe pensar que pueda existir tal volumen de trabajos no recuperados por nosotros, por lo que podemos desechar el sesgo de publicación como una posible amenaza contra la validez de nuestra investigación.

La eficacia global

Los resultados de eficacia fundamentales se obtuvieron a partir de las escalas que medían la reducción de las obsesiones, compulsiones y rituales, si bien también se calcularon tamaños del efecto para la mejora de la ansiedad, la depresión, el ajuste social y se incluyó una última categoría en la que se agruparon otros constructos psicológicos medidos. Las escalas que medían síntomas obsesivos y compulsivos más frecuentemente utilizadas fueron el *Maudsley Obsessive Compulsive Inventory*, *MOCI* (Hodgson y Rachman, 1977), la *Yale-Brown Obsessive Compulsive Scale*, *Y-BOCS* (Goodman et al., 1989), el *Compulsive Activity Checklist*, *CAC* (Marks, Hallam, Connloy y Philpott, 1977), la *Anxiety Discomfort Scale* (Marks et al., 1977), la *Four Target Ritual Scale* (Marks et al., 1980), el *Behaviour Avoidance Test*, *BAT* (Marks et al., 1977) y el *Leyton Obsessional Inventory*, *LOI* (Cooper, 1970).

La Tabla 1 presenta los resultados globales obtenidos por los grupos de tratamiento y por los grupos de control para los diferentes constructos psicológicos medidos y para las medidas de autoinforme, por una parte, las medidas de los clínicos y los asesores independientes⁵, por otra, y todas ellas combinadas. La medida de resultado de mayor interés para nuestros propósitos eran los síntomas obsesivos y compulsivos y, según se observa en la tabla, las intervenciones alcanzaron, en general, un tamaño del efecto de magnitud alta ($d_+ = 1,443$) siguiendo la clasificación de Cohen

5 Aunque inicialmente registramos por separado los índices del tamaño del efecto para las medidas de los clínicos, o terapeutas, y las de los asesores independientes (cuando éstos se utilizaron) la similitud de los resultados obtenidos nos indujo a promediar sus tamaños del efecto y presentarlos conjuntamente.

(1988), si bien fue mayor para los informes de los clínicos ($d_+ = 1,939$) que para los autoinformes ($d_+ = 1,104$)⁶. No obstante, se observa una clara heterogeneidad entre los resultados en los tres casos [Global: $Q(38) = 68,492$, $p = 0,001$; Autoinformes: $Q(35) = 54,227$, $p = 0,020$; Clínicos: $Q(31) = 71,476$, $p = 0,00005$]. Por el contrario, los tres grupos de control alcanzaron tamaños del efecto medios estadísticamente no significativos, tanto en los autoinformes ($d_+ = 0,045$) como en los informes de los clínicos ($d_+ = 0,300$) como de forma global ($d_+ = 0,209$).

Los datos que acabamos de presentar asumen el modelo de efectos fijos. El tamaño del efecto medio obtenido de forma global para los síntomas obsesivo-compulsivos partiendo del modelo de efectos aleatorios arrojó un valor similar, aunque ligeramente superior ($d_+ = 1,507$), alcanzándose homogeneidad en torno a dicha media [$Q(38) = 43,326$, $p = 0,254$]. En el seguimiento (recuérdese que, en promedio, los seguimientos registrados fueron de unos 6 meses), y volviendo al modelo de efectos fijos, los resultados de eficacia se mantuvieron e incluso se incrementaron ligeramente en los informes de los clínicos ($d_+ = 2,069$) y de forma global ($d_+ = 1,508$), si bien disminuyeron levemente en los autoinformes ($d_+ = 1,058$).

Las escalas más utilizadas para medir la depresión fueron el *Beck Depression Inventory*, *BDI* (Beck et al., 1961), la *Hamilton Depression Rating Scale*, *HDRS* (Hamilton, 1969) y la *Self-Rating Depression Scale*, *SDS* (Zung, 1965). Los tamaños del efecto obtenidos con estas escalas, aunque también estadísticamente significativos, fueron de menor magnitud que los obtenidos con las medidas de las obsesiones/compulsiones, tanto de forma global ($d_+ = 0,598$) como para los autoinformes ($d_+ = 0,602$) y los informes de los clínicos ($d_+ = 0,729$). En los tres casos se observó homogeneidad en torno al tamaño del efecto medio. Los dos únicos estudios que reportaron datos de autoinforme para las medidas de depresión obtuvieron un tamaño del efecto medio no significativo ($d_+ = 0,263$). En el seguimiento, los resultados de eficacia se mantuvieron estables (Global: $d_+ = 0,658$; Autoinformes: $d_+ = 0,649$; Clínicos: $d_+ = 0,737$).

La ansiedad se midió con el *Fear Questionnaire*, *FQ* (Marks et al., 1977), el *Free-Floating Anxiety* (Gelder y Marks, 1966), la *Anxious Mood Scale* (Watson y Marks, 1971) y el *State-Trait Anxiety Inventory*, *STAI* (Spielberger, Gorsuch y Lushene, 1970). Los tamaños del efecto obtenidos fueron de magnitud similar a los alcanzados con las medidas de depresión (Global: $d_+ = 0,643$; Autoinformes: $d_+ = 0,573$; Clínicos: $d_+ = 0,814$). Así mismo, en los tres casos también se observó homogeneidad en torno al tamaño del efecto medio. En el seguimiento, los resultados de eficacia se mantuvieron estables (Global: $d_+ = 0,625$; Autoinformes: $d_+ = 0,573$; Clínicos: $d_+ = 0,844$).

El ajuste social se registró, básicamente, mediante el *Hostility and Direction of Hostility Questionnaire*, *HDHQ* (Caine, Foulds y Hope, 1967). Los grupos de tratamiento reflejaron una clara eficacia cuya magnitud se situó entre la alcanzada con

6 De hecho, al aplicar la prueba *T* de diferencias entre medias dependientes sobre los índices *d* para medidas de autoinforme y para medidas de los clínicos, recogidas de los 30 estudios que nos permitieron calcular ambos índices, obtuvimos un resultado estadísticamente significativo a favor de las medidas de los clínicos [$T(29) = 7,673$, $p < 0,0001$].

Tabla 1
Resultados de eficacia global según el constructo psicológico medido (obsesiones/compulsiones, ansiedad, depresión, ajuste social y otros) y el modo de registro (autoinformes, clínicos y combinación de ambos).

Medidas de resultado	Grupos de tratamiento										Grupos de control			
	I. C. al 95%										I. C. al 95%			
	k	d ₊	Li	Ls	Q	p	k	d ₊	Li	Ls	k	d ₊	Li	Ls
• Global (autoinformes + clínicos)	40	1,136	1,031	1,241	66,462	0,004	3	0,184	-0,285	0,653				
Obsesiones/compulsiones	39	1,443	1,323	1,563	68,492	0,001	3	0,209	-0,275	0,693				
Ansiedad	16	0,643	0,479	0,807	19,707	0,183	-	-	-	-				
Depresión	34	0,598	0,503	0,693	35,182	0,365	2	0,263	-0,233	0,759				
Ajuste social	17	0,810	0,640	0,980	35,122	0,004	2	-0,166	-1,061	0,729				
Otros	21	0,626	0,512	0,740	49,453	0,0003	2	0,096	-0,374	0,566				
• Autoinformes	40	0,895	0,799	0,991	64,235	0,007	3	0,099	-0,356	0,554				
Obsesiones/compulsiones	36	1,104	0,998	1,209	54,227	0,020	1	0,045	-0,482	0,572				
Ansiedad	8	0,573	0,431	0,714	12,242	0,093	-	-	-	-				
Depresión	34	0,602	0,506	0,698	38,180	0,246	2	0,263	-0,233	0,759				
Ajuste social	13	0,801	0,618	0,984	30,989	0,002	2	-0,166	-1,061	0,729				
Otros	17	0,591	0,474	0,708	41,674	0,0004	2	0,096	-0,375	0,567				
• Clínicos	32	1,776	1,620	1,932	54,258	0,006	3	0,300	-0,190	0,790				
Obsesiones/compulsiones	32	1,939	1,772	2,106	71,476	0,00005	3	0,300	-0,190	0,790				
Ansiedad	10	0,814	0,581	1,047	8,930	0,444	-	-	-	-				
Depresión	13	0,729	0,538	0,919	9,214	0,684	-	-	-	-				
Ajuste social	6	1,202	0,875	1,529	9,138	0,104	-	-	-	-				
Otros	6	1,258	0,930	1,586	10,304	0,067	-	-	-	-				

k: Número de estudios. d₊: Tamaño del efecto medio. Li y Ls: Límites confidenciales inferior y superior del intervalo de confianza al 95% en torno al índice d₊. Q: Prueba de homogeneidad de los tamaños del efecto en torno a su media. p: Nivel crítico de probabilidad asociado a la prueba Q, asumiendo una distribución Chi-cuadrado de Pearson con k-1 grados de libertad.

los síntomas obsesivos y compulsivos y las medidas de depresión y ansiedad, siendo ligeramente superiores con las medidas de los clínicos ($d_+ = 1,202$) que con los autoinformes ($d_+ = 0,801$). Además, se observó heterogeneidad entre los tamaños del efecto en las medidas de autoinforme y en las medidas globales. Los dos grupos de control que midieron el ajuste social, con autoinformes, alcanzaron un tamaño del efecto no significativo ($d_+ = -0,166$). En el seguimiento, los resultados de eficacia se mantuvieron estables, tanto con los autoinformes ($d_+ = 1,193$) como con las medidas de los clínicos ($d_+ = 1,273$), manteniendo la heterogeneidad.

Finalmente, las escalas que se incluyeron en la categoría «otros constructos» fueron, básicamente, la *Symptom Checklist, SCL-90* (Derogatis, 1977) y la *Quality of Life Scale, QOL* (Marks, 1986). Se obtuvieron, nuevamente, tamaños del efecto significativos con los grupos de tratamiento y claramente superiores con las medidas de los clínicos ($d_+ = 1,258$) que con los autoinformes ($d_+ = 0,591$). En el seguimiento los resultados se mantuvieron estables (Global: $d_+ = 0,739$; Autoinformes: $d_+ = 0,696$; Clínicos: $d_+ = 1,001$).

Eficacia diferencial de los tratamientos

Una vez comprobada la existencia de una clara eficacia global de las intervenciones sobre el TOC, el paso siguiente consistió en analizar la eficacia diferencial de las diferentes técnicas de intervención, así como sus combinaciones entre ellas. A partir de este punto centramos nuestra atención en la medida global de los síntomas obsesivo-compulsivos, ya que es éste el constructo psicológico de mayor interés en este trastorno. La Tabla 2 presenta los tamaños del efecto medios obtenidos para cada técnica y combinación de técnicas, tanto psicológicas como farmacológicas⁷, para las medidas de autoinforme, de los clínicos y tomando ambas globalmente.

Centrándonos en las medidas globales, el análisis de varianza aplicado sobre las siete categorías representadas en el meta-análisis con más de un estudio reflejó la existencia de diferencias significativas entre los tamaños del efecto medios [$Q_B(6) = 31,649$, $p = 0,00002$] y una adecuada especificación del modelo, al no ser significativa la prueba de homogeneidad global intra-clase [$Q_W(30) = 38,742$, $p = 0,132$], así como ninguna de las pruebas de homogeneidad intra-clase individuales (pruebas Q_{Wij} en la Tabla 2). Con respecto a los autoinformes, las seis categorías de tratamiento con más de un estudio también mostraron diferencias significativas [$Q_B(5) = 12,384$, $p = 0,030$] y una buena especificación del modelo [$Q_W(25) = 27,682$, $p = 0,323$]. Así mismo, las cuatro categorías de tratamiento con más de un estudio en las medidas de los clínicos alcanzaron diferencias significativas [$Q_B(3) = 36,849$, $p < 0,0001$], si bien el modelo no está bien especificado [$Q_W(25) = 50,381$, $p = 0,002$]. Resultados similares se obtuvieron aplicando el modelo de efectos aleatorios. Estos resultados

7 Según los análisis de varianza aplicados sobre las 12 categorías de tratamiento que figuran en la Tabla 2, no se observaron diferencias significativas en cuanto a las edades medias de las muestras de pacientes, la distribución por género, los años de padecimiento del problema, la mortalidad experimental y la calidad del estudio.

Tabla 2
Tamaños del efecto medios en función de las técnicas de intervención.

Técnicas	GLOBAL						AUTOINFORMES						CLÍNICOS					
	k	d.	I.C. al 95%			Q _{wj}	k	d.	I.C. al 95%			Q _{wj}	k	d.	I.C. al 95%			Q _{wj}
			Li	Ls	Li				Ls	Li	Ls				Li	Ls		
Exposición	20	1,262	1,098	1,426	23,531	17	0,922	0,764	1,080	19,460	18	1,758	1,536	1,980	29,852*			
Reestructuración cognitiva	5	1,529	1,196	1,862	6,683	5	1,068	0,672	1,463	5,228	5	2,001	1,421	2,580	16,580**			
Farmacológica	1	1,327	0,459	2,195	-	1	1,135	0,331	1,939	-	1	1,711	0,702	2,720	-			
Otros tratam. psicológicos	1	0,623	-0,581	1,827	-	1	0,623	-0,581	1,827	-	-	-	-	-	-			
No tratamiento (control)	3	0,209	-0,275	0,693	1,830	1	0,045	-0,482	0,572	-	3	0,300	-0,775	1,374	1,686			
EPR + RC	2	1,750	1,341	2,159	3,162	2	1,444	-0,903	3,791	0,461	1	2,808	1,852	3,764	-			
EPR + Farmac	3	2,044	1,355	2,733	2,189	3	1,460	0,253	2,667	1,656	3	2,428	0,703	4,145	2,263			
RC + Farmac	1	2,953	1,505	4,401	-	1	2,353	1,149	3,557	-	1	3,354	1,738	4,970	-			
EPR + RC + Farmac	1	3,520	2,333	4,706	-	1	1,182	0,621	1,742	-	1	4,300	2,880	5,719	-			
EPR + RC + otros	2	1,584	1,045	2,123	0,038	2	1,421	-1,878	4,720	0,734	1	2,195	1,106	3,284	-			
EPR + Farmac + Otros	1	1,963	1,389	2,537	-	1	1,963	1,389	2,537	-	-	-	-	-	-			
EPR + RC + Farmac + Otros	2	1,318	0,885	1,751	1,309	2	1,231	-1,461	3,922	0,141	1	2,565	1,388	3,742	-			

* $p < 0,05$. ** $p < 0,01$. EPR: Exposición. RC: Reestructuración Cognitiva. Farmac: Tratamiento farmacológico. Otros: Otras técnicas psicológicas. d.: Tamaño del efecto medio. Li y Ls: Límites confidenciales inferior y superior en torno al índice d. Q_{wj}: Prueba de homogeneidad intra-clase, que se distribuye según Chi-cuadrado con $k - 1$ grados de libertad. p: Nivel crítico de probabilidad asociado a la prueba Q_{wj}.

nos permitieron examinar las posibles diferencias entre las diferentes técnicas mediante comparaciones a posteriori, controlando la inflación de la tasa de Error Tipo I por el procedimiento de Bonferroni, para lo cual nos centramos en las medidas globales de resultado. En primer lugar, no encontramos diferencias significativas entre los tamaños del efecto medios de las categorías EPR, RC, tratamiento psicofarmacológico y otras técnicas psicológicas entre sí, cuando fueron aplicadas cada una de ellas por separado ($p > 0,05$), si bien el tamaño del efecto medio más elevado lo obtuvo la categoría RC ($d_+ = 1,529$) seguida del tratamiento psicofarmacológico ($d_+ = 1,327$) y de las técnicas de EPR ($d_+ = 1,262$). La comparación de cada una de ellas con la categoría grupo de control ($d_+ = 0,209$) reflejó un resultado significativo a favor de aquéllas ($p < 0,05$), a excepción de la categoría «otras técnicas». No obstante, los resultados obtenidos para las intervenciones psicofarmacológicas (un estudio) y RC (5 estudios) deben tomarse con extrema cautela por el escaso número de estudios disponibles.

Por otra parte, el escaso número de estudios en las diferentes combinaciones entre técnicas limita también la generalizabilidad de nuestros resultados. Así, tan sólo haremos alusión a la importancia que el uso de psicofármacos⁸ parece ejercer en los resultados de eficacia. En efecto, la incorporación de antidepresivos a los tratamientos psicológicos reflejó una mejora en eficacia estadísticamente significativa, tanto cuando se combinó con EPR ($d_+ = 2,044$; $z = 5,233$, $p < 0,001$) como con RC ($d_+ = 2,953$; $z = 2,490$, $p < 0,05$). Así mismo, obtuvimos un resultado estadísticamente significativo al comparar las cinco combinaciones de técnicas en las que figuraba el tratamiento con antidepresivos frente a las dos combinaciones que no lo incluían ($z = 2,503$, $p < 0,05$). Por el contrario, la comparación de las combinaciones que incluían EPR frente a las que no la incluían no alcanzó la significación estadística, como tampoco la obtuvo esta misma comparación referida a la RC.

La importante contribución que los tratamientos farmacológicos parecen ejercer sobre los resultados de eficacia nos llevó a continuación a examinar los diferentes tipos de psicofármacos administrados. Los siete estudios que administraron de forma experimental psicofármacos utilizaron antidepresivos serotoninérgicos, en concreto, fluvoxamina (5 estudios) y clomipramina (2 estudios), no encontrándose diferencias significativas entre sus tamaños del efecto medios [$Q_B(1) = 1,567$, $p = 0,210$], aunque los resultados con clomipramina ($d_+ = 2,855$) fueron superiores a los obtenidos con fluvoxamina ($d_+ = 1,989$). No obstante, los buenos resultados obtenidos con los tratamientos psicofarmacológicos, tanto solos como en combinación con los psicológicos, deben ponderarse teniendo en cuenta las mayores tasas de abandonos que suelen producir. Así, encontramos menores tasas de abandonos con los tratamientos psicológicos (EPR: 14,4%; RC: 16,7%; combinación de ambas: 8,3%) que con los psicofármacos (psicofármacos solos: 35%; EPR + psicofármacos: 23%; RC + psicofármacos: 41,7%; EPR + RC + psicofármacos: 20%),

8 El lector debe tener presente que cuando hablamos de psicofármacos en el contexto de nuestro meta-análisis, nos estamos refiriendo exclusivamente a los antidepresivos clomipramina y fluvoxamina, que son los únicos de los que tenemos datos.

aunque el reducido número de estudios por combinación aconseja precaución en la interpretación de estas tasas.

Por último, comparamos los resultados de eficacia alcanzados con las distintas modalidades de las técnicas de EPR, para lo cual nos centramos en los 20 estudios que habían aplicado exclusivamente alguna variante de estas técnicas. No encontramos diferencias significativas al comparar la exposición guiada por el terapeuta ($d_+ = 1,357$) con la auto-exposición ($d_+ = 1,115$) y la auto-exposición asistida ($d_+ = 0,906$) [$Q_B(3) = 4,807$, $p = 0,186$], pero sí obtuvimos un efecto marginalmente significativo al comparar la exposición *in vivo* con la exposición en imaginación [$Q_B(2) = 5,529$, $p = 0,063$], siendo más altos los efectos al combinar ambos tipos de exposición ($d_+ = 1,493$) que con la exposición *in vivo* ($d_+ = 1,207$) o en imaginación ($d_+ = 0,763$) por separado.

Otras variables moderadoras

Además de las diferentes tipologías de los tratamientos analizadas en el epígrafe anterior, se aplicaron análisis de varianza y de regresión (por mínimos cuadrados ponderados) para determinar el papel que otras variables moderadoras, de las múltiples que codificamos, pudieran jugar en la heterogeneidad entre los tamaños del efecto. Para identificar aquellas variables más influyentes, tomamos como criterio el que tuvieran que alcanzar resultados significativos tanto desde el modelo de efectos fijos como de efectos aleatorios. Siguiendo este criterio estricto de significación estadística, encontramos un reducido grupo de variables influyentes en la eficacia de los tratamientos psicológicos. Las Tablas 3 y 4 presentan los resultados.

En primer lugar, la intensidad de las intervenciones psicológicas mostró una relación curvilínea, en forma de U invertida, con el tamaño del efecto [$Q_R(2) = 16,626$, $p = 0,0002$; $R_{aj}^2 = 0,203$] (Tabla 4), según la cual los mejores resultados se obtienen cuando los participantes reciben unas 10 horas de tratamiento a la semana ($d' = 2,427$). Así mismo, la magnitud (nº total de horas de tratamiento recibidas por los participantes) de la intervención mostró una relación positiva significativa [$Q_R(1) = 15,880$, $p = 0,0001$; $R_{aj}^2 = 0,215$] (Tabla 4). Otras variables relacionadas con las características de los tratamientos psicológicos que no resultaron significativas fueron (Tabla 3): la homogeneidad de la intervención (aunque los tratamientos homogéneos mostraron un tamaño del efecto medio superior a los no homogéneos), la inclusión de programa de seguimiento (aunque se obtuvieron mayores tamaños del efecto cuando se incluyó), el modo de la intervención (aunque las intervenciones directas mostraron un tamaño del efecto superior a las indirectas) y el control de psicofármacos durante la intervención psicológica (aunque se obtuvieron tamaños del efecto mayores cuando no se controló).

En segundo lugar, tres variables relacionadas con las características de los terapeutas presentaron resultados significativos: se alcanzaron tamaños del efecto más altos cuando los investigadores coincidieron con los terapeutas ($d_+ = 1,808$) frente a cuando no coincidieron ($d_+ = 1,339$); la titulación mayoritaria de los terapeutas también afectó a los tamaños del efecto, siendo mejores los resultados para la titulación

Tabla 3

Análisis de varianza ponderados sobre las variables moderadoras cualitativas.

Variable moderadora	<i>k</i>	<i>d₊</i>	<i>Q_B</i>	<i>GL</i>	<i>p</i>
• Variables de tratamiento:					
* Homogeneidad del tratamiento			0,817 <i>0,198</i>	1 <i>1</i>	0,366 <i>0,656</i>
Sí	29	1,490			
No	10	1,378			
* Programa de seguimiento			3,988 <i>1,981</i>	1 <i>1</i>	0,046 <i>0,159</i>
Sí	7	1,763			
No	31	1,408			
* Modo de la intervención			4,044 <i>1,558</i>	1 <i>1</i>	0,044 <i>0,212</i>
Indirecto	4	1,018			
Directo	34	1,497			
* Control de psicofármacos			1,213 <i>0,497</i>	1 <i>1</i>	0,271 <i>0,481</i>
Sí se controló	17	1,342			
No se controló	18	1,483			
• Variables del terapeuta:					
* Coincidencia investigador/es-terapeuta/s			8,966 <i>4,997</i>	1 <i>1</i>	0,003 <i>0,025</i>
Sí	12	1,808			
No	19	1,339			
* Formación de los terapeutas			16,935 <i>12,571</i>	3 <i>3</i>	0,007 <i>0,006</i>
Psicólogo/a	19	1,349			
Psiquiatra	6	2,450			
Enfermero/a	5	1,529			
Mixto	2	1,498			
* Experiencia de los terapeutas			10,295 <i>7,227</i>	2 <i>2</i>	0,006 <i>0,027</i>
Alta	10	1,790			
Media	8	1,365			
Baja	12	1,263			
• Variables de los participantes:					
* Psicopatologías asociadas			0,141 <i>0,359</i>	1 <i>1</i>	0,707 <i>0,549</i>
Sí	12	1,589			
No	12	1,528			
• Variables metodológicas:					
* Tipo de diseño			7,016 <i>3,464</i>	2 <i>2</i>	0,030 <i>0,177</i>
Pre-experimental	8	1,230			
Cuasi-experimental	2	1,572			
Experimental	29	1,566			
* Tipo de asignación			4,648 <i>2,561</i>	1 <i>1</i>	0,031 <i>0,109</i>
No aleatoria	10	1,301			
Aleatoria	29	1,566			

En cursiva figuran los resultados del modelo de efectos aleatorios.

Tabla 4
Análisis de regresión de las variables moderadoras cuantitativas.

Variable moderadora	<i>k</i>	<i>B</i>	<i>Q_R</i>	<i>GL</i>	<i>p</i>	<i>R_{aj}²</i>
• Variables de tratamiento:						
Duración (nº semanas)	39	0,016	2,057	1	0,151	0,004
	39	<i>0,010</i>	<i>0,418</i>	1	<i>0,518</i>	<i>0,000</i>
Intensidad (nº horas/semana)	37	*	16,626	2	0,0002	0,203
Magnitud (nº total horas)	34	0,045	15,880	1	0,0001	0,215
	34	<i>0,048</i>	<i>9,939</i>	1	<i>0,002</i>	<i>0,190</i>
• Variables de los participantes:						
Edad (años)	38	-0,009	0,188	1	0,664	0,000
	38	<i>-0,003</i>	<i>0,013</i>	1	<i>0,908</i>	<i>0,000</i>
	33	0,002	0,102	1	0,749	0,000
Género (% hombres)	33	<i>0,002</i>	<i>0,551</i>	1	<i>0,458</i>	<i>0,000</i>
	36	-0,008	0,126	1	0,722	0,000
Historia del problema (años)	36	<i>0,002</i>	<i>0,006</i>	1	<i>0,935</i>	<i>0,000</i>
• Variables metodológicas:						
Calidad del estudio	39	0,091	5,009	1	0,025	0,048
	39	<i>0,108</i>	<i>3,246</i>	1	<i>0,071</i>	<i>0,048</i>
	39	0,003	0,210	1	0,647	0,000
Mortalidad experimental	39	<i>0,004</i>	<i>0,199</i>	1	<i>0,655</i>	<i>0,000</i>

* Dada la relación curvilínea entre esta variable y el tamaño del efecto, se ajustó un modelo de regresión polinómica de segundo orden (modelo de efectos fijos), siendo la ecuación predictiva: $d' = 1,037 + 0,266 \times \text{Intens} - 0,0127 \times \text{Intens}^2$.

En cursiva figuran los resultados del modelo de efectos aleatorios. *k*: nº de estudios. *B*: coeficiente de regresión. *Q_R*: prueba de contraste del modelo, que se distribuye según chi-cuadrado de Pearson con grados de libertad $GL = \text{nº de predictores}$. *GL*: grados de libertad asociados a *Q_R*. *p*: nivel crítico de probabilidad asociado a la prueba *Q_R*. *R_{aj}²*: coeficiente de terminación ajustado.

de psiquiatría ($d_+ = 2,450$), seguida de la de enfermería ($d_+ = 1,529$), mixta ($d_+ = 1,498$) y psicología ($d_+ = 1,349$); y se observó una relación positiva y significativa entre los resultados de eficacia y la experiencia de los terapeutas, distinguiendo entre baja ($d_+ = 1,263$), media ($d_+ = 1,365$) y alta ($d_+ = 1,790$).

En lo que respecta a las variables relacionadas con las características de las muestras de participantes, ninguna de las analizadas mostró resultados significativos: la edad, el género, la historia del problema y las psicopatologías asociadas (ver Tablas 3 y 4). Por otra parte, encontramos una variable metodológica relacionada con la magnitud de los efectos (Tabla 4): la calidad del estudio, siendo positiva tal relación, es decir, a mayor puntuación en nuestra escala de calidad se observó un mayor tamaño del efecto [$Q_R(1) = 5,009$, $p = 0,025$; $R_{aj}^2 = 0,048$]. Hay que destacar, así mismo, la presencia de ciertas diferencias entre los estudios que asignaron aleatoriamente los participantes a las condiciones experimentales ($d_+ = 1,566$) frente a los que no la aplicaron ($d_+ = 1,301$).

Un modelo predictivo

Finalmente, proponemos un modelo predictivo que, con un reducido número de variables, sea capaz de explicar buena parte de la heterogeneidad de los tamaños del efecto encontrada en los estudios. Para la selección de los predictores nos basamos en criterios conceptuales y estadísticos, sobre los cuales aplicamos un análisis de regresión múltiple jerárquico en el que, en el primer paso, se introdujo la variable calidad del estudio para parcializar su influjo sobre el resto de predictores; en el segundo paso se introdujeron las dos variables relacionadas con la intensidad de la intervención (Intens e Intens²), y en el tercer paso del análisis introdujimos los cuatro predictores dicotómicos (del tipo 1-0) que identificaban en cada estudio si éste incluía técnicas de exposición (EPR), de reestructuración cognitiva (RC), algún otro tratamiento psicológico (Otros) o psicofármacos (Farmac). La Tabla 5 presenta los resultados.

Tabla 5
Resultados del análisis de regresión jerárquico.

Variables introducidas	Q_{Racum}	G L	p	Q_{Rparcial}	GL	p	R²_{acum}	R²_{parcial}
Calidad	4,371	1	0,036	4,371	1	0,036	0,038	0,038
Intens + Intens ²	20,659	3	0,0001	16,288	2	0,0003	0,244	0,206
EPR+RC+Farmac+Otros	41,844	7	< 0,0001	21,185	4	0,0003	0,531	0,287

Q_{Racum}: prueba de contraste del modelo completo. Q_{Rparcial}: prueba de significación de las nuevas variables introducidas. p: niveles críticos de probabilidad asociados a las pruebas de contraste. R²_{acum}: coeficiente de determinación total del modelo. R²_{parcial}: coeficiente de determinación parcializado.

Los tres bloques de variables provocaron un incremento significativo de la varianza explicada y, lo que es más importante, una vez controlado el influjo de las variables calidad e intensidad de la intervención, el bloque de variables formado por la identificación de los diferentes tratamientos (EPR, RC, otros tratamientos psicológicos y psicofármacos) supuso un incremento del 28,7% de la varianza explicada, lo que viene a demostrar la existencia de una clara y significativa eficacia diferencial entre las diferentes técnicas de intervención. El modelo completo explicó un 53,1% de la varianza total de los tamaños del efecto. El modelo predictivo nos permite, además, hacer predicciones tentativas sobre qué combinaciones de tratamientos son las más eficaces, siendo las mejores las que combinan RC con antidepresivo ($d' = 3,567$), seguida de los antidepresivos solos ($d' = 3,282$) y de la combinación de éstos con EPR ($d' = 2,663$). Por debajo de éstos se encontrarían las técnicas de RC solas ($d' = 2,640$) y EPR sola ($d' = 2,378$)⁹.

⁹ El modelo predictivo obtenido fue el siguiente: $d' = 0,368 + 0,093 \times \text{Calidad} + 0,205 \times \text{Intens} - 0,009 \times \text{Intens}^2 + 0,023 \times \text{ERP} + 0,285 \times \text{RC} + 0,927 \times \text{Farmac} - 0,205 \times \text{Otros}$.

Discusión

El meta-análisis llevado a cabo nos permite alcanzar varias conclusiones. En primer lugar, hemos observado claros resultados de eficacia, en general, del tratamiento psicológico y psicofarmacológico del TOC sobre la reducción de los síntomas obsesivo-compulsivos, así como el mantenimiento de tal eficacia a lo largo del tiempo, al menos a corto plazo (6 meses de seguimiento, en promedio). No obstante, se hace preciso poner a prueba la consistencia temporal con períodos más largos de seguimiento, ya que 9 de los 22 estudios de nuestro meta-análisis utilizaron seguimientos de tan sólo un mes y sólo 7 estudios presentaron datos de seguimiento entre 9 y 19 meses. Dado el frecuente carácter crónico de este trastorno, una cuestión crucial por determinar es la bondad de los tratamientos a lo largo del tiempo.

En segundo lugar, el tratamiento del TOC produce también resultados beneficiosos sobre la depresión, la ansiedad y el ajuste social, aunque no tan fuertes como sobre la reducción de los síntomas obsesivos y compulsivos. Esto demuestra los efectos específicos de la terapia, que da lugar a mejores resultados sobre los constructos que se desean mejorar. En tercer lugar, es importante considerar el tipo de registro de las conductas, ya que los informes de los clínicos presentan magnitudes del efecto claramente superiores a las de los autoinformes (Abramowitz, 1996, 1997; Christensen *et al.*, 1987; Van Balkom *et al.*, 1994). Por tanto, los estudios de eficacia siempre deben aportar ambos tipos de registro.

Por otra parte, la heterogeneidad exhibida por los resultados de los estudios se explica en gran parte por la eficacia diferencial de las técnicas de intervención. Tanto las técnicas de EPR como la RC y los antidepresivos por separado se muestran eficaces. Así pues, nuestros resultados nos llevan a concluir, en la misma línea que Abramowitz (1997) y en contra de Van Balkom *et al.* (1994) que, según la evidencia disponible, las técnicas de EPR y las de RC presentan una eficacia similar y, en consecuencia, la idea de que la EPR es superior en el tratamiento del TOC no se sostiene; antes al contrario, el tamaño del efecto medio alcanzado por las técnicas de RC se sitúa por encima de las de EPR. No obstante, estamos con Caballo y Mateos (2000) en que se necesita más investigación con las técnicas cognitivas que pongan a prueba esta afirmación, ya que en nuestro meta-análisis sólo hemos encontrado cinco estudios que probaran su efecto de forma individual (Cottraux *et al.*, 2001; Emmelkamp y Beens, 1991; Emmelkamp, Visser y Hoekstra, 1988; Van Balkom *et al.*, 1998; Van Oppen *et al.*, 1995). En cualquier caso, las técnicas de RC parecen ser una alternativa prometedora para aquellos participantes que no responden al tratamiento con técnicas de EPR y para aquéllos que padecen sólo de obsesiones sin compulsiones, donde el componente cognitivo probablemente ejercerá un importante papel (Van Oppen y Arntz, 1994).

La comparación de las diferentes modalidades de EPR nos conducen a conclusiones similares a las alcanzadas por Abramowitz (1996) en su meta-análisis sobre la eficacia de esta técnica específica sobre el TOC: la combinación de la exposición *in vivo* con la exposición en imaginación ofrece mejores resultados que ambas modalidades por separado, probablemente porque permite combatir tanto las obsesiones

como las compulsiones. Por otra parte, aunque no encontramos diferencias significativas, nuestros resultados van en la línea de Abramowitz (1996) al afirmar que la exposición guiada por el terapeuta se muestra por encima de la auto-exposición.

Coincidiendo con las conclusiones alcanzadas en otras revisiones sobre el tratamiento del TOC, nuestros resultados nos permiten afirmar que son más eficaces las intervenciones psicológicas más intensas y de mayor magnitud (Cruzado, 1997; Vallejo, 2001), estimando en unas 10 horas a la semana la intensidad más eficaz, al menos en una primera fase del tratamiento.

La combinación de cualquiera de estas intervenciones psicológicas con la administración de antidepresivos ofrece los mejores resultados y debe ser considerada como el tratamiento de elección. Esta conclusión no coincide con la propuesta por Van Balkom *et al.* (1994), ya que en su meta-análisis no hallaron evidencia de mejores resultados al incorporar psicofármacos a las intervenciones conductuales, si bien los tamaños del efecto medios apuntaban en esa dirección. Una explicación de esta discrepancia puede estar en las técnicas estadísticas aplicadas en el análisis, que son menos potentes que las utilizadas en nuestro meta-análisis. Sin embargo, resulta un tanto paradójico que sean escasos los estudios que han combinado técnicas psicológicas con psicofármacos, a pesar de ser estas combinaciones presumiblemente más eficaces; de hecho, hemos podido incluir en el meta-análisis sólo tres estudios que combinaban técnicas de EPR con antidepresivos (Cottraux *et al.*, 1990; Marks *et al.*, 1980; Van Balkom *et al.*, 1998) y un solo estudio que combinaba RC con antidepresivos (Van Balkom *et al.*, 1998). Una propuesta integradora, en la línea de la formulada recientemente por Foster y Eisler (2001), consistiría en la administración en una primera fase del tratamiento de un psicofármaco para ayudar a mejorar de forma relativamente rápida a los participantes, mientras que el uso concomitante y más prolongado de tratamiento de EPR y/o RC permitiría consolidar la reducción de los síntomas. De esta forma, se paliarían los efectos secundarios de los fármacos y, probablemente, disminuiría la tasa de abandonos que, como indican los datos empíricos, suele ser mayor con los tratamientos farmacológicos.

En cualquier caso, parece muy recomendable la administración de algún antidepresivo, si bien queda por determinar si sus efectos beneficiosos se mantienen en el tiempo, ya que no tenemos datos de seguimiento al respecto. Por otra parte, queda abierta la cuestión de si la clomipramina o la fluvoxamina son o no igualmente eficaces, al no disponer de suficientes estudios. No obstante, nuestros resultados van en la línea de los meta-análisis de Abramowitz (1997), Modigh (1987), Piccinelli *et al.* (1995) y Van Balkom *et al.* (1994), según los cuales la clomipramina es el antidepresivo serotoninérgico más eficaz para el TOC. La exclusión en nuestro meta-análisis de aquellos estudios que sólo aplicaban tratamiento farmacológico, ya que no era ese nuestro objetivo prioritario, ha limitado nuestras posibilidades de análisis en ese sentido, por lo que sería recomendable complementar estos resultados con los de otro meta-análisis que incorpore los estudios netamente farmacológicos. Sólo de esa forma será posible, por ejemplo, comparar los resultados de eficacia entre diferentes tipos de antidepresivos y no sólo de los serotoninérgicos, como ha sido nuestro caso, y tomar en consideración los mayores efectos secundarios que la clomipramina parece tener respecto de otros antidepresivos más recientes como la fluvoxamina, la

fluoxetina o la sertralina (Abramowitz, 1997); o determinar, así mismo, si los antidepresivos en el tratamiento del TOC son más eficaces cuando los sujetos sufren de depresión como psicopatología asociada.

Otros aspectos relevantes del tratamiento psicológico del TOC hay que buscarlos en las características de los terapeutas en lo referente a su nivel de experiencia y formación. No obstante, la falta de información en los estudios acerca de características relevantes limitan las posibilidades de detectar otros aspectos interesantes. Nos estamos refiriendo a características de las intervenciones tales como la homogeneidad y grado de adherencia al tratamiento, y el control del consumo de psicofármacos cuando el tratamiento aplicado es sólo psicológico; así como a aspectos de las muestras de pacientes, tales como la presencia de psicopatologías asociadas al TOC (en especial la depresión), el haber recibido tratamientos previos, las expectativas de resultado y motivación de los sujetos y la presencia de pacientes caracterizados por la ausencia de compulsiones. En esta misma línea, la calidad metodológica de los estudios es un factor que está directamente relacionado con los resultados de eficacia obtenidos en los estudios y, en consecuencia, tales resultados deben interpretarse teniendo en cuenta la corrección metodológica de los mismos.

Finalmente, un problema de nuestro meta-análisis está en que, al no disponer de grupos de control en la mayoría de los estudios, el índice del tamaño del efecto utilizado compara pretest con posttest, por lo que las ganancias observadas pueden deberse tanto a los efectos específicos como inespecíficos del tratamiento. Este mismo problema lo han padecido los meta-análisis de Christensen *et al.* (1987), Van Balkom *et al.* (1994) y Abramowitz (1996). Esto suele ocurrir en ámbitos de investigación donde ya está demostrado que ciertas intervenciones son eficaces y el objetivo de los estudios no es ya probar su eficacia, sino la eficacia diferencial entre tratamientos alternativos para encontrar el mejor o los mejores. No obstante, para controlar los efectos inespecíficos disponemos de los tamaños del efecto medios obtenidos con los tres estudios que incorporaron grupos de control, de tal forma que el cálculo de su diferencia reflejaría una estimación de los efectos propios de las intervenciones. Si efectuamos estos cálculos vemos que tanto los tratamientos conductuales y cognitivo-conductuales como los antidepresivos aquí analizados siguen presentando una clara eficacia sobre el TOC. Pero dado el escaso número de grupos de control (sólo tres estudios), y la necesidad de combinar en dicha categoría los grupos activos con los inactivos, nuestros resultados deben tomarse como meramente orientativos acerca de la eficacia diferencial de las intervenciones. La escasa inclusión de grupos de control, del tipo que sea, en la literatura europea es una deficiencia que futuras investigaciones en este campo deberían subsanar.

Referencias

(Los estudios precedidos con un asterisco fueron incluidos en el meta-análisis.)

Abramowitz, J. S. (1996). Variants of exposure and response prevention in the treatment of obsessive-compulsive disorder: A meta-analysis. *Behavior Therapy*, 27, 583-600.

- Abramowitz, J. S. (1997). Efficacy of psychological and pharmacological treatments for obsessive-compulsive disorder: A quantitative review. *Journal of Consulting & Clinical Psychology, 65*, 44-52.
- Abramowitz, J. S. (1998). Does cognitive-behavioral therapy cure obsessive-compulsive disorder? A meta-analytic evaluation of clinical significance. *Behavior Therapy, 29*, 339-355.
- American Psychiatric Association (1994). *Manual diagnóstico y estadístico de los trastornos mentales*. Barcelona: Masson (orig. en inglés, APA, 1994).
- *Bachofen, M., Nakagawa, A., Marks, I. M., Park, J. M., Greist, J. H., Baer, L., Wenzel, K. W., Parkin, J. R. y Dottl, S. L. (1999). Home self-assessment and self-treatment of obsessive-compulsive disorder using a manual and a computer-conducted telephone interview: Replication of a U.K.-U.S. study. *Journal of Clinical Psychiatry, 60*, 545-549.
- Ball, S. G., Baer, L. y Otto, M. W. (1996). Symptom subtypes of obsessive-compulsive disorder in behavioral treatment studies: A quantitative review. *Behaviour Research & Therapy, 34*, 47-51.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. E. y Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571.
- Caballo, V. E. y Mateos, P. M. (2000). El tratamiento de los trastornos de ansiedad a las puertas del siglo XXI. *Psicología Conductual, 8*, 173-215.
- Caine, T. M., Foulds, G. A. y Hope, K. (1967). *Manual of the Hostility and Direction of Hostility Questionnaire (HDHQ)*. Londres: University of London Press.
- Christensen, H., Hadzi-Pavlovic, D., Andrews, G. y Mattick, R. (1987). Behavior therapy and tricyclic medication in the treatment of obsessive-compulsive disorder: A quantitative review. *Journal of Consulting & Clinical Psychology, 55*, 701-711.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2ª ed.). Hillsdale, NJ: Erlbaum.
- Cooper, H. y Hedges, L. V. (Eds.) (1994). *The handbook of research synthesis*. Nueva York: Sage.
- Cooper, J. (1970). The Leyton Obsessional Inventory. *Psychological Medicine, 1*, 48-64.
- *Cottraux, J., Mollard, E., Bouvard, M., Marks, I. M., Sluys, M., Nury, A. M., Douge, R. y Cialdella, P. (1990). A controlled study of fluvoxamine and exposure in obsessive-compulsive disorder. *International Clinical Psychopharmacology, 5*, 17-30.
- *Cottraux, J., Note, I., Yao, S. N., Lafont, S., Note, B., Mollard, E., Bouvard, M., Sauteraud, A., Bourgeois, M. y Dartigues, J. F. (2001). A randomized controlled trial of cognitive therapy versus intensive behaviour therapy in obsessive-compulsive disorder. *Psychotherapy & Psychosomatics*, en prensa.
- Cox, B. J., Swinson, R. P., Morrison, B. y Lee, P. S. (1993). Clomipramine, fluoxetine, and behavior therapy in the treatment of obsessive-compulsive disorder: A meta-analysis. *Journal of Behavior Therapy & Experimental Psychiatry, 24*, 149-153.
- Cruzado, J. A. (1997). Técnicas de intervención con pacientes obsesivo-compulsivos. *Ansiedad y Estrés, 3*, 289-318.
- Derogatis, L. R. (1977). *SCL-90: Administration, scoring and procedures manual-I for the revised version and other instruments of the psychopathology rating scales*. Baltimore, MD: The Johns Hopkins University School of Medicine.
- Douglass, H. M., Moffitt, T. E., Dar, R., McGee, R. y Silva, P. (1995). Obsessive-compulsive disorder in a birth cohort of 18-year olds: prevalence and predictors. *Journal of the American Academy of Child & Adolescent Psychiatry, 34*, 1424-1431.
- *Drummond, L. M. (1993). The treatment of severe, chronic, resistant obsessive-compulsive disorder: An evaluation of an in-patient programme using behavioural psychotherapy in combination with other treatments. *British Journal of Psychiatry, 163*, 223-229.
- *Emmelkamp, P. M. G. y Beens, H. (1991). Cognitive therapy with obsessive-compulsive disorder: A comparative evaluation. *Behaviour Research & Therapy, 29*, 293-300.

- *Emmelkamp, P. M. G., de Haan, E. y Hoogduin, C. A. L. (1990). Marital adjustment and obsessive-compulsive disorder. *British Journal of Psychiatry*, 156, 55-60.
- *Emmelkamp, P. M. G. y de Lange, I. (1983). Spouse involvement in the treatment of obsessive-compulsive patients. *Behaviour Research & Therapy*, 21, 341-346.
- *Emmelkamp, P. M. G. y Giesselbach, P. (1981). Treatment of obsessions: Relevant v. irrelevant exposure. *Behavioural Psychotherapy*, 9, 322-329.
- *Emmelkamp, P. M. G., van Linden van den Heuvel, C., Rüphan, M. y Sanderman, R. (1989). Home-based treatment of obsessive-compulsive patients: Intersession interval and therapist involvement. *Behaviour Research & Therapy*, 27, 89-93.
- *Emmelkamp, P. M. G., van der Helm, M., van Zanten, B. L. y Plochg, I. (1980). Treatment of obsessive-compulsive patients: The contribution of self-instructional training to the effectiveness of exposure. *Behaviour Research & Therapy*, 18, 61-66.
- *Emmelkamp, P. M. G. y van der Heyden, H. (1980). Treatment of harming obsessions. *Behaviour Analysis & Modification*, 9, 341-346.
- *Emmelkamp, P. M. G., Visser, S. y Hoekstra, R. J. (1988). Cognitive therapy vs exposure in vivo in the treatment of obsessive-compulsives. *Cognitive Therapy & Research*, 12, 103-114.
- Foa, E. B., Steketee, G. S. y Milby, J. B. (1980). Differential effects of exposure and response prevention in obsessive-compulsive washers. *Journal of Consulting & Clinical Psychology*, 48, 71-79.
- Foster, P. S. y Eisler, R. M. (2001). An integrative approach to the treatment of obsessive-compulsive disorder. *Comprehensive Psychiatry*, 42, 24-31.
- Gelder, M. G. y Marks, I. M. (1966). Severe agoraphobia: A controlled prospective trial of behaviour therapy. *British Journal of Psychiatry*, 112, 309-319.
- Glass, G. V., McGaw, B. y Smith, M. L. (1981). *Meta-analysis for social research*. Beverly Hills, CA: Sage.
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R. y Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Archives of General Psychiatry*, 46, 1006-1011.
- Hamilton, M. (1969). Standardized assessment and recording of depressive symptoms. *Psychiatrica Neurologia & Neurochirurgia*, 72, 201-205.
- Hedges, L. V. y Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hedges, L. V. y Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, 3, 486-504.
- Hodgson, R. J. y Rachman, S. (1977). Obsessional-compulsive complaints. *Behaviour Research & Therapy*, 15, 389-395.
- *Hohagen, F., Winkelmann, G., Rasche-Räuchle, H., Hand, I., König, A., Münchau, N., Hiss, H., Geiger-Kabisch, C., Käßler, C., Schramm, P., Rey, E., Aldenhoff, J. y Berger, M. (1998). Combination of behaviour therapy with fluvoxamine in comparison with behaviour therapy and placebo. *British Journal of Psychiatry*, 173(Suppl. 35), 71-78.
- *Ito, L. M., de Araujo, L. A., Hemsley, D. R. y Marks, I. M. (1995). Beliefs and resistance in obsessive-compulsive disorder: Observations from a controlled study. *Journal of Anxiety Disorders*, 9, 269-281.
- Karno, M. y Golding, J. M. (1991). Obsessive-compulsive disorders. En L. N. Robins y D. A. Reiger (Eds.), *Psychiatric disorders in America: The epidemiological catchment area study* (pp. 204-219). Nueva York: The Free Press.
- *Keijsers, G. P. J., Hoogduin, C. A. L. y Schaap, C. P. D. R. (1994). Predictors of treatment outcome in the behavioral treatment of obsessive-compulsive disorder. *British Journal of Psychiatry*, 165, 781-786.

- Lipsey, M. W. (1994). Identifying potentially interesting variables and analysis opportunities. En H. Cooper y L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 111-123). Nueva York: Sage.
- *Lovell, K., Marks, I. M., Noshirvani, H. y O'Sullivan, G. (1994). Should treatment distinguish anxiogenic from anxiolytic obsessive-compulsive ruminations? *Psychotherapy & Psychosomatics*, *61*, 150-155.
- Marks, I. M. (1986). *Behavioural psychotherapy: Maudsley pocketbook of clinical management*. Bristol: John Wright.
- Marks, I. M., Hallam, R., Connloy, J. y Philpott, R. (1977). *Nursing in behavioural psychotherapy: An advanced clinical role for nurses*. Londres: Royal College of Nursing.
- *Marks, I. M., O'Dwyer, A. M., Meehan, O., Greist, J., Baer, L. y McGuire, P. (2000). Subjective imagery in obsessive-compulsive disorder before and after exposure therapy. *British Journal of Psychiatry*, *176*, 387-391.
- *Marks, I. M., Stern, R. S., Mawson, D., Cobb, J. y McDonald, R. (1980). Clomipramine and exposure for obsessive-compulsive rituals: I. *British Journal of Psychiatry*, *136*, 1-25.
- Modigh, K. (1987). Antidepressant drugs in anxiety disorders. *Acta Psychiatrica Scandinavica*, *76*, 57-71.
- Morris, S. B. (2000). Distribution of the standardized mean change effect size for meta-analysis of repeated measures. *British Journal of Mathematical & Statistical Psychology*, *53*, 17-29.
- *Nakagawa, A., Marks, I. M., Park, J. M., Bachofen, M., Baer, L., Dottl, S. L. y Greist, J. H. (2000). Self-treatment of obsessive-compulsive disorder guided by manual and computer-conducted telephone interview. *Journal of Telemedicine & Telecare*, *6*, 22-26.
- Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics*, *8*, 157-159.
- Orwin, R. G. (1994). Evaluating coding decisions. En H. Cooper y L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 139-162). Nueva York: Sage.
- Piccinelli, M., Pini, S., Bellantuono, C. y Wilkinson, G. (1995). Efficacy of drug treatment in obsessive-compulsive disorder: A meta-analytic review. *British Journal of Psychiatry*, *166*, 424-443.
- Rasmussen, S. A. y Eisen, J. L. (1990). Epidemiology of obsessive compulsive disorder. *Journal of Clinical Psychiatry*, *51*(Supl. 2), 10-13.
- Reiger, D., Narrow, W. E. y Raye, D. S. (1990). The epidemiology of anxiety disorders: The epidemiologic catchment area experience. *Journal of Psychiatric Research*, *24*(Supl. 2), 3-14.
- Rosa, A. I., Olivares, J. y Sánchez, J. (1999). La terapia de conducta en el contexto español e internacional: Situación actual y factores implicados en su eficacia. *Psicología Conductual*, *7*, 215-252.
- Rosenberg, M. S., Adams, D. C. y Gurevitch, J. (2000). *MetaWin 2.0: Statistical software for meta-analysis*. Sunderland, MA: Sinauer Assoc.
- Ruhmland, M. y Margraf, J. (2001). Effektivität psychologischer therapien von spezifischer phobie und zwangsstörung: Metaanalysen auf störungsebene [Efficacy of psychological treatments for specific phobia and obsessive compulsive disorder]. *Verhaltenstherapie*, *11*, 14-26.
- Sánchez, J. y Ato, M. (1989). Meta-análisis: Una alternativa metodológica a las revisiones tradicionales de la investigación. En J. Arnau y H. Carpintero (Coords.), *Tratado de psicología general. 1: Historia, teoría y método* (pp. 617-669). Madrid: Alhambra.
- Spielberger, C. D., Gorsuch, R. L. y Lushene, R. E. (1970). *Manual for the Stait-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- SPSS 11.0 (2001). *Statistical Package for the Social Sciences*. SPSS, Inc.
- Stanley, M. A. y Turner, S. M. (1995). Current status of pharmacological and behavioral treatment of obsessive-compulsive disorder. *Behavior Therapy*, *26*, 163-186.

- Stoll, A. L., Tohen, M. y Baldessrini, R. J. (1992). Increasing frequency of the diagnosis of obsessive-compulsive disorder. *American Journal of Psychiatry*, 149, 638-640.
- *Thornicroft, G., Colson, L. y Marks, I. M. (1991). An In-patient Behavioural Psychotherapy Unit: Description and audit. *British Journal of Psychiatry*, 158, 362-367.
- Vallejo, M. A. (2001). Tratamientos psicológicos eficaces para el trastorno obsesivo compulsivo. *Psicothema*, 13, 419-427.
- *Van Balkom, A. J. L. M., de Haan, E., van Oppen, P., Spinhoven, P., Hoogduin, K. A. L. y van Dyck, R. (1998). Cognitive and behavioral therapies alone versus in combination with fluvoxamine in the treatment of obsessive-compulsive disorder. *Journal of Nervous & Mental Disease*, 186, 492-499.
- Van Balkom, A. J. L. M., van Oppen, P., Vermeulen, A. W. A., van Dyck, R., Nauta, M. C. E. y Vorst, H. C. M. (1994). A meta-analysis of the treatment of obsessive compulsive disorder: A comparison of antidepressants, behavior, and cognitive therapy. *Clinical Psychology Review*, 14, 359-381.
- *Van den Hout, M., Emmelkamp, P. M. G., Kraaykamp, H. y Griez, E. (1988). Behavioral treatment of obsessive-compulsives: Inpatient vs outpatient. *Behaviour Research & Therapy*, 26, 331-332.
- Van Oppen, P. y Arntz, A. (1994). Cognitive therapy for obsessive-compulsive disorder. *Behaviour Research & Therapy*, 32, 79-87.
- *Van Oppen, P., de Haan, E., van Balkom, A. J. L. M., Spinhoven, P., Hoogduin, K. y van Dyck, R. (1995). Cognitive therapy and exposure in vivo in the treatment of obsessive compulsive disorder. *Behaviour Research & Therapy*, 33, 379-390.
- Watson, J. P. y Marks, I. M. (1971). Relevant and irrelevant fear in flooding: A cross-over study of phobic patients. *Behavior Therapy*, 2, 275-295.
- Weismann, M. M., Bland, R. C., Canino, G. J. y Greenwald, S. (1994). The Cross National Collaborative Group: The cross national epidemiology of obsessive-compulsive disorder. *Journal of Clinical Psychiatry*, 55(Supl. 3), 5-10.
- Yaryura-Tobias, J. A. (1997). *Trastornos obsesivo-compulsivos*. Madrid: Harcourt Brace.
- Zung, W. W. K. (1965). A self-rating depression scale. *Archives of General Psychiatry*, 12, 63-70.

