

Meta-análisis e Intervención Psicosocial Basada en la Evidencia

Meta-analysis and Evidence-Based Psychosocial Intervention

Julio Sánchez-Meca, Fulgencio Marín-Martínez y José Antonio López-López

Universidad de Murcia, Spain

Resumen. Las intervenciones psicosociales que se aplican en la práctica profesional deberían ser aquellas que han recibido las mejores evidencias científicas acerca de su eficacia. El enfoque de la *Intervención Psicosocial Basada en la Evidencia* es una herramienta metodológica que pretende concienciar a los profesionales y a los gestores políticos de la necesidad de que su práctica profesional esté guiada por las mejores evidencias. Para este propósito las revisiones sistemáticas y los meta-análisis de los estudios evaluativos empíricos juegan un papel primordial, ya que permiten sintetizar los resultados de numerosas investigaciones sobre un mismo problema para determinar cuáles son los mejores tratamientos o intervenciones para resolverlo. En este artículo se presenta una panorámica de qué son los meta-análisis y qué información pueden ofrecer a la práctica profesional. Se presentan las fases en que se lleva a cabo un meta-análisis: (a) formulación del problema, (b) búsqueda de los estudios, (c) codificación de los estudios, (d) cálculo del tamaño del efecto, (e) técnicas estadísticas de integración y (f) publicación del estudio. El alcance de los meta-análisis y de sus resultados se ilustra con un ejemplo y se discuten sus implicaciones para la práctica profesional.

Palabras clave: meta-análisis, Práctica Basada en la Evidencia, revisión sistemática, tamaño del efecto.

Abstract. Psychosocial interventions that are applied in practice should be those that have received the best scientific evidence about their effectiveness. *Evidence-Based Psychosocial Intervention* is a methodological tool that aims to raise awareness among professionals and policy makers of the need for professional practice to be guided by the best evidence. For this purpose, systematic reviews and meta-analyses of empirical evaluation studies play an important role as they allow us to synthesize the results of numerous studies on the same issue to determine which are the best treatments and interventions for solving the problem. This article presents an overview of the meta-analyses and the information they can provide for professional practice. The phases in which a meta-analysis is carried out are outlined as follows: (a) formulating the problem, (b) searching for the studies, (c) coding the studies, (d) calculating the effect size, (e) statistical techniques of integration and (f) publishing the study. The scope of meta-analyses and their results are illustrated with an example and their implications for professional practice are discussed.

Keywords: effect size, Evidence-Based Practice, meta-analysis, systematic review.

La práctica profesional debería estar basada en las mejores evidencias científicas. Esta premisa fundamental debería aplicarse a cualquier ámbito profesional y, como tal, al campo de la intervención psicosocial. Los profesionales de este campo deberían decidir qué programa, tratamiento o intervención aplicar en función de las evidencias alcanzadas en estudios evaluativos empíricos debidamente diseñados e implementados. Además, las políticas sociales, educativas y sanitarias deberían también decidirse tomando en consideración la evidencia científica acumulada sobre cómo mejor tratar o prevenir los problemas sociosanitarios. Con este propósito ha surgido el enfoque de la *Práctica Basada en la Evidencia*, como una herramienta dirigida a lograr que los programas, tratamientos e intervenciones que se apliquen en la práctica

profesional rutinaria sean aquellos que estén basados en las mejores evidencias o pruebas científicas (Frías Navarro y Pascual Llobell, 2003; Pascual Llobell, Frías Navarro y Monterde, 2004; Vázquez y Nieto, 2003). Este enfoque, que surgió en la Medicina en la década de 1990, se ha ido extrapolando a otras disciplinas, como la Psicología, la Criminología, el Trabajo Social o la Educación. Pero para que el enfoque de la *Práctica Basada en la Evidencia* resulte operativo, se hace preciso disponer de revisiones sistemáticas de los estudios empíricos que sean capaces de ofrecer una panorámica de cuáles son los programas más eficaces y efectivos para resolver o prevenir un problema psicosocial. Una *revisión sistemática* es una revisión objetiva de una pregunta formulada de forma clara para cuya respuesta es preciso integrar los estudios empíricos que se han llevado a cabo sobre ella. Si en este proceso de integración se aplican métodos estadísticos sobre los resultados de los estudios empíricos, entonces esa revisión sistemática se denomina *meta-análisis*, que es el tipo de revisión

Correspondencia: Julio Sánchez-Meca, Departamento de Psicología Básica y Metodología. Facultad de Psicología. Campus de Espinardo. Universidad de Murcia. 30100-Murcia, Spain. E-mail: jsmeca@um.es - Web: www.um.es/metaanalysis

más potente posible (Sánchez-Meca, 2010; Sánchez-Meca y Botella, 2010). En el ámbito de la intervención psicosocial, un meta-análisis permite determinar qué programas son más efectivos, bajo qué condiciones y para qué tipos de participantes, receptores o pacientes. Los meta-análisis ofrecen, pues, una información crucial que, utilizada adecuadamente por los gestores y políticos, pueden contribuir a optimizar la práctica profesional. Como resultado del auge que está experimentando el enfoque de la práctica basada en la evidencia en las Ciencias Sociales y de la Salud, son múltiples las iniciativas que se están desarrollando en el plano internacional para promover este enfoque, así como la realización de revisiones sistemáticas y meta-análisis. En esta línea, especial mención merece la Colaboración Campbell, una organización internacional dirigida a promover la realización de meta-análisis y revisiones sistemáticas de alta calidad, aunando esfuerzos de muchos equipos de investigación, gestores y profesionales en los ámbitos de la Educación, la Criminología y los Servicios Sociales (Petrosino, Boruch, Soydan, Duggan y Sánchez-Meca, 2001; Sánchez-Meca, Boruch, Petrosino y Rosa-Alcázar, 2002).¹

Los profesionales involucrados en el ámbito de la intervención psicosocial deberían disponer de los conocimientos adecuados para aplicar el enfoque de la *Intervención Psicosocial Basada en la Evidencia* y de qué es un meta-análisis, cómo se lleva a cabo y cómo hacer una lectura crítica de este tipo de investigación. El propósito de este artículo es presentar una panorámica de qué es el meta-análisis, cuáles son las fases en que se lleva a cabo y en qué aspectos debemos ser especialmente cuidadosos para hacer una lectura crítica de ellos. Para ilustrar la presentación del meta-análisis, utilizaremos el recientemente publicado por Sánchez-Meca, Rosa-Alcázar y López-Soler (2011) sobre la eficacia de los tratamientos psicológicos de niños y adolescentes que han sido víctimas de abuso sexual.

Fases de un meta-análisis

Pueden ser varias las razones de por qué hacer un meta-análisis. La más común es por la existencia de un cuerpo de estudios evaluativos sobre un mismo problema que arrojan resultados heterogéneos, e incluso contradictorios, por ejemplo, acerca de qué programas, tratamientos o intervenciones son más apropiados para resolver, paliar o mejorar un problema psicológico, psicosocial, educativo, o de salud. Aunque no existan contradicciones aparentes en los resultados de un conjunto de estudios, también es justificable hacer un meta-análisis si se pretende obtener una estimación global de la eficacia de un programa o tratamiento a partir de toda la evidencia acumulada. Sea cual sea la razón, la realización de un meta-análisis requiere una

serie de fases o etapas similares a las de cualquier investigación científica.

A continuación, se presentan dichas etapas y se ilustran con un meta-análisis realizado por Sánchez-Meca *et al.* (2011) sobre la eficacia de los tratamientos psicológicos en menores que han sufrido abuso sexual. Para una mayor profundización en las fases de un meta-análisis pueden consultarse los textos de Borenstein, Hedges, Higgins y Rothstein (2009), Cooper (2010), Cooper, Hedges y Valentine (2009), Higgins y Green (2008), Lipsey y Wilson (2001), Littell, Corcoran y Pillai (2008) o Petticrew y Roberts (2006). En castellano pueden consultarse los textos de Botella y Gambara (2002) y Martín, Tobías y Seoane (2006), o los capítulos de Sánchez-Meca y Ato (1989) y Sánchez-Meca (1999, 2008).

1. Formulación del problema

La primera fase de un meta-análisis consiste en plantear el problema objeto de estudio de forma clara y precisa y definir conceptual y operativamente los constructos psicológicos, psicosociales, educativos, etc., implicados en la pregunta. Por ejemplo, en el meta-análisis de Sánchez-Meca *et al.* (2011) sobre el abuso sexual infantil y adolescente, el objetivo era estimar el grado de eficacia de los diferentes tratamientos psicológicos que se han puesto a prueba para mejorar el bienestar y la salud mental de los menores que han sufrido abusos sexuales. Para ello, en esta primera fase se definió qué se entiende por menores, qué son los abusos sexuales a menores, qué diferentes tipos de tratamientos psicológicos se han aplicado a los menores y qué constructos psicológicos se han evaluado en los menores para determinar la eficacia de dichos tratamientos.

Si existen resultados discrepantes entre los estudios empíricos, entonces un objetivo fundamental del meta-análisis será examinar las causas de esa heterogeneidad. Para ello, se debe comprobar si diferentes características de los estudios, tales como variables relacionadas con los participantes, los tratamientos aplicados, el contexto de aplicación o la metodología de los estudios, están relacionadas con los resultados de eficacia de los programas. Estos análisis pueden ir guiados por el planteamiento de hipótesis específicas, si las investigaciones y teorías previas así lo permiten.

2. Búsqueda de los estudios

En un meta-análisis, la unidad de análisis son los estudios empíricos que han investigado el problema objeto de estudio. Una vez formulado el problema, la siguiente etapa consiste en determinar los criterios que tiene que cumplir un estudio para que pueda ser incluido en el meta-análisis. Estos criterios de selección

dependerán del propósito del meta-análisis, pero hay una serie de criterios que no pueden faltar en ningún estudio de este tipo. Así, en un meta-análisis sobre la eficacia de intervenciones en el ámbito psicosocial, como el de Sánchez-Meca *et al.* (2011), será preciso plantear como criterios de selección de los estudios aspectos tales como a qué tipos de participantes se aplicaron las intervenciones, qué intervenciones se aceptarán en el meta-análisis, qué diseños evaluativos tienen que haber utilizado los estudios, si se aceptarán sólo estudios con grupo de control, qué tamaño muestral mínimo deben tener los grupos implicados en el estudio, si los estudios deben incluir medidas pretest o sólo posttest, qué variables de resultado deben haberse medido, cuál es el rango temporal de realización del estudio, si se aceptarán estudios publicados y no publicados y en qué idiomas puede estar escrito el estudio para ser aceptado.

Una vez fijados los criterios de selección, debe realizarse una búsqueda de la literatura lo más completa posible, con objeto de localizar los estudios que cumplen con los criterios de selección. Para ello, deben combinarse estrategias de búsqueda formales e informales. Dentro de las fuentes formales, no puede faltar la consulta de bases de datos electrónicas, tales como *PsycInfo*, *MedLine* o *ERIC*, mediante la combinación de palabras clave escogidas concienzudamente. Las búsquedas electrónicas se pueden completar con la consulta directa de revistas especializadas en el campo, de revisiones teóricas y sistemáticas previas sobre el problema de investigación y de las referencias de los propios estudios que se vayan localizando. Estas estrategias de búsqueda formal se deben completar, a su vez, con el uso de fuentes informales para tratar de localizar estudios no publicados o de difícil localización. Para ello, se pueden consultar libros de actas de congresos, informes técnicos de organismos públicos o privados, así como contactar con investigadores reconocidos en el campo y con organismos y entidades para solicitarles estudios no publicados.

En el meta-análisis de Sánchez-Meca *et al.* (2011), se consultaron las bases electrónicas *PsycInfo* y *MedLine* mediante la combinación de una serie de palabras clave, así como las referencias de varios meta-análisis previos, revisiones sistemáticas y revisiones teóricas relacionadas con el abuso sexual infantil. También se enviaron cartas a investigadores reconocidos en el campo y se consultaron las referencias de los estudios que se fueron localizando. Esta búsqueda de estudios implicó revisar más de 2.700 referencias, dando como resultado la selección de 33 estudios evaluativos que cumplían con los criterios de selección. En total, estos estudios incluían 44 grupos de menores que habían recibido algún tratamiento psicológico y 7 grupos que actuaron como grupos de control, lo que supuso una muestra total de 1.141 menores. Este dato da una idea de la representatividad de los resultados de

un meta-análisis, en comparación con las limitaciones que tienen los estudios evaluativos individuales, al tener éstos que disponer necesariamente de muestras de participantes mucho menores.

3. Codificación de los estudios

Es habitual que los estudios integrados en un meta-análisis presenten resultados heterogéneos, a pesar de estar investigando todos ellos el mismo problema. Con objeto de investigar dicha heterogeneidad, se elabora un *Manual de Codificación* de las características de los estudios, o variables moderadoras, que pueden estar relacionadas con la eficacia de los tratamientos. Tomando como base dicho Manual, se confecciona un *Protocolo de Registro* de las variables moderadoras y se aplica a los estudios empíricos seleccionados. Dependiendo de la naturaleza y objetivos del meta-análisis, las características a codificar variarán, pero en un meta-análisis sobre la eficacia de intervenciones psicosociales pueden distinguirse hasta cinco tipos de variables moderadoras: variables de tratamiento, de los participantes, del contexto, metodológicas y extrínsecas (Lipsey y Wilson, 2001). Dentro de las *variables de tratamiento*, y tomando como ejemplo el meta-análisis de Sánchez-Meca *et al.* (2011) sobre el abuso sexual a menores, cabe mencionar el modelo teórico de intervención (e.g., conductual, cognitivo-conductual, psicodinámico, humanista, etc.), la duración e intensidad del programa, si el programa disponía de un manual de aplicación o si la intervención iba dirigida sólo al menor o si también incluía a los tutores o a la familia. De entre las *características de los participantes* cabe mencionar la edad media de la muestra de menores, la distribución por género, su distribución étnica, el tipo de agresión sexual sufrida o la relación de la víctima con el agresor (intra-familiar versus extra-familiar). Las *características contextuales* tienen que ver con el lugar y las condiciones en las que se aplicó el tratamiento, tales como si se aplicó en el hogar, en un centro de menores o en un centro educativo. Las *características metodológicas* se refieren al diseño y al procedimiento experimental llevado a cabo en el estudio. Así, cabe mencionar como características relevantes si el estudio incluyó un grupo de control y si, en caso afirmativo, éste fue activo o inactivo, si se asignaron aleatoriamente los participantes a las condiciones de tratamiento, si se utilizaron evaluadores enmascarados, el tamaño muestral, la mortalidad experimental, si los instrumentos de evaluación empleados en el estudio están debidamente validados. En este punto, es recomendable que el meta-analista aplique alguna escala de calidad metodológica que tenga en cuenta todos estos aspectos del diseño. El análisis de las características metodológicas de los estudios permite comprobar si los resultados de eficacia pueden estar sufrien-

do sesgos en sus estimaciones. Finalmente, las *características extrínsecas* se refieren a aspectos de los estudios que, en principio, no tienen por qué afectar a los resultados de eficacia de los estudios, ya que no tienen nada que ver con la aplicación del método científico, pero se ha demostrado que, en ocasiones, están asociadas a los resultados. Pueden clasificarse dentro de esta categoría el hecho de que el estudio esté o no publicado, si hay alguna fuente de financiación (pública o privada) que pueda suponer un conflicto de intereses, la formación de los investigadores (e.g., psicólogo, educador, pedagogo, trabajador social, etc.) o el género de los investigadores (o del investigador principal).

Dependiendo de la complejidad de las variables moderadoras, el proceso de codificación de éstas está sujeto a error y, por ello, se hace preciso realizar un análisis de la fiabilidad de este proceso mediante la codificación por dos, o más, codificadores independientes de los estudios empíricos, o de una parte representativa de éstos. La obtención de índices de acuerdo tales como *kappa* de Cohen, para las variables moderadoras cualitativas, o de la correlación intra-clase, para las continuas, nos permite valorar el grado de fiabilidad de este proceso (Orwin y Vevea, 2009).

4. Cálculo del tamaño del efecto

Además de registrar las características de los estudios, la realización de un meta-análisis requiere registrar los resultados numéricos de la eficacia de los tratamientos. Dado que los estudios empíricos suelen utilizar instrumentos de evaluación diferentes, aunque todos ellos hagan referencia a constructos similares, se hace preciso utilizar un índice estadístico que sea capaz de poner los resultados de eficacia de los estudios en las mismas unidades métricas. Ese índice estadístico es el *tamaño del efecto*, que puede definirse como un indicador cuantitativo que refleja el grado en que existe el fenómeno que se está investigando (Cohen, 1988). En el caso de los estudios sobre la eficacia de tratamientos, el tamaño del efecto es un indicador del grado en que el tratamiento ha sido efectivo.

Se han propuesto en la literatura meta-analítica diversos índices del tamaño del efecto para cuantificar la eficacia de los tratamientos (Borenstein *et al.*, 2009; Sánchez-Meca, 2008). La elección del índice más apropiado está en función del tipo de diseño de los estudios y de la naturaleza de las variables implicadas. En el ámbito de la eficacia de intervenciones psicosociales, los diseños de investigación más utilizados son aquéllos en los que los participantes son asignados a dos, o más, condiciones de tratamiento, tales como un grupo tratado versus un grupo de control, o bien dos o más grupos tratados con programas alternativos, junto con uno o más grupos de control. La asignación de los participantes a las condiciones de tratamiento puede

ser al azar (diseños experimentales) o no aleatoria (diseños cuasi-experimentales). Además, es habitual que se tomen medidas postest, justo al finalizar las intervenciones, y opcionalmente pueden incorporarse medidas de seguimiento (e.g., varios meses después de finalizar la intervención), e incluso medidas pretest, o de línea base. A estas diferentes combinaciones hay que añadir la naturaleza de las variables dependientes medidas en los participantes, de forma que podemos distinguir básicamente entre medidas continuas (e.g., puntuaciones en un test de ansiedad, de depresión, de autoestima, etc.) o bien medidas dicotómicas de éxito versus fracaso del tratamiento.

Dependiendo de las condiciones del diseño de investigación utilizado en los estudios empíricos, diferentes índices del tamaño del efecto pueden utilizarse para cuantificar el efecto del tratamiento. Si los estudios de un meta-análisis han asignado a los participantes a dos, o más, condiciones de tratamiento y la variable dependiente es de naturaleza continua, los índices del tamaño del efecto más apropiados son los que se enmarcan dentro de la *familia de índices d* (Ray y Shadish, 1996). La familia *d* se define como un conjunto de índices del tamaño del efecto consistentes en calcular la diferencia entre dos medias (usualmente, grupo tratado versus grupo de control) dividida por una estimación de la variabilidad de las puntuaciones (una desviación típica), con objeto de unificar la métrica de las diferentes variables dependientes utilizadas en los estudios.

Dentro de la familia *d* de tamaños del efecto existen varios índices diferentes que se pueden utilizar en función de las características específicas del diseño del estudio. El diseño más conocido, y utilizado, es aquél en el que los participantes son asignados (aleatoriamente o no) a un grupo de intervención versus un grupo de control y se toman medidas en el postest sobre alguna variable dependiente continua. Para este tipo de diseños el índice de la familia *d* más apropiado es la *diferencia media tipificada*, que se calcula mediante la ecuación (Hedges y Olkin, 1985):

$$d = c(m) \frac{\bar{y}_T - \bar{y}_C}{S}, \quad (1)$$

donde \bar{y}_T y \bar{y}_C son las medias de los grupos tratado y control, respectivamente; S es la desviación típica conjunta, que se obtiene promediando las varianzas de los dos grupos:

$$S = \sqrt{\frac{(n_T - 1) S_T^2 + (n_C - 1) S_C^2}{n_T + n_C - 2}}, \quad (2)$$

siendo n_T y n_C los tamaños muestrales de los grupos tratado y de control, respectivamente, y S_T^2 y S_C^2 las respectivas varianzas. El factor $c(m)$ corrige el ligero sesgo

positivo para muestras pequeñas mediante (Hedges y Olkin, 1985):

$$c(m) = 1 - \frac{3}{4(n_T + n_C) - 9} \tag{3}$$

Valores $d = 0$ significan que las medias de los dos grupos son iguales, por lo que no existe efecto del tratamiento. Por tanto, cuanto más se aleje el valor d de 0, mayor será la diferencia entre las medias de los dos grupos y, en consecuencia, mayor evidencia habrá de un efecto del tratamiento, positivo o negativo.

Un índice d puede interpretarse como una puntuación típica y, como tal, representa la distancia, en unidades típicas, que separa a las medias de los dos grupos (tratado vs. control). Puede utilizarse una clasificación orientativa formulada por Jacob Cohen (1988) para ayudar a interpretar la magnitud del efecto. Según este autor, en el campo de las Ciencias Sociales en general cabe esperar que valores absolutos del índice d en torno a .20, .50 y .80 se corresponderían con magnitudes del efecto de tamaño bajo, medio y alto, respectivamente. Esta propuesta ha sido avalada por los numerosos meta-análisis que se han realizado en Ciencias Sociales (Psicología, Educación, Criminología, Trabajo Social, etc.), ya que las magnitudes del efecto que suelen encontrarse en las diferentes disciplinas suelen situarse en este rango de valores d . Este mismo índice d definido en la ecuación (1) puede utilizarse para comparar los resultados alcanzados por los grupos tratado y de control en algún seguimiento posterior, en lugar de en el posttest.

No siempre los estudios incluyen grupos de control. En ocasiones, por razones éticas o de otra índole, no es posible asignar a un grupo de participantes a un grupo de control y mantenerlos durante un cierto tiempo sin recibir tratamiento. Éste fue el caso de la mayoría de los estudios encontrados en el meta-análisis de Sánchez-Meca *et al.* (2011) sobre la eficacia de los tratamientos de menores víctimas de abuso sexual: por razones éticas, apenas se dispuso de grupos de control en los estudios, por lo que el índice d definido en la ecuación (1) no es utilizable. Si no hay un grupo de control pero los estudios incorporan medidas pretest y posttest, entonces es posible aplicar otro índice de la familia d denominado *índice de cambio medio estandarizado*, d_{MR} , mediante el cual se cuantifica el efecto del tratamiento calculando la diferencia entre las medias del pretest, \bar{y}_{Pre} , y el posttest, \bar{y}_{Post} , dividida por la desviación típica del pretest, S_{Pre} (Morris y DeShon, 2002):

$$d_{MR} = c(n - 1) \frac{\bar{y}_{Pre} - \bar{y}_{Post}}{S_{Pre}} \tag{4}$$

donde el factor de corrección, $c(n - 1)$, se obtiene mediante:

$$c(n - 1) = 1 - \frac{3}{4n - 5} \tag{5}$$

siendo n el tamaño muestral.

La interpretación de este índice sigue las mismas pautas que la diferencia media tipificada, d . No obstante, al tratarse del cambio del pretest al posttest (no hay grupo de comparación), este índice no puede interpretarse como el resultado de un grupo (tratado) en comparación con otro grupo diferente (control), sino como el cambio que se produce del pretest al posttest en un grupo que recibe un tratamiento. En cualquier caso, y teniendo en cuenta esta diferencia, puede utilizarse el criterio normativo propuesto por Cohen (1988), para ayudar a interpretar el significado de este índice. El índice d_{MR} es metodológicamente más débil que la diferencia media tipificada, d , ya que al no incluir un grupo de control la estimación del efecto del tratamiento puede sufrir sesgos debidos a fuentes de invalidez, tales como los efectos de la maduración de los participantes, de la historia o los efectos reactivos a las pruebas de evaluación aplicadas en dos ocasiones.

El índice de la familia d más completo es el que se aplicaría si el diseño del estudio está compuesto por dos grupos (tratado vs. control) con medidas pretest y posttest: la diferencia de cambios medios tipificada, d_c , que se calcula mediante la ecuación (Morris, 2008):

$$d_c = c(m) \left[\frac{(\bar{y}_{Pre}^T - \bar{y}_{Post}^T) - (\bar{y}_{Pre}^C - \bar{y}_{Post}^C)}{S_{Pre}} \right] \tag{6}$$

donde \bar{y}_{Pre}^T y \bar{y}_{Post}^T son las medias pretest y posttest del grupo tratado; \bar{y}_{Pre}^C y \bar{y}_{Post}^C son las medias pretest y posttest del grupo de control, y S_{Pre} es la desviación típica promedio en el pretest de los grupos tratado y control, que se obtiene mediante:

$$S_{Pre} = \sqrt{\frac{(n_T - 1)(S_{Pre}^T)^2 + (n_C - 1)(S_{Pre}^C)^2}{n_T + n_C - 2}} \tag{7}$$

siendo n_T y n_C los tamaños muestrales de los grupos tratado y control, respectivamente, y S_{Pre}^T y S_{Pre}^C las desviaciones típicas en el pretest de los grupos tratado y de control, respectivamente. El factor $c(m)$ corrige el ligero sesgo positivo para muestras pequeñas mediante la ecuación (3).

La interpretación que podemos plantear de este índice del tamaño del efecto es similar a cualquiera de los índices de la familia d y, en particular, al primero de ellos planteado a propósito del diseño sólo posttest con grupo de control. En consecuencia, todo lo dicho respecto de su interpretación es también aplicable al índice d_c , con la salvedad de que lo que estamos comparando es en qué grado el cambio que se ha producido del pretest al posttest en el grupo tratado difiere del

cambio del pretest al postest producido en el grupo de control.

En el meta-análisis de Sánchez-Meca *et al.* (2011) sobre menores víctimas de abuso sexual, la escasez de grupos de control obligó a proponer como índice del tamaño del efecto el índice de cambio medio tipificado, d_{MR} , entre las medias del pretest y el postest de cada grupo de intervención. Pero, dada la diversidad de variables dependientes utilizadas en los diferentes estudios, hubo que calcular un índice d_{MR} sobre cada una de ellas, para posteriormente poder hacer meta-análisis separados. Así, se calculó un tamaño del efecto para medidas de: (a) conductas sexualizadas impropias y sentimientos o pensamientos acerca del abuso sexual, (b) ansiedad y estrés, (c) depresión, (d) autoestima y autoconcepto, (e) problemas de conducta y ajuste general y (f) otras medidas, dando lugar a un total de seis meta-análisis llevados a cabo para las respectivas variables dependientes.

Los estudios empíricos no suelen reportar los tamaños del efecto, por lo que el meta-analista tendrá que calcularlos. Las principales dificultades que surgen en el cálculo vienen del hecho de que los estudios pueden proporcionar información estadística muy variada, tales como medias, desviaciones típicas, pruebas *T* de diferencias entre medias, pruebas *F* de ANOVA, niveles *p* de probabilidad, proporciones, tablas de contingencia, etc. A partir de la información estadística aportada en los estudios, el meta-analista tiene que ser capaz de obtener una estimación del tamaño del efecto, para lo cual existen fórmulas y procedimientos mediante los cuales es posible facilitar estos cálculos (Glass, McGaw y Smith, 1981; Grissom y Kim, 2005; Ray y Shadish, 1996; Sánchez-Meca y Ato, 1989).

Dada la complejidad de los cálculos de los tamaños del efecto, éstos pueden sufrir errores por parte del meta-analista. Es por ello que se recomienda realizar un análisis de la fiabilidad de dichos cálculos siguiendo una estrategia similar a la comentada anteriormente para estudiar la fiabilidad del proceso de codificación de las variables moderadoras de los estudios.

Aunque hemos centrado nuestra atención en los índices de la familia *d*, existen otras familias de tamaños del efecto igualmente útiles. Si los estudios tienen un diseño de dos grupos (tratado vs. control) y la variable dependiente es dicotómica, la familia de índices de riesgo ofrece los mejores indicadores de la eficacia del tratamiento. Los tres índices de riesgo de uso más común en meta-análisis son la diferencia de riesgos, la razón de riesgos (o riesgo relativo) y el *odds ratio* (o razón de posibilidades). Cuando los estudios de un meta-análisis han aplicado diseños correlacionales en los que se cuantifica el grado de relación entre dos variables, por ejemplo, para estudiar factores de riesgo, o de protección, en un determinado problema social o psicológico, los índices de la familia *r*, o de coeficientes de correlación, son los más apropiados y, en particular, el coeficiente de correlación de Pearson. Un

abordaje de estas familias de índices del tamaño del efecto va más allá de los propósitos de este trabajo, por lo que remitimos al lector a otras fuentes más especializadas (e.g., Deeks y Altman, 2001; Fleiss, 1994; Rosenthal, Rosnow y Rubin, 2000).

5. Análisis estadístico e interpretación

La característica más destacada del meta-análisis es el uso de métodos estadísticos para integrar los resultados de los estudios. Una vez codificadas las variables moderadoras de los estudios y calculados sus tamaños del efecto, la siguiente fase consiste en analizar estadísticamente los datos. Para ello, se dispone de una matriz de datos en la que las variables moderadoras de los estudios actúan como potenciales variables independientes, o predictoras, de la heterogeneidad de los estudios, y los tamaños del efecto constituyen la variable dependiente cuya variabilidad pretende explicarse.

Es una práctica generalizada en meta-análisis aplicar métodos estadísticos en los que el peso de cada estudio está en función de la precisión de sus estimaciones del tamaño del efecto, lo cual depende fundamentalmente del tamaño muestral. De esta forma, los estudios con los tamaños muestrales más elevados aportan estimaciones más precisas del tamaño del efecto y, en consecuencia, contribuyen con un mayor peso específico en los análisis. Una presentación detallada de las fórmulas de los procedimientos estadísticos propios del meta-análisis queda fuera del alcance de este artículo. Remitimos al lector interesado a otras fuentes más especializadas (e.g., Borenstein *et al.*, 2009; Cooper *et al.*, 2009; Sánchez-Meca, 2008; Sánchez-Meca y Marín-Martínez, 2010a, 2010b).

El primer paso del análisis estadístico consiste en construir un gráfico denominado *forest plot*, en el que se presentan, de forma gráfica y numérica, los tamaños del efecto individuales obtenidos en cada estudio, junto con su correspondiente intervalo de confianza (generalmente, al 95% de nivel de confianza). Este gráfico, que se ha creado específicamente para los estudios meta-analíticos, permite visualizar el perfil de los tamaños del efecto obtenidos, su precisión y su variabilidad.

El segundo paso de los análisis estadísticos consiste en calcular un promedio de los tamaños del efecto obtenidos en los estudios, con objeto de obtener un resultado resumen de la eficacia global de los tratamientos. Este promedio se suele acompañar con un intervalo de confianza al 95%, cuya amplitud nos informa de la precisión de la estimación global del efecto. Además, los límites confidenciales del intervalo nos permiten comprobar si el efecto global es estadísticamente significativo (comprobando si el efecto nulo, generalmente el valor 0, se encuentra o no dentro del intervalo de confianza).

Como ejemplo ilustrativo, en la Figura 1 se presen-

ta un *forest plot* con los tamaños del efecto medios obtenidos con cada variable dependiente de resultado en el meta-análisis de Sánchez-Meca *et al.* (2011) sobre el tratamiento del abuso sexual infantil. El *forest plot* presenta el tamaño del efecto medio por separado para los grupos que recibieron tratamiento (cuadrados negros) y para los grupos de control (cuadrados blancos). Así, se puede observar cómo para el cómputo global de todas las variables dependientes, los grupos tratados alcanzaron un efecto medio $d = .64$, estadísticamente significativo (límites confidenciales: $.54$ y $.75$) y de magnitud media según el criterio de Cohen (1988), mientras que los grupos de control obtuvieron un efecto medio de tan sólo $d = .08$, estadísticamente no significativo (límites confidenciales: $-.13$ y $.29$) y de magnitud prácticamente nula. La comparación de los efectos medios alcanzados por los grupos tratados y de control para las diferentes variables de resultado conducen a la misma conclusión de un efecto beneficioso, estadística y prácticamente, de los tratamientos psicológicos en comparación con la no intervención sobre los menores víctimas de abuso sexual.

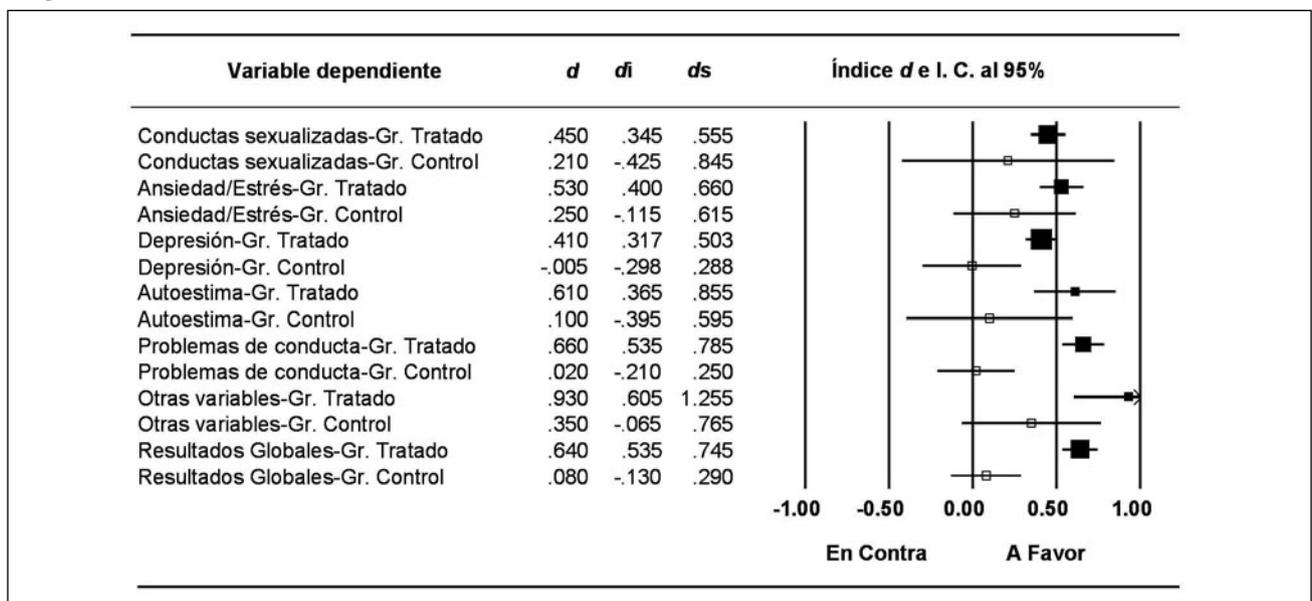
Uno de los problemas a los que se enfrenta el meta-análisis, y las revisiones de la investigación en general, es el sesgo de publicación, según el cual en algunos ámbitos es posible que sólo se publiquen estudios que han alcanzado resultados estadísticamente significativos, mientras que los estudios con resultados no significativos (generalmente, por utilizar un tamaño muestral bajo y/o por estar estimando un efecto bajo) no logran alcanzar el estatus de publicación. Si un meta-análisis sólo contiene estudios publicados, entonces la estimación del efecto medio puede estar sesgada al alza, es decir, el tamaño del efecto medio puede estar ofrecien-

do un resultado de eficacia de los tratamientos superior al que realmente se hubiera obtenido si se hubiera podido disponer de estudios no publicados. Un meta-análisis debería comprobar, mediante procedimientos estadísticos ideados para tal fin, si el tamaño del efecto medio obtenido puede estar afectado por el sesgo de publicación (Rothstein, Sutton y Borenstein, 2005).

En el meta-análisis de Sánchez-Meca *et al.* (2011) sobre el abuso sexual infantil, dado que todos los estudios estaban publicados, se aplicó el *índice de tolerancia a los resultados nulos (Fail-safe N)*, mediante el cual se calcula cuántos estudios no publicados con efecto nulo deberían existir para que el tamaño del efecto obtenido en el meta-análisis se convierta en el valor 0, es decir, en un efecto nulo. Si el número que se obtiene resulta ser muy elevado, entonces es posible descartar el sesgo de publicación como una amenaza contra la validez de los resultados del meta-análisis, ya que no será razonable asumir que puedan existir tantos estudios no publicados en ese ámbito de investigación. En concreto, para el cómputo global de todas las variables dependientes, el índice de tolerancia a los resultados nulos que se obtuvo fue de 181, un valor excesivo para pensar que el sesgo de publicación pueda ser una amenaza, por lo que pudo descartarse la existencia de dicho sesgo en los resultados del meta-análisis.

Si los estudios presentan tamaños del efecto muy heterogéneos entre sí, entonces el efecto medio no representará adecuadamente al conjunto de estudios. Para comprobar si los tamaños del efecto son excesivamente heterogéneos se aplica una prueba de homogeneidad, denominada estadístico Q que, si resulta estadísticamente significativa, revelará que los estudios son muy heterogéneos entre sí, por lo que el tamaño

Figura 1. *Forest plot* de los tamaños del efecto medios obtenidos para las diferentes variables dependientes con los grupos tratados (cuadrados negros) y con los grupos de control (cuadrados blancos). d : índice de cambio medio estandarizado promedio. d_i y d_s : límites confidenciales inferior y superior para el tamaño del efecto medio



del efecto medio no representa bien al conjunto de estudios. La prueba de homogeneidad se complementa con el cálculo del índice I^2 que, en tantos por ciento, cuantifica la cantidad de heterogeneidad existente entre los tamaños del efecto de los estudios. Valores de I^2 en torno a 25%, 50% y 75% indican heterogeneidad baja, media y alta, respectivamente (Higgins y Thompson, 2002).

En el meta-análisis de Sánchez-Meca *et al.* (2011), la prueba de homogeneidad calculada sobre el cómputo global de todas las variables dependientes arrojó un resultado estadísticamente significativo [$Q(43) = 74.571, p = .002$], con un índice $I^2 = 42.3\%$, de magnitud media, por lo que el tamaño del efecto medio no representaba bien al conjunto de los estudios.

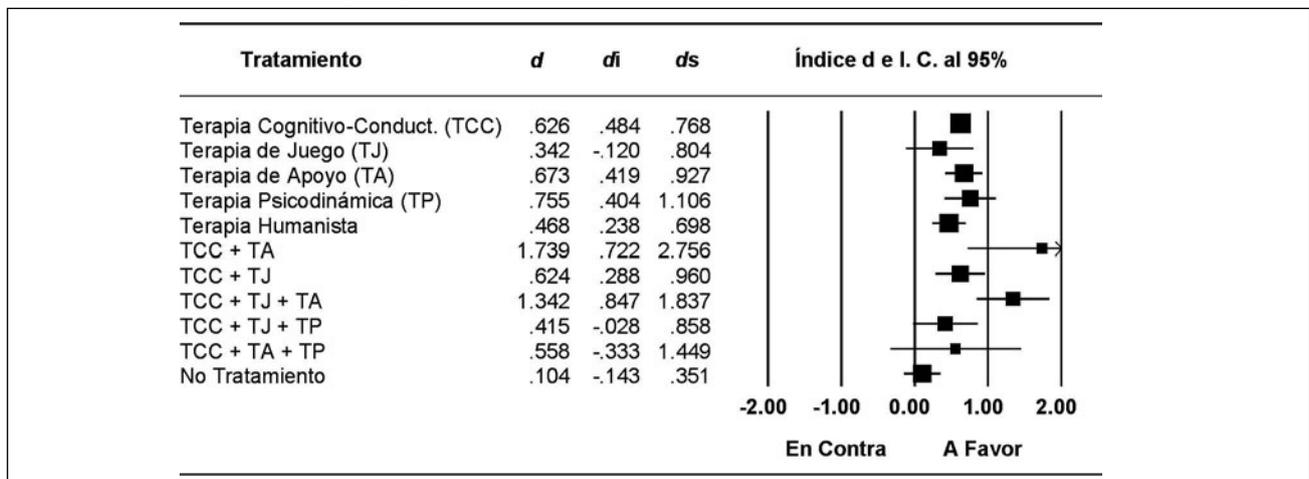
Cuando los tamaños del efecto de los estudios no son homogéneos entre sí, ello se deberá a que los estudios no son réplicas idénticas el uno del otro, sino que difieren en características importantes, tales como el tipo de tratamiento o intervención aplicado, su duración, la edad de los participantes, su distribución por género, las características metodológicas de los diseños utilizados y un sinnúmero de variables más. En consecuencia, el siguiente paso es analizar el influjo de variables moderadoras que puedan dar cuenta de la variabilidad de los tamaños del efecto. El análisis de las variables moderadoras se lleva a cabo aplicando análisis de varianza cuando la variable moderadora es cualitativa (e.g., el tipo de tratamiento) y análisis de regresión cuando es continua (e.g., la edad media de los participantes en las muestras). Lo habitual es comenzar estos análisis tomando cada variable moderadora de una en una y, si el meta-análisis contiene un número suficientemente elevado de estudios, terminar estos análisis aplicando un modelo de regresión múltiple que contenga el conjunto de variables moderadoras que expliquen la mayor parte de la variabilidad de los tamaños del efecto. En todos los casos, los análisis

estadísticos se basan en procedimientos de estimación ponderada para darle mayor peso a los estudios con los tamaños muestrales más altos y, por tanto, más precisos.

Como ejemplo ilustrativo de cómo se analiza el influjo de una variable moderadora cualitativa, la Figura 2 presenta un *forest plot* construido a partir de los resultados del meta-análisis de Sánchez-Meca *et al.* (2011), en el que se comparan los tamaños del efecto medios obtenidos por los diferentes tratamientos psicológicos, y combinaciones entre ellos, tomando como variable dependiente el tamaño del efecto obtenido en los estudios para el cómputo global de todas las variables de resultado.

Como puede observarse en dicha figura, los tratamientos psicológicos aplicados sobre menores víctimas de abuso sexual se enmarcan dentro de cinco modelos de intervención: terapia cognitivo-conductual específica de abuso (TCC), terapia de juego (TJ, que es un tipo particular del enfoque psicodinámico, pero que clasificamos aparte debido a su uso frecuente), terapia de apoyo (TA), terapias basadas en el modelo psicodinámico (TP) y terapias humanistas. A su vez, se encontraron numerosos estudios que combinaron TCC con otras terapias, tales como TJ, TA y TP, por lo que se clasificaron aparte. Los tamaños del efecto medios más elevados se alcanzaron cuando se combinó TCC con terapia de apoyo ($d = 1.74$) y si se añadió terapia de juego ($d = 1.34$). Cuando no se combinaron diferentes modelos de tratamiento, los resultados más favorables fueron para el enfoque psicodinámico ($d = .75$), la terapia de apoyo ($d = .67$) y la TCC ($d = .63$). La inspección de los intervalos de confianza permite comprobar si el efecto medio para una combinación concreta de tratamientos es o no es estadísticamente significativo (comprobando si el valor 0, o efecto nulo, se encuentra dentro de sus límites confidenciales). Por ejemplo, un modelo de tratamiento que no parece resultar efectivo

Figura 2. *Forest plot* de los tamaños del efecto medios obtenidos para los diferentes tratamientos psicológicos, y combinaciones de tratamientos, del abuso sexual infantil para la medida global de resultado de eficacia. d : índice de cambio medio estandarizado promedio para cada tratamiento. d_i y d_s : límites confidenciales inferior y superior para el tamaño del efecto medio



es el basado en terapia de juego sola, sin ningún otro componente de tratamiento adicional, ya que ésta obtuvo un efecto medio $d = .34$ con límites confidenciales $-.12$ y $.80$, no resultando estadísticamente significativo. Así mismo, la última categoría de tratamiento incluida en este análisis (y en la Figura 2) fueron los grupos de control, es decir, los que no recibieron ningún tipo de tratamiento, observándose un tamaño del efecto medio prácticamente nulo, $d = .10$, con un intervalo de confianza no significativo (límites confidenciales: $-.14$ y $.35$).

Además de la comparación visual de los efectos medios y los límites confidenciales obtenidos por las diferentes combinaciones de tratamiento, el análisis de varianza de una variable moderadora cualitativa, como es el ejemplo que nos ocupa, aporta también una prueba estadística para determinar si los efectos medios de las diferentes categorías de la variable moderadora son estadísticamente diferentes entre sí. Esta prueba, representada habitualmente como el estadístico Q_B , cumple la misma función que el estadístico F de un ANOVA típico, si bien su distribución no es F de Snedecor, sino Chi-cuadrado de Pearson con grados de libertad iguales al número de categorías de la variable moderadora menos uno. Junto con el estadístico Q_B , el ANOVA diseñado para su aplicación en meta-análisis permite también comprobar si el modelo está bien especificado mediante el estadístico Q_W , que se distribuye según Chi-cuadrado de Pearson con grados de libertad iguales al número de estudios menos el número de categorías de la variable moderadora. Si el estadístico Q_W resulta ser estadísticamente significativo, entonces el modelo no está bien especificado, lo que significa que deben existir otras variables moderadoras afectando a la variabilidad de los tamaños del efecto.

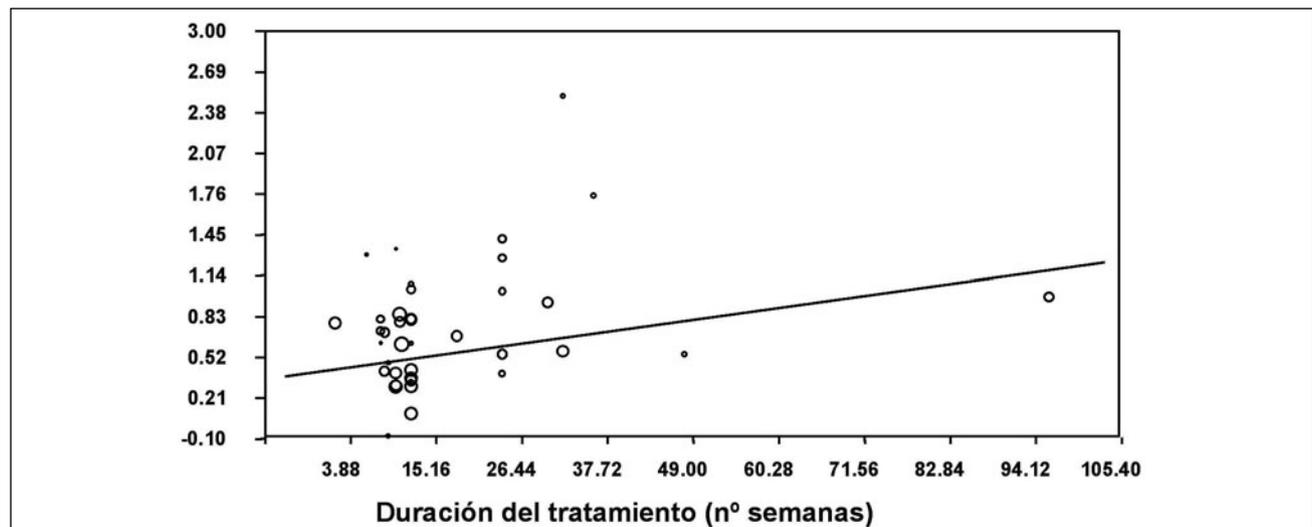
Siguiendo con el análisis de la variable moderadora ‘tipo de tratamiento’ sobre menores víctimas de abuso

sexual, el resultado de la prueba Q_B fue estadísticamente significativo [$Q_B(10) = 32.59, p < .001; \omega^2 = .33$] y con un porcentaje de varianza explicada, nada despreciable, del 33%. Además, la prueba Q_W arrojó un resultado estadísticamente no significativo [$Q_W(40) = 37.18, p > .05$], por lo que puede concluirse que el modelo está bien especificado. Estos resultados permiten concluir que las diferentes modalidades de tratamiento, y sus combinaciones, puestas a prueba en la literatura del abuso sexual infantil arrojan resultados de eficacia claramente diferentes entre sí, por lo que la práctica profesional con estos menores debería tomar en consideración estas evidencias científicas a la hora de decidir qué tratamientos son los más adecuados para mejorar la salud mental y el bienestar personal de los menores que han sido víctimas de abusos sexuales.

El análisis de variables moderadoras continuas se aborda desde modelos de regresión ponderada (simple y múltiple), que últimamente se están denominando modelos de ‘meta-regresión’. Además de obtener una estimación de la pendiente de regresión que relaciona la variable moderadora con los tamaños del efecto, el estadístico Q_R , que se distribuye según Chi-cuadrado de Pearson con grados de libertad igual al número de predictores del modelo, permite comprobar si existe una asociación estadísticamente significativa entre ambos. Así mismo, el estadístico Q_E , que se distribuye según Chi-cuadrado de Pearson con grados de libertad iguales al número de estudios menos el número de predictores, permite comprobar si el modelo está bien especificado. Los estadísticos Q_R y Q_E se corresponden con las sumas de cuadrados ponderados del modelo de regresión.

En el meta-análisis de Sánchez-Meca *et al.* (2011) una de las variables moderadoras continuas que se analizaron fue la duración (en semanas) del tratamiento. La Figura 3 presenta un diagrama de dispersión de la

Figura 3. Representación gráfica, mediante un diagrama de dispersión, de la meta-regresión simple realizada sobre los tamaños del efecto en función de la duración del tratamiento (nº de semanas)



relación entre la duración del tratamiento y el tamaño del efecto obtenido. Se observó una tendencia positiva entre ellas, que resultó estadísticamente significativa y con un 10% de varianza explicada [$Q_R(1) = 4.24$ $p = .039$; $R^2 = .10$]. Además, el resultado de la prueba de especificación del modelo indicó que éste estaba bien especificado [$Q_E(40) = 39.92$ $p = .385$]. Este resultado permite concluir que los tratamientos más prolongados logran alcanzar efectos más beneficiosos que los más breves, esto dicho dentro del rango de duraciones de los programas analizados, que se situó entre las 2 y 96 semanas, con una duración mediana de 12 semanas.

Modelo de efectos fijos versus modelo de efectos aleatorios. Los métodos estadísticos ponderados típicamente aplicados en meta-análisis se clasifican, a su vez, en función del modelo estadístico asumido, distinguiendo entre los modelos de efectos fijos y de efectos aleatorios. La elección entre uno u otro depende de varios factores. El más importante de ellos es el grado de generalización que el meta-analista pretende alcanzar en sus conclusiones. Si el meta-analista quiere generalizar sus resultados a una población mayor de estudios a los representados en el propio meta-análisis, entonces el modelo de efectos aleatorios es el adecuado (Hedges y Vevea, 1998; Raudenbush, 2009). Pero asumir este modelo implica también asumir que los estudios incluidos en el meta-análisis constituyen una muestra razonablemente representativa de esa población mayor de estudios realizados en dicho campo de investigación, o que podrían hacerse en el futuro. Además, los procesos de estimación de los parámetros en el modelo de efectos aleatorios requieren que el meta-análisis disponga de un número mínimo razonable de estudios para que dichas estimaciones sean lo suficientemente precisas como para que sean informativas. Aunque el número mínimo de estudios requerido para poder aplicar un modelo de efectos aleatorios dependerá de las condiciones concretas del meta-análisis, podemos decir que, en general, el mínimo de estudios aconsejable estará en torno a 30 (Borenstein *et al.*, 2009).

Si el meta-analista pretende generalizar los resultados a una población de estudios idénticos a los incluidos en el meta-análisis, entonces debe aplicarse el modelo de efectos fijos. Este modelo es más restrictivo en el grado de generalización de los resultados (Konstantopoulos y Hedges, 2009). Por regla general, el meta-analista deseará alcanzar unas conclusiones lo más generalizables posible, por lo que el modelo más apropiado suele ser el de efectos aleatorios (Schmidt, Oh y Hayes, 2009). Pero si el número de estudios del meta-análisis es pequeño, o no se puede definir razonablemente bien la población mayor de estudios hacia la que se desea generalizar los resultados, entonces el modelo de efectos fijos será el más indicado (Borenstein, Hedges, Higgins y Rothstein, 2010).

La elección de un modelo u otro tiene también consecuencias en los métodos estadísticos aplicados y en las

estimaciones de los parámetros. En el modelo de efectos aleatorios los procedimientos estadísticos dan lugar a intervalos de confianza más anchos y a resultados más conservadores en los contrastes de hipótesis, en comparación con el modelo de efectos fijos. Así, por ejemplo, en el meta-análisis de Sánchez-Meca *et al.* (2011) se aplicaron modelos de efectos aleatorios, ya que se pretendía generalizar los resultados a una población mayor de estudios ya realizados o que se puedan realizar en el futuro. Si se hubiera aplicado un modelo de efectos fijos, el tamaño del efecto medio que se habría obtenido para el cómputo global de todas las variables dependientes hubiera sido $d = .600$, y sus límites confidenciales .529 y .671. Si comparamos la anchura del intervalo de confianza con los dos modelos estadísticos, el de efectos aleatorios es claramente mayor ($d = .64$; límites confidenciales: .535 y .745; anchura del intervalo = .21) que el de efectos fijos (.142).

6. Publicación

La fase final de un meta-análisis, como de cualquier investigación, es publicarlo. Dado que un meta-análisis es un tipo de investigación empírica, su publicación sigue el mismo formato que el de cualquier estudio empírico: (1) introducción, (2) método, (3) resultados y (4) discusión y conclusiones (Botella y Gambara, 2006; Clarke, 2009; Sánchez-Meca y Botella, 2010). En la introducción del informe debe plantearse con claridad y contextualizarse el problema objeto de estudio. Deben definirse conceptual y operativamente los constructos implicados en el problema: las intervenciones, los participantes o receptores de los programas y el contexto de aplicación de los tratamientos. En la *Introducción* también se debe hacer una revisión teórica de la literatura, se tienen que plantear los objetivos concretos del meta-análisis y, en su caso, las hipótesis de partida.

En la sección de *Método*, se incluyen varios apartados. En primer lugar, se deben especificar los criterios de selección de los estudios y las estrategias de búsqueda de los estudios que se han utilizado, donde no puede faltar el rango temporal que ha cubierto dicha búsqueda. En otro apartado de la sección *Método* se debe detallar el proceso de codificación de las variables moderadoras y características de los estudios que posteriormente se pondrán en relación con los tamaños del efecto, así como los resultados del análisis de la fiabilidad del proceso de codificación. En este apartado debe ofrecerse el Manual de Codificación de las variables, por si algún lector estuviese interesado en solicitarlo a los autores. En el proceso de codificación de variables no puede faltar la inclusión de un conjunto de indicadores de la calidad metodológica de los estudios, con objeto de poder comprobar si los efectos de los estudios pueden estar sujetos a sesgos de estimación. El siguiente apartado debe explicar el índice del tama-

ño del efecto utilizado, sobre qué variables dependientes se ha calculado y si se han incluido cálculos del tamaño del efecto para las medidas de seguimiento. El último apartado de la sección *Método* debe explicar el modelo estadístico asumido en los cálculos meta-analíticos (de efectos fijos versus aleatorios), una justificación de su elección y que técnicas estadísticas concretas se han aplicado (e.g., intervalos de confianza, pruebas de heterogeneidad, ANOVAs, modelos de regresión, etc.).

La sección de *Resultados* debe comenzar con una descripción de las características de los estudios, para lo cual se presentan tablas de frecuencias en las que se reportan las variables moderadoras de los estudios. De esta forma, es posible comprobar qué tipos de tratamiento se han aplicado y con qué frecuencia, con qué duración e intensidad, cuál ha sido el rango de variación de las edades de los participantes, su distribución por género o por etnia, qué tipos de diseños han utilizado los estudios, los tamaños muestrales, la mortalidad experimental, etc. Esta descripción ofrece una panorámica de los estudios evaluativos integrados en el meta-análisis, así como definir el estudio prototípico.

El siguiente apartado de la sección de *Resultados* debe ofrecer una panorámica de los tamaños del efecto obtenidos en los estudios, para lo cual el *forest plot* es una técnica gráfica idónea. En dicho gráfico se suele incluir el tamaño del efecto medio obtenido por el conjunto de estudios, junto con su intervalo de confianza. Con objeto de comprobar si el sesgo de publicación puede ser una amenaza contra la validez del tamaño del efecto medio obtenido, el siguiente apartado debe presentar los resultados de algún análisis del sesgo de publicación. A continuación, se presentan los resultados de las pruebas de heterogeneidad, y si éstas apuntan a la presencia de una elevada heterogeneidad entre los tamaños del efecto entonces el siguiente apartado debe presentar los resultados de los ANOVAs y de los modelos de regresión que examinan el influjo de variables moderadoras que puedan explicar al menos parte de la variabilidad de los tamaños del efecto.

El reporte de un meta-análisis finaliza con la sección de *Discusión y Conclusiones*. Este apartado debe comenzar con un breve resumen de los principales resultados alcanzados. A continuación, éstos se ponen en relación con los de estudios y revisiones previas del tema, discutiéndose las coincidencias y contradicciones. Los meta-análisis sobre la eficacia de intervenciones en el ámbito psicosocial deben incluir una sección sobre cuáles son las principales implicaciones clínicas de los resultados de cara a la práctica profesional. El informe del meta-análisis debe finalizar con algunas indicaciones y recomendaciones para la investigación futura en el campo, tales como la detección de deficiencias metodológicas en los diseños de los estudios o vacíos en la literatura que deberían solventarse. No debemos olvidar que el meta-análisis es una mirada al

pasado para ayudar a dirigir adecuadamente la investigación futura.

En la sección de *Referencias* se deben incluir todas las referencias de los estudios incluidos en el meta-análisis (generalmente, precedidas por un asterisco para su mejor identificación). Finalmente, si las limitaciones de espacio lo permiten, es aconsejable incorporar un *apéndice* que recoja la base de datos meta-analítica completa, es decir, los tamaños del efecto individuales de los estudios junto con las variables moderadoras. De esta forma, cualquier lector interesado puede replicar los resultados del meta-análisis.

Conclusión y reflexiones finales

El propósito de este artículo era ofrecer una panorámica de qué es el meta-análisis, cómo se lleva a cabo, qué información es capaz de ofrecernos y cuál es su vinculación con el enfoque de la *Práctica Basada en la Evidencia*. No cabe duda de que los meta-análisis sobre la eficacia de intervenciones en el ámbito psicosocial están aportando una información de gran utilidad para su puesta en práctica por los profesionales que tienen que tomar decisiones día a día acerca de cómo mejor intervenir o tratar en los problemas sociales, educativos y psicológicos que forman parte de su quehacer cotidiano. Al basarse en una muestra muy amplia de participantes, las conclusiones de los meta-análisis son más fiables y generalizables que las de los estudios evaluativos individuales. Los estudios individuales pueden considerarse como las piezas del puzzle, mientras que un meta-análisis puede configurar dicho puzzle. No obstante, los meta-análisis están también sujetos a deficiencias y a sesgos en sus estimaciones, por lo que es fundamental saber hacer una lectura crítica de un meta-análisis, siendo capaz de depurar su calidad metodológica. Una cuestión crucial para la correcta interpretación de los resultados de un meta-análisis es la calidad metodológica de los estudios evaluativos incluidos en él. Si los estudios empíricos son metodológicamente débiles, las conclusiones del meta-análisis también lo serán. Por ello, los meta-análisis más válidos son aquéllos basados en estudios experimentales, en los que los participantes se han asignado aleatoriamente a las condiciones de tratamiento. En cualquier caso, los meta-análisis están ofreciendo evidencias de gran valor que pueden ser aprovechadas desde el enfoque de la *Práctica Basada en la Evidencia* para repercutir en la práctica profesional de los profesionales del ámbito de la intervención psicosocial.

Agradecimientos

Este artículo ha sido financiado por el Ministerio de Ciencia e Innovación y los Fondos FEDER (Proyecto nº: PSI2009-12172).

Referencias

- Borenstein, M. J., Hedges, L. V., Higgins, J. P. T. y Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.
- Borenstein, M. J., Hedges, L. V., Higgins, J. P. T. y Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*, 97-111.
- Botella, J. y Gambara, H. (2002). *Qué es el meta-análisis*. Madrid: Biblioteca Nueva.
- Botella, J. y Gambara, H. (2006). Doing and reporting a meta-analysis. *International Journal of Clinical and Health Psychology, 6*, 425-440.
- Clarke, M. (2009). Reporting format. En H. Cooper, L. V. Hedges y J. C. Valentine (Eds.), *The handbook of research synthesis and metaanalysis* (2ª ed.) (pp. 279-293). Nueva York: Russell Sage Foundation.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2ª ed.). Hillsdale, NJ: Erlbaum.
- Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (3ª ed.). Thousand Oaks, CA: Sage.
- Cooper, H., Hedges, L. V. y Valentine, J. C. (Eds.) (2009). *The handbook of research synthesis and meta-analysis* (2ª ed.). Nueva York: Russell Sage Foundation.
- Deeks, J. J. y Altman D. G. (2001). Effect measures for meta-analysis of trials with binary outcomes. En M. Egger, D. Smith y D. G. Altman (Eds.), *Systematic reviews in health care* (pp. 313-335). Londres: BMJ Pub. Group.
- Fleiss, J. L. (1994). Measures of effect size for categorical data. En H. Cooper y L.V. Hedges (Eds.), *The handbook of research synthesis* (pp. 245-260). Nueva York: Russell Sage Foundation.
- Frías Navarro, M. D. y Pascual Llobell, J. (2003). Psicología clínica basada en pruebas: Efecto del tratamiento. *Papeles del Psicólogo, 85*, 11-18.
- Glass, G. V., McGaw, B. y Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Grissom, R. y Kim, J. J. (2005). *Effect sizes for research. A broad practical approach*. Mahwah, NJ: LEA.
- 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.
- Higgins, J. P. T. y Green, S. (Eds.) (2008). *Cochrane handbook for systematic reviews of interventions*. Chichester, UK: Wiley-Blackwell.
- Higgins, J. P. T. y Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*, 1539-1558.
- Konstantopoulos, S. y Hedges, L. V. (2009). Analyzing effect sizes: Fixed-effects models. En H. Cooper, L. V. Hedges y J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2ª ed.) (pp. 279-293). Nueva York: Russell Sage Foundation.
- Lipsey, M. W. y Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Littell, J. H., Corcoran, J. y Pillai, V. (2008). *Systematic reviews and meta-analysis*. Oxford, UK: Oxford University Press.
- Martín, J. L. R., Tobías, A. y Seoane, T. (Coords.) (2006). *Revisión sistemática en las ciencias de la vida*. Toledo: FISCAM.
- Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*, 364-386.
- Morris, S. B. y DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*, 105-125.
- Orwin, R. G. y Vevea, J. L. (2009). Evaluating coding decisions. En H. Cooper, L. V. Hedges y J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2ª ed.) (pp. 177-203). Nueva York: Russell Sage Foundation.
- Pascual Llobell, J., Frías Navarro, M. D. y Monterde, H. (2004). Tratamientos psicológicos con apoyo empírico y práctica clínica basada en la evidencia. *Papeles del Psicólogo, 87*, 1-8.
- Petrosino, A., Boruch, R. F., Soydan, H., Duggan, L. y Sánchez-Meca, J. (2001). Meeting the challenges of evidence-based policy: The Campbell Collaboration. *Annals of the American Academy of Political and Social Science, 578*, 14-34.
- Petticrew, M. y Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Malden, MA: Blackwell.
- Raudenbush, S. W. (2009). Random effects model. En H. Cooper, L. V. Hedges y J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 295-315). Nueva York: Russell Sage Foundation.
- Ray, J. W. y Shadish, W. R. (1996). How interchangeable are different estimators of effect size? *Journal of Consulting and Clinical Psychology, 64*, 1316-1325.
- Rosenthal, R., Rosnow, R. L. y Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach*. Cambridge, UK: Cambridge University Press.
- Rothstein, H. R., Sutton, A. J. y Borenstein, M. J. (Eds.) (2005). *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. Chichester, UK: Wiley.
- Sánchez-Meca, J. (1999). Meta-análisis para la investigación científica. En F. J. Sarabia-Sánchez (Coord.), *Metodología para la investigación en marketing y dirección de empresas* (pp. 173-201). Madrid: Pirámide.
- Sánchez-Meca, J. (2008). Meta-análisis de la investigación. En M. A. Verdugo, M. Crespo, M. Badía y B. Arias (Coords.), *Metodología en la investigación sobre discapacidad: Introducción al uso de las ecuaciones estructurales*. Salamanca: Publicaciones del INICO.
- Sánchez-Meca, J. (2010). Cómo realizar una revisión sistemática y un meta-análisis. *Aula Abierta, 38*, 53-64.
- Sánchez-Meca, J. y Ato, M. (1989). Meta-análisis: Una alternativa metodológica a las revisiones tradicionales de la

- investigación. En J. Arnau y H. Carpintero (Eds.), *Tratado de Psicología General* (pp. 617-669). Madrid: Alhambra.
- Sánchez-Meca, J., Boruch, R. F., Petrosino, A. y Rosa-Alcázar, A. I. (2002). La Colaboración Campbell y la práctica basada en la evidencia. *Papeles del Psicólogo*, 83, 44-48.
- Sánchez-Meca, J. y Botella, J. (2010). Revisiones sistemáticas y meta-análisis: Herramientas para la práctica profesional. *Papeles del Psicólogo*, 31, 7-17.
- Sánchez-Meca, J. y Marín-Martínez, F. (2010a). Meta-analysis. En P. Peterson, E. Baker y B. McGaw (Eds.), *International Encyclopedia of Education*, Volumen 7 (3ª ed.) (pp. 274-282). Oxford: Elsevier.
- Sánchez-Meca, J. y Marín-Martínez (2010b). Meta-analysis in psychological research. *International Journal of Psychological Research*, 3, 151-163.
- Sánchez-Meca, J., Rosa-Alcázar, A. I. y López-Soler, C. (2011). The psychological treatment of sexual abuse in children and adolescents: A meta-analysis. *International Journal of Clinical and Health Psychology*, 11, 67-93.
- Schmidt, F. L., Oh, I.-S. y Hayes, T. L. (2009). Fixed versus random effects models in meta-analysis: Model properties and an empirical comparison of difference in results. *British Journal of Mathematical and Statistical Psychology*, 62, 97-128.
- Vázquez, C. y Nieto, M. (2003). Psicología (clínica) basada en la evidencia (PBE): Una revisión conceptual y metodológica. En J. L. Romero (Ed.), *Psicópolis: Paradigmas actuales y alternativos en la psicología contemporánea*. Barcelona: Kairos.

Notas

¹ Puede consultarse su sitio web oficial en: www.campbellcollaboration.org.

Manuscrito Recibido: 23/11/2010

Revisión Recibida: 19/01/2011

Manuscrito Aceptado: 20/01/2011