

## Recomendaciones para el Reporte de Revisiones Sistemáticas y Meta-análisis

María Rubio-Aparicio<sup>1</sup>, Julio Sánchez-Meca<sup>1\*</sup>, Fulgencio Marín-Martínez<sup>1</sup> y José Antonio López-López<sup>2</sup>

<sup>1</sup> Departamento de Psicología Básica y Metodología, Universidad de Murcia (España).

<sup>2</sup> Department of Population Health Sciences, Bristol Medical School, University of Bristol (Reino Unido).

**Título:** Recomendaciones para el Reporte de Revisiones Sistemáticas y Meta-análisis.

**Resumen:** El meta-análisis es una metodología esencial que permite a los investigadores sintetizar la evidencia científica disponible sobre una determinada cuestión de investigación. Debido a su amplia aplicabilidad en la mayoría de los campos de investigación aplicados, es realmente importante que los meta-análisis se escriban y se reporten de manera apropiada. En este artículo proponemos pautas para reportar los resultados de un meta-análisis en una revista científica como *Anales de Psicología*. Concretamente, se detalla la estructura para reportar un meta-análisis siguiendo sus fases. Además, proporcionamos recomendaciones relacionadas con las tareas usuales en meta-análisis. Un meta-análisis reciente centrado en el campo de la psicología es usado para ilustrar las guías propuestas. Finalmente, presentamos algunas observaciones finales.

**Palabras clave:** Meta-análisis; síntesis de la investigación; tamaño del efecto; calidad de la investigación

**Title:** Guidelines for Reporting Systematic Reviews and Meta-analyses.

**Abstract:** Meta-analysis is an essential methodology that allows researchers to synthesize the scientific evidence available on a given research question. Due to its wide applicability in most applied research fields, it is really important that meta-analyses be written and reported appropriately. In this paper we propose some guidelines to report the results of a meta-analysis in a scientific journal as *Annals of Psychology*. Concretely, the structure for reporting a meta-analysis following its different stages is detailed. In addition, some recommendations related to the usual tasks when conducting a meta-analysis are provided. A recent meta-analysis focused on the psychological field is used to illustrate the guidelines proposed. Finally, some concluding remarks are presented.

**Keywords:** Meta-analysis; research synthesis; effect size; research quality

### Introducción

La producción científica ha crecido exponencialmente en las últimas décadas en la mayoría de campos científicos. Como consecuencia de este notable aumento de conocimiento, las tareas de síntesis y revisión han ganado importancia para entender el estado de la cuestión sobre un determinado tema. En este contexto, el meta-análisis ha emergido como una metodología que permite a los investigadores integrar cuantitativamente los resultados de un conjunto de estudios primarios sobre un tópico común. Desde que Glass (1976) acuñó el término meta-análisis para referirse a este tipo de investigación metodológica hasta la actualidad, los meta-análisis han ganado popularidad en diferentes áreas de investigación como la educación, la psicología y las ciencias de la salud.

Actualmente, la mayor parte de las conclusiones sobre la acumulación del conocimiento en psicología están basadas en meta-análisis. La psicología basada en la evidencia es una herramienta metodológica que enfatiza la importancia de la evidencia científica en la práctica psicológica. El objetivo de esta estrategia consiste en modificar la forma de trabajo de los psicólogos, pues estos profesionales aplicados deben tener en cuenta la mejor evidencia científica para tomar sus decisiones sobre un determinado problema (Sánchez-Meca y Botella, 2010). Por ejemplo, una decisión importante para un psicólogo puede ser qué tratamiento es más efectivo para prevenir el trastorno obsesivo-compulsivo (TOC) en niños y adolescentes, así como qué moderadores pudieran afectar a dicha efectividad.

El meta-análisis es una metodología esencial para sintetizar la evidencia científica disponible sobre una cuestión de

investigación determinada en un momento determinado. Los meta-análisis deben ser llevados a cabo con el mismo rigor científico que los estudios empíricos, es decir, objetividad, sistematización y replicabilidad. La calidad del reporte de los meta-análisis no es siempre óptima. De forma similar que en la investigación primaria, la pobre calidad del reporte es un tema preocupante también en la investigación meta-analítica. Varias revisiones sistemáticas de meta-análisis publicados han encontrado una gran heterogeneidad entre los patrones de reporte (e.g., Cafri, Kromrey, y Brannick, 2010; Rubio-Aparicio, Marín-Martínez, Sánchez-Meca, y López-López, en prensa; Schmidt, Oh, y Hayes, 2009; Valentine, Cooper, Pall, Tyson, y Robinson, 2010).

Es importante hacer una distinción entre calidad metodológica y calidad del reporte de un meta-análisis. La calidad metodológica hace referencia a todos los aspectos que se deben tener en cuenta para garantizar la validez interna de un meta-análisis, con el propósito de evitar sesgos en sus resultados. La calidad del reporte consiste en incluir cada uno de los pasos y decisiones que se toman cuando se hace un meta-análisis, con el objetivo de asegurar la reproducibilidad por otros investigadores. Por tanto, un meta-análisis metodológicamente bueno no necesariamente tiene que estar bien reportado; y un meta-análisis bien reportado puede presentar deficiencias metodológicas. En este artículo, presentamos un conjunto de recomendaciones para garantizar un buen reporte.

Se han elaborado diferentes guías con el objetivo de ayudar a los investigadores a mejorar tanto el reporte como la realización de revisiones sistemáticas y meta-análisis. De todas ellas, la declaración PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*) es la más popular para revisiones sistemáticas y meta-análisis sobre la efectividad de intervenciones en el ámbito de la salud. Moher, Liberatti, Tetzlaff y Altman (2009) elaboraron esta guía como una actualización de la declaración QUOROM desarrollada por

\* Correspondence address [Dirección para correspondencia]:  
Julio Sánchez-Meca. Departamento de Psicología Básica y Metodología. Facultad de Psicología. Campus de Espinardo, Universidad de Murcia, 30100 Murcia (España). E-mail: [jsmeca@um.es](mailto:jsmeca@um.es)

Moher et al. (1999). La guía PRISMA se compone de 27 ítems sobre el reporte de un meta-análisis clasificados en título, resumen, introducción, metodología, resultados, discusión y fuentes de financiación. Además, esta guía se ha ampliado a revisiones sistemáticas y meta-análisis que incorporan métodos actuales y novedosos de estos últimos años, tales como el meta-análisis en red (Hutton et al., 2015) y el meta-análisis con datos individuales de pacientes (Stewart et al., 2015). Para más detalles, véase la web PRISMA (<http://www.prisma-statement.org/>).

Otra herramienta elaborada en el campo de la salud es la guía AMSTAR (*Assessment of Multiple Systematic Reviews*; Shea et al., 2007). La guía AMSTAR consiste en 11 ítems diseñados para estudiar la calidad metodológica, más que la calidad del reporte de un meta-análisis. Existe evidencia de que el uso de las guías PRISMA y/o AMSTAR ha incrementado la calidad del reporte de revisiones sistemáticas y meta-análisis en ciencias de la salud (e.g., Panic, Leoncini, de Belvis, Ricciardi, y Boccia, 2013).

Las guías PRISMA y AMSTAR están diseñadas sobre todo para su aplicación en meta-análisis sobre la efectividad de intervenciones. Dado que también se hacen meta-análisis de estudios observacionales, Stroup et al. (2000) propusieron la guía MOOSE (*Meta-analysis Of Observational Studies in Epidemiology*) como una herramienta para revisores, editores y lectores de estudios observacionales. La guía MOOSE se compone de 35 ítems sobre calidad del reporte distribuidos en: título, resumen, introducción, método, resultados y discusión.

En el campo de la psicología, la guía MARS (*Meta-Analysis Reporting Standard*) se desarrolló para orientar a los investigadores a hacer y reportar meta-análisis sobre la eficacia de tratamientos psicológicos (APA Publications y Communications Board Working Group on Journal Article Reporting Standards, 2008). Harris Cooper fue el encargado de dirigir el desarrollo de esta guía basándose en las guías QUOROM, PRISMA y MOOSE, y en las aportaciones recibidas en una reunión sobre meta-análisis celebrada en Potsdam (Alemania) (Cook, Sackett, y Spitzer, 1995). La guía MARS se compone de 74 ítems agrupados en título, resumen, introducción, método, resultados y discusión.

Finalmente, Sánchez-Meca et al. (2017, Julio) propusieron recientemente la guía REGEMA (*REliability GEneralization Meta-Analysis*) como un esfuerzo por ofrecer una serie de recomendaciones a los investigadores interesados en publicar un tipo especial de meta-análisis llamado “generalización de la fiabilidad” (GF). Un meta-análisis GF tiene como objetivo integrar estadísticamente los coeficientes de fiabilidad reportados en los estudios que han aplicado un determinado instrumento de medida (psicológico, psicosocial, educacional, etc.), para obtener una estimación de su fiabilidad media, valorar la heterogeneidad exhibida en las estimaciones de la fiabilidad en las diferentes aplicaciones y buscar características de los estudios que pudieran explicar la variabilidad observada en tales estimaciones de fiabilidad (Sánchez-Meca, López-López, y López-Pina, 2013). La guía RE-

GEMA se compone de 29 ítems que pueden utilizarse tanto para reportar como valorar críticamente un meta-análisis GF.

El objetivo principal de este trabajo fue ofrecer una serie de recomendaciones para el reporte adecuado de los meta-análisis a la luz de las guías ya propuestas en la literatura (PRISMA, AMSTAR, MOOSE, Botella y Gambará, 2006). Proponemos una estructura tentativa que todo meta-análisis bien reportado debería seguir acorde con la calidad científica y la transparencia requerida en investigación. Las guías PRISMA, AMSTAR, MOOSE, MARS y REGEMA están diseñadas para tipos específicos de meta-análisis. Aquí, nosotros presentamos recomendaciones de propósito general que pueden ser aplicadas independientemente del objetivo del meta-análisis. Por lo tanto, nuestras recomendaciones pretenden ser una herramienta complementaria a aquellas ya propuestas. En la próxima sección, explicamos con detalle la estructura de un meta-análisis. Además, utilizamos un meta-análisis reciente sobre la efectividad de tratamientos psicológicos con propósitos ilustrativos. Por último, ofrecemos algunas conclusiones finales.

## Escribiendo un meta-análisis

Un meta-análisis es una investigación científica y por tanto, su estructura es muy similar a la de cualquier estudio empírico, aunque con ciertas especificaciones. Básicamente, un meta-análisis se divide en cinco secciones principales: (1) Introducción, (2) método, (3) resultados, (4) discusión, y (5) referencias.

Antes de comenzar explicando cada sección es importante tener en cuenta algunas recomendaciones en la elección del título y en la composición del resumen. El título debe incluir el término “meta-análisis” para mejorar la indexación e identificación del trabajo. Por ejemplo, en el meta-análisis sobre la eficacia del tratamiento cognitivo conductual para niños y adolescentes con TOC (Iniesta-Sepúlveda, Rosa-Alcázar, Sánchez-Meca, Parada-Navas, y Rosa-Álcazar, 2017), el título era “Cognitive-behavioral high parental involvement treatments for pediatric OCD: A meta-analysis”. Por otro lado, el resumen debe contener, de forma abreviada y estructurada, los objetivos principales, los criterios de elegibilidad de los estudios, y los principales resultados y conclusiones.

### Introducción

En esta sección, debe definirse de forma clara y objetiva la pregunta de investigación. Esto implica especificar los constructos cuyas relaciones van a ser objeto de estudio, así como definir todas las variables implicadas en tales relaciones, incluyendo no sólo variables dependientes e independientes, sino también posibles variables moderadoras. Además, deben describirse los modelos teóricos y los estudios previos centrados en el objeto de investigación, enfatizando los puntos discrepantes en la literatura. Por último, debe in-

dicarse el propósito del meta-análisis, especificando los objetivos y las hipótesis, en caso de que las hubiera.

Por ejemplo, en el meta-análisis sobre TOC (Iniesta-Sepúlveda et al., 2017) se definieron constructos tales como el trastorno obsesivo compulsivo, el tratamiento familiar cognitivo-conductual y las herramientas de medida para evaluar la eficacia. También se describió la investigación previa y se indicó el propósito del meta-análisis.

## Método

La sección método debe incluir las siguientes subsecciones: criterios de selección de los estudios, estrategia de búsqueda de los estudios, codificación de las variables moderadoras, cálculo del tamaño del efecto y análisis estadístico. A continuación, presentamos una descripción breve de la información que debe incluir cada una de estas subsecciones.

### *Criterios de selección de los estudios*

Los criterios de selección son los criterios que los estudios primarios deben cumplir para ser incluidos en el meta-análisis. Aunque los criterios de selección dependen del objeto de estudio del meta-análisis, existen varios criterios de selección que pueden presentarse en cualquier meta-análisis, tales como características de las muestras de participantes, rango de edades considerado, tipo de diseño en los estudios primarios, restricciones de idioma y un tamaño muestral mínimo.

En los meta-análisis sobre la eficacia de tratamientos, el enfoque PICOS permite identificar claramente los cinco componentes principales que se deben tener en cuenta cuando se definen los criterios de selección de los estudios: *Participantes* (población/es objetivo a la/s que pertenece/n la/s muestra/s de participantes), *Intervenciones* (tratamiento/s, intervención/es, programa/s a examinar en el meta-análisis), *Comparaciones* (tipo/s de grupo/s de control que se acepta/n en los estudios –grupos control activo vs. inactivo), *Resultados* (variables dependientes que deben reportarse en los estudios) y *Diseño del estudio* (tipos de diseños aceptados: experimental, cuasi-experimental, diseños de un solo grupo) (Liberati et al., 2009).

Por ejemplo, en el meta-análisis sobre TOC de Iniesta-Sepúlveda et al. (2017), los estudios tuvieron que cumplir los siguientes criterios: (a) examinar la eficacia de las terapias cognitivo-conductuales en participantes con TOC mayores de 19 años y diagnosticados mediante un criterio estandarizado; (b) examinar intervenciones con un nivel alto de implicación parental (programas de tratamiento que incluyeran técnicas centradas en los padres y en las cuales los padres se involucran en al menos el 70% de las sesiones); (c) incluir al menos un grupo tratado con medidas pretest y posttest y, opcionalmente, medidas de seguimiento; (d) el tamaño muestral en el posttest debía ser mayor de cuatro participantes; por tanto, los diseños de caso único eran excluidos; (e) se requería que el estudio incluyera la CY-BOCS como medida de re-

sultado; (f) tamaño/s del efecto reportado/s, o al menos información suficiente para calcularlo/s, y (g) estudios escritos en inglés o español.

### *Búsqueda de los estudios*

Para identificar los estudios que cumplen con los criterios de selección, se realiza una búsqueda bibliográfica combinando procedimientos formales e informales. Las bases bibliográficas electrónicas consultadas (e.g., PsycINFO, MedLine, ERIC, Google Académico, Web of Science, Scopus) deben reportarse, incluyendo las palabras claves utilizadas y cómo fueron combinadas (quizás en un apéndice). Para asegurarnos que estamos siendo exhaustivos en este proceso, es conveniente complementar esta estrategia de búsqueda formal con búsquedas manuales en revistas específicas y libros centrados en el tema de interés, revisando las referencias enumeradas en los estudios seleccionados e intentado contactar con investigadores reconocidos en el campo. En esta sección el meta-analista debe indicar de forma explícita si intentó localizar estudios no publicados.

Las bases de datos electrónicas consultadas por Iniesta-Sepúlveda y colaboradores (2017) fueron Medline, PsycINFO, Psychology and Behavioral Sciences Collection (PBSC) y Consejo Superior de Investigaciones Científicas (CSIC), combinando las siguientes palabras clave en el título y en el resumen: ((obsessive-compulsive) or (OCD)) and ((treatment) or (cognitive behavioral therapy) or (CBT) or (exposure response prevention) or (ERP)) and ((family) or (parents)). Además, contactaron con expertos en el campo con el objetivo de localizar estudios no publicados.

Además, es recomendable presentar un diagrama de flujo que resuma el proceso de cribado y selección de los estudios (véase, por ejemplo, Figura 1). Alternativamente, los resultados del proceso de selección de los estudios pueden reportarse en la sección Resultados.

### *Codificación de las variables moderadoras*

El objetivo de esta sección es describir las variables moderadoras extraídas de los estudios primarios. Para tal fin, se debe elaborar un manual de codificación y un protocolo de registro de las características de los estudios, los cuales deben estar disponibles para la comunidad científica, y así garantizar la transparencia y replicabilidad del proceso de codificación. La información relevante extraída de cada estudio primario incluye variables numéricas que serán usadas en los análisis principales (véase la siguiente sección), pero también incluye posibles variables moderadoras de la asociación de interés. Aunque la lista de posibles variables moderadoras varía de un meta-análisis a otro, pueden distinguirse tres categorías generales: variables metodológicas, sustantivas, y extrínsecas (Lipsey, 2009).

Las características sustantivas son aquellas relacionadas con el objeto de estudio del meta-análisis, tales como características sociodemográficas de la muestra y características de

tratamiento. Las variables metodológicas se refieren al tipo de diseño y a los métodos aplicados en los estudios. Estas incluyen aspectos relacionados con la calidad metodológica (o riesgo de sesgo) de los estudios. En el contexto de estudios de intervención, ejemplos de tales características metodológicas son la asignación aleatoria o no de los participantes a los grupos, las tasas de mortalidad en cada grupo, y el cegamiento o no de los participantes y de los evaluadores. El meta-analista debe siempre considerar la posible influencia de la calidad metodológica sobre los resultados antes de comenzar la integración estadística, así como hacer mención explícita a si decidió incluir/excluir estudios en los análisis (e.g., análisis de sensibilidad) en función de su calidad metodológica (Chacón-Moscote, Sanduete-Chaves, y Sánchez-Martín, 2016). Por último, las variables extrínsecas son aquellas características que, en principio, no tienen por qué afectar a los resultados, ya que no tienen nada que ver con el objeto de estudio, pero se ha demostrado que, en ocasiones, están asociadas a los resultados. Son ejemplo de tales variables el estatus de publicación (publicado o no publicado), la formación del autor principal, el país donde se realizó el estudio y la fuente de financiación.

(porcentaje de hombres), la duración media del TOC (en años), y el porcentaje de minorías étnicas. Las características metodológicas codificadas incluyeron el criterio diagnóstico, la mortalidad del pretest al posttest y el control de intervenciones farmacológicas.

En la práctica, el proceso de codificación de los estudios es a menudo complejo debido a la ambigüedad de la información reportada en los estudios primarios. Por tanto, se debe analizar la fiabilidad del proceso de codificación. Para tal propósito, todos los estudios primarios, o bien una muestra aleatoria de los mismos, deben codificarse por dos (o más) codificadores independientes previamente entrenados. La fiabilidad interjueces puede evaluarse utilizando índices tales como las correlaciones intraclass y los coeficientes kappa para moderadores continuos y categóricos, respectivamente. En el meta-análisis, la fiabilidad de este proceso de codificación debe reportarse incluyendo, al menos, los valores mínimo y máximo de acuerdo interjueces y la fiabilidad media.

*Cálculo de un índice del tamaño del efecto*

Un punto esencial en meta-análisis es calcular uno o más tamaños del efecto que resuma/n cuantitativamente los resultados de cada estudio. Diferentes índices del tamaño del efecto pueden aplicarse dependiendo del diseño del estudio y de la naturaleza de las variables medidas (e.g., continuas, dicotómicas). Los índices del tamaño del efecto más frecuentemente utilizados en los meta-análisis son: la familia *d*, la familia *r* y los índices de riesgo.

En cuanto a los índices *d*, uno de los diseños más comunes en psicología y áreas relacionadas consiste en la comparación de dos grupos en una o más variables continuas. En un diseño de dos grupos (e.g., experimental vs. control), la diferencia de medias tipificada es el tamaño del efecto usualmente aplicado, el cual nos permite homogeneizar los resultados cuando diferentes instrumentos de medida (e.g., test psicológicos diferentes que miden un mismo constructo) son utilizados en los estudios. La diferencia de medias tipificada se define como la diferencia entre las dos medias dividida entre una desviación típica conjunta (Hedges y Olkin, 1985; Rubio-Aparicio et al., en prensa). Es importante tener en cuenta que el orden de los grupos en el numerador determina que la mayoría de las estimaciones de los tamaños del efecto sean que signo positivo o negativo. Es recomendable ordenar las medias de tal forma que sean en su mayoría positivas. Esto debe hacerse explícito en el meta-análisis, introduciendo una frase del tipo: "...valores positivos indican que el grupo tratado presenta mejores resultados que el grupo control". Esto ocurre cuando se diseña la intervención para incrementar los valores (por ejemplo, en habilidades sociales), pero el signo debería estar en el orden opuesto cuando lo esperable es reducir los valores (por ejemplo, en ansiedad). En un diseño de medidas repetidas, en el que se toman medidas pretest y posttest continuas en una única muestra de participantes (e.g., antes y después de una intervención, o antes y en el seguimiento), el tamaño del efecto más apropiado es el cambio medio tipificado. Por ejemplo, Iniesta-Sepúlveda et al. (2017) utilizaron como índice del tamaño del efecto el cambio medio tipificado, calculado como la diferencia entre las medias del pretest y del posttest dividida entre la desviación típica del pretest. Si los estudios inclu-

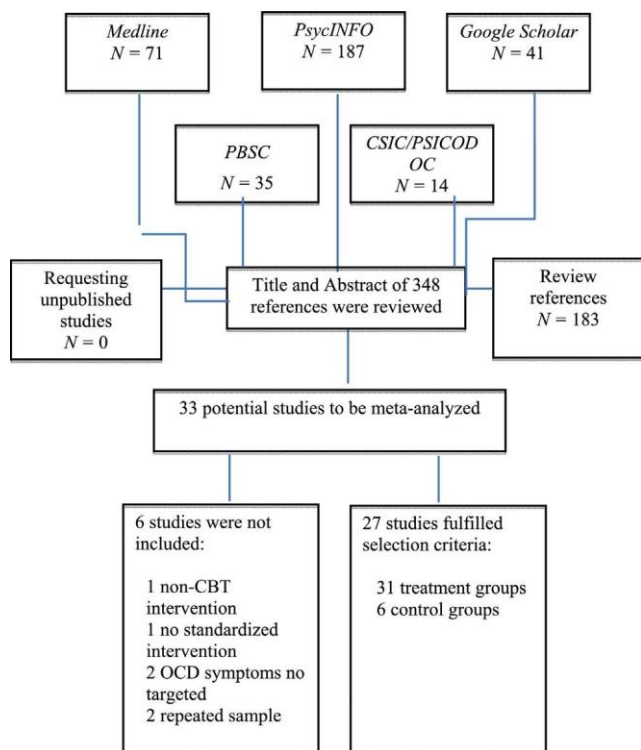


Figura 1. Diagrama de flujo de la estrategia de búsqueda en el meta-análisis de Iniesta-Sepúlveda et al. (2017).

En el meta-análisis de Iniesta-Sepúlveda et al. (2017), las características sustantivas codificadas fueron las técnicas de comportamiento aplicadas, la duración del tratamiento (número de semanas), la inclusión o no de un programa de seguimiento, la edad media de la muestra (en años), la distribución del género

yen grupo experimental y grupo control con medidas pretest y postest, el índice del tamaño del efecto más apropiado es la diferencia de cambio medios tipificados (cf. e.g., Botella y Sánchez-Meca, 2015; Morris, 2008; Rubio-Aparicio et al., en prensa).

Respecto a los índices  $r$ , cuando los estudios primarios aplican un diseño correlacional para analizar el grado de asociación entre dos variables, el índice del tamaño del efecto que debe usarse es un coeficiente de correlación (e.g., el coeficiente de correlación de Pearson, su transformación a  $Z$  de Fisher, el coeficiente  $\phi$ , el coeficiente de correlación biserial puntual, etc.).

Por último, los índices de riesgo se usan para sintetizar resultados dicotómicos. Estos incluyen la diferencia de riesgos, definida como la diferencia entre las proporciones de fracaso (o éxito) en los dos grupos; la razón de riesgos, definida como la razón entre las dos proporciones (riesgos), y la razón de ventajas, definida como la razón entre las ventajas de los dos grupos (Sánchez-Meca, Marín-Martínez, y Chacón-Moscoso, 2003). La posición de las proporciones de los dos grupos en las formulas determina la interpretación de este tipo de tamaños del efecto. Por tanto, cuando el evento de interés es el fracaso, una diferencia de riesgo negativa entre las proporciones del grupo tratado y el grupo control indica un resultado a favor del tratamiento; y las razones de riesgo y las razones de ventaja son menores que uno. Si las proporciones de interés se refieren al éxito, entonces es esperable obtener diferencias de riesgo positivas, y razones de riesgo y razones de ventaja mayores que uno.

Una vez que hemos seleccionado el índice del tamaño del efecto más apropiado dadas las características de los estudios, se calcula el tamaño del efecto de cada estudio individual. En la literatura podemos encontrar las diferentes fórmulas para calcular las varianzas muestrales en función del tipo de índice del tamaño del efecto (e.g., Sánchez-Meca y Marín-Martínez, 2010; véase Tabla 1). En el reporte del meta-análisis, debe aparecer una descripción clara sobre cómo se han extraído/calculado los tamaños del efecto de los estudios primarios y cómo interpretarlos. Además, al igual que en el proceso de codificación de las características de los estudios, en el cálculo de los tamaños del efecto es recomendable analizar la fiabilidad interjueces y reportar los resultados de dicho análisis (e.g., mínimo, máximo, y fiabilidad media).

#### *Análisis estadístico*

Una vez que el proceso de extracción de datos ha terminado, el siguiente paso consiste en describir los análisis estadísticos. Especificar el marco estadístico (frecuentista o Bayesiano) y el modelo meta-analítico asumido en los análisis son aspectos esenciales en esta sección. En este sentido, dos principales modelos estadísticos son típicamente considerados en meta-análisis: el modelo de efecto fijo y el modelo de efectos aleatorios.

Cuando el meta-analista aplica un modelo de efecto fijo está asumiendo que todos los estudios en el meta-análisis estiman un tamaño del efecto paramétrico común, la única fuente de variabilidad entre los tamaños del efecto es el error de muestreo debido a la selección aleatoria de los participantes en cada estudio (Konstantopoulos y Hedges, 2009). En el modelo de efectos aleatorios se asume que cada estudio en el meta-análisis estima su propio tamaño del efecto paramétrico, y que los estudios son seleccionados aleatoriamente de una población de estudios.

Como consecuencia, el modelo de efectos aleatorios asume dos fuentes de variabilidad: la variabilidad inter-estudios y la variabilidad intra-estudio. Actualmente, existe un amplio consenso en considerar el modelo de efectos aleatorios más realista que el modelo de efecto fijo en la mayoría de situaciones, debido a las diferencias metodológicas y sustantivas encontradas habitualmente entre los estudios que conforman el meta-análisis (Borenstein, Hedges, Higgins, y Rothstein, 2010; Hedges y Vevea, 1998; Raudenbush, 1994, 2009).

La elección del modelo influye en los procedimientos estadísticos utilizados para integrar la información y en la generalización de los resultados (Hedges y Vevea, 1998). El meta-analista debe explicar la razón de asumir un modelo de efecto fijo o un modelo de efectos aleatorios en el análisis de sus datos. Por ejemplo, en el meta-análisis de Iniesta-Sepúlveda et al. (2017) se asumieron modelos de efectos aleatorios para adaptarse a la variabilidad exhibida en los tamaños del efecto. Independientemente del modelo estadístico asumido, los tamaños del efecto individuales se ponderan en función de su precisión. El meta-analista debe especificar el procedimiento de ponderación aplicado en los análisis (e.g., ponderar por la inversa de la varianza, por el tamaño muestral, o no ponderar). Si procede, las fórmulas utilizadas para calcular la varianza intra-estudio de cada tamaño del efecto, la varianza inter-estudios, u otro método de ponderación deberían ser reportadas en el artículo (o ser incluidas como material suplementario).

Otros aspectos relevantes que tienen que reportarse en esta sección son cómo se ha calculado el tamaño del efecto medio y el intervalo de confianza y si se ha construido un forest plot u otros gráficos para presentar los resultados. Además, el meta-analista debe especificar cómo ha estudiado la heterogeneidad entre los tamaños del efecto (e.g., con la prueba  $Q$ , el índice  $I^2$  u otras estrategias analíticas). Este es un punto importante, pues la prueba  $Q$  tiene pobre potencia estadística para detectar la verdadera heterogeneidad entre los tamaños del efecto cuando los meta-análisis incluyen un número pequeño de estudios (Sánchez-Meca y Marín-Martínez, 1997). Por tanto, es recomendable complementar la conclusión estadística alcanzada con la prueba  $Q$  con el índice  $I^2$  (Higgins y Thompson, 2002), el cual cuantifica la heterogeneidad exhibida en los tamaños como un porcentaje (Huedo-Medina, Sánchez-Meca, Marín-Martínez, y Botella, 2006).

Si se encuentra una heterogeneidad considerable entre los tamaños del efecto (como suele ocurrir), entonces el meta-analista debe describir cómo ha realizado la búsqueda de variables moderadoras. El modelo estadístico asumido en estos análisis debe también especificarse. Una recomendación general cuando se lleva a cabo el análisis de variables moderadoras es adoptar un modelo de efectos mixtos, en el que los tamaños del efecto se toman como una variable de efectos aleatorios, y los niveles de estudio de las variables moderadoras – o los niveles individuales de los moderadores, si los datos de los participantes individuales están disponibles – se toman como efectos fijos. Otros aspectos que deben puntualizarse son si se han aplicado modelos de meta-regresión para analizar variables continuas y ANOVAs ponderadas para el análisis de variables cualitativas, y si se han calculado las proporciones de varianza explicadas por las variables moderadoras (López-López, Marín-Martínez, Sánchez-Meca, Van den Noortgate, y Viechtbauer, 2014).

Si procede, deben describirse análisis adicionales, tales como análisis de sensibilidad para detectar valores atípicos o para abordar la dependencia estadística entre los tamaños del efecto de la misma muestra o estudio. Las estructuras dependientes están presentes en los meta-análisis llevados a cabo en psicología y ciencias sociales, y algunas estrategias para tratar este problema consisten en incluir un único tamaño del efecto por muestra en cada meta-análisis, calcular una media aritmética de los tamaños del efecto dependientes, o aplicar técnicas meta-analíticas que tienen en cuenta estructuras dependientes tales como el meta-análisis multivariado, modelos de tres niveles y aproximaciones de estimación robusta de la varianza (López-López, Van den Noortgate, Tanner-Smith, Wilson, y Lipsey, 2017).

Si el meta-analista considera que el sesgo de publicación podría ser una amenaza para sus resultados, debe explicitar qué métodos ha aplicado para estudiar tal sesgo (e.g., funnel plots con el método de “recorte-y-relleno” para imputar efectos perdidos, el índice de tolerancia a los resultados nulos, la prueba de Egger, la correlación de rangos de Begg y Mazumbar, etc.; cf. e.g., Rothstein, Sutton, y Borenstein, 2005).

Un último aspecto que debe incluirse en esta sección es el programa estadístico utilizado para llevar a cabo los análisis estadísticos. Algunas de las elecciones más populares son metafor en R, Comprehensive Meta-Analysis 3.3 (Borenstein, Hedges, Higgins, y Rothstein, 2014), y RevMan 5.3 (Review Manager, 2014).

## Resultados

Los principales resultados que cualquier meta-análisis debe incluir son: las características descriptivas de los estudios individuales y sus tamaños del efecto, la estimación del tamaño del efecto promedio y su intervalo de confianza, la evaluación de la heterogeneidad, la valoración de la posible existencia de sesgo de publicación, y el análisis de variables moderadoras.

### *Distribución de los tamaños del efecto, tamaño del efecto promedio y heterogeneidad*

Como un primer paso, es recomendable presentar, como un apéndice o material suplementario, la base de datos con los tamaños del efecto y las características de cada uno de los estudios individuales incluidos en el meta-análisis. En general, esta base de datos está formada por tantas filas como estudios hayamos incluido en el meta-análisis, y las diferentes columnas están formadas por el identificador del estudio, algunas de las características codificadas de los estudios, y las estimaciones de los tamaños del efecto y tamaños muestrales. Un ejemplo de este tipo es el Apéndice A en el meta-análisis de Iniesta-Sepúlveda et al. (2017).

Después de la sección descriptiva, debe presentarse el tamaño del efecto medio junto con su intervalo de confianza y las medidas de heterogeneidad (e.g., la prueba  $Q$  y el índice  $I^2$ ). Si se han considerado varias medidas de resultado y/o grupos (e.g., grupo tratado y grupo control), es realmente importante presentar sus resultados de forma separada. Además, se puede ilustrar la distribución de los tamaños del efecto utilizando técnicas gráficas. En la Figura 2 se muestra un ejemplo de forest plot, uno de los gráficos más comúnmente utilizados en meta-análisis. Sin

embargo, si contamos con un número muy grande de tamaños del efecto, los histogramas o los diagramas de tallo y hoja son gráficos más interesantes para mostrar simultáneamente los tamaños del efecto y su distribución.

### *Análisis de variables moderadoras*

Si los tamaños del efecto individuales mostraron gran heterogeneidad, deben analizarse las características de los estudios que pudieran estar afectando a tal variabilidad. Los resultados de análisis de moderadores se presentan de forma separada para moderadores categóricos y continuos.

Para moderadores categóricos, los resultados se presentan en formato ANOVA, reportando para cada categoría del moderador el tamaño del efecto promedio, el intervalo de confianza y las medidas de heterogeneidad, y opcionalmente las comparaciones múltiples entre los tamaños del efecto promedio de algunas categorías (ajustando la inflación de la tasa de error Tipo I) en moderadores con más de dos categorías. Para moderadores continuos, se presentan los coeficientes de regresión, el error típico y los límites confidenciales. Además, para ambos tipos de moderadores, deben reportarse las pruebas de significación, las pruebas de especificación del modelo, y la proporción de varianza explicada. Como ejemplo, véanse las Tablas 2, 4 y 5 del meta-análisis de Iniesta-Sepúlveda et al. (2017).

Es aconsejable complementar los análisis de moderadores individuales (ANOVAs y meta-regresión) con modelos de meta-regresión múltiple que incluyan simultáneamente las variables moderadoras más relevantes (Rubin, 1992). El ajuste de modelos de meta-regresión múltiple permite estudiar la influencia de una variable moderadora controlando el efecto de otros moderadores también importantes. No siempre es posible ajustar un modelo predictivo como un paso adicional en meta-análisis. Por ejemplo, el ajuste de un modelo de meta-regresión múltiple no tiene sentido cuando el número de estudios incluidos en el meta-análisis no es lo suficientemente grande como para lograr estimaciones precisas, o cuando ninguna variable moderadora alcanza individualmente la significación estadística. En el meta-análisis de Iniesta-Sepúlveda et al., (2017), la Tabla 3 presenta los resultados de dos modelos de meta-regresión múltiple.

### *Análisis adicionales*

En esta sección, si procede, deben reportarse los resultados de cualquier análisis adicional realizado. Un meta-análisis debe incluir análisis estadísticos adicionales relativos al estudio del sesgo de publicación, la estimación de la potencia estadística en análisis de moderadores (Cafri et al. 2010) y los análisis de sensibilidad comparando los resultados obtenidos con diferentes procedimientos meta-analíticos (Botella y Sánchez-Meca, 2015).

Un análisis adicional altamente recomendado es el estudio del sesgo de publicación. El meta-analista debe evaluar hasta qué punto el sesgo de publicación puede ser una amenaza para las conclusiones alcanzadas en su meta-análisis. Existen diferentes técnicas para estudiar la existencia de sesgo de publicación: el índice de tolerancia a los resultados nulos (Rothstein et al. 2005), la prueba de Egger (Egger, Smith, Schneider, y Minder, 1997), el método de “recorte-y-relleno” (Duval y Tweedie, 2000), y el funnel plot (Light y Pillemer, 1984). Nosotros reco-

mendamos la combinación de al menos tres de estos procedimientos: el funnel plot, que sugiere la existencia de sesgo de publicación cuando la distribución de los tamaños del efecto contra sus errores típicos no es simétrica; la prueba de Egger, que evalúa la asimetría del funnel plot basándose en un modelo de

regresión simple, y el método de “recorte-y-relleno”, que permite estimar los tamaños del efecto “perdidos” que harían que el funnel plot fuera simétrico, y una vez añadidos tales valores perdidos vuelve a estimar el tamaño del efecto medio.

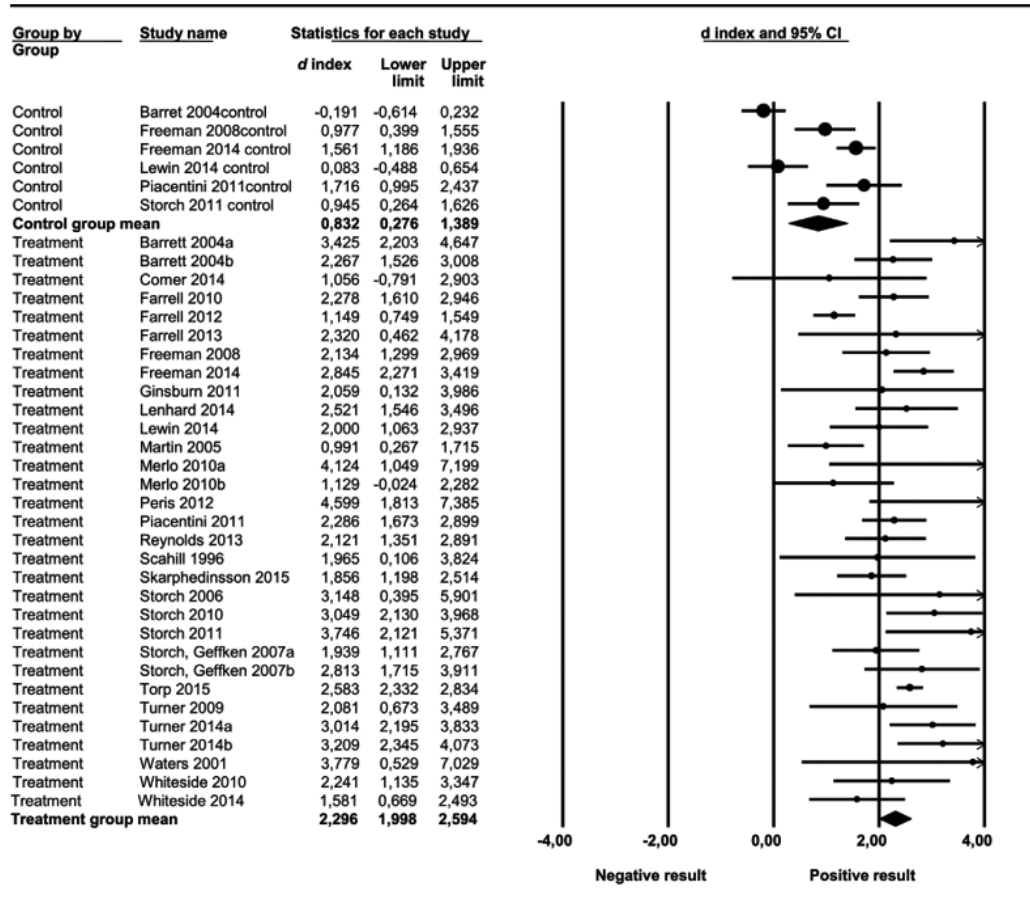


Figura 2. Forest plot de los tamaños del efecto obtenidos para los síntomas del TOC agrupados en función del tipo de intervención (grupo tratado vs. grupo control). Extraído de Iniesta-Sepúlveda et al. (2017).

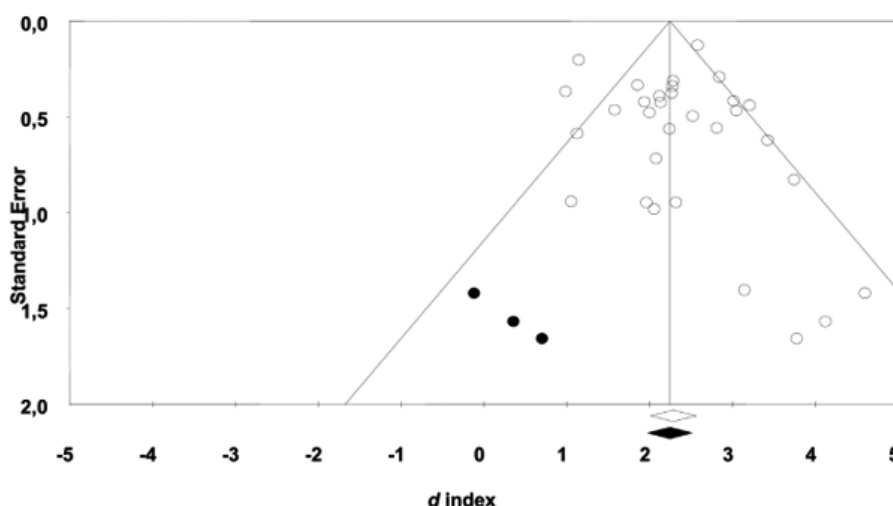
En nuestro ejemplo, Iniesta-Sepúlveda y colaboradores (2017) aplicaron estos tres procedimientos. La Figura 3 presenta el funnel plot de los índices *d* aplicando el método de “recorte-y-relleno”. Aunque el método de “recorte-y-relleno” imputó tres valores para lograr la simetría del funnel plot, la diferencia entre el efecto medio original y el efecto obtenido después de incorporar esos tres valores fue despreciable. Por tanto, se descartó el sesgo de publicación como una amenaza a los resultados de este meta-análisis.

**Discusión**

En esta sección, la primera afirmación que debe aparecer es un resumen de los principales resultados del meta-análisis, tales como el tamaño del efecto promedio y los análisis de moderadores. Después, tales resultados deben discutirse a la luz de la evidencia previa. En psicología y áreas afines, los meta-análisis

sobre la efectividad de diferentes tratamientos son muy comunes. En este contexto, los análisis de moderadores proporcionan información valiosa sobre bajo qué condiciones cada tratamiento es más efectivo. Esta es una de las ventajas más importantes de este tipo de investigación, pues los estudios primarios se encuentran limitados a sus condiciones examinadas. Independientemente del tipo de meta-análisis, en la sección Discusión el meta-analista debe concluir si se han alcanzado los objetivos propuestos y, si procede, si los resultados meta-analíticos confirman o no las hipótesis previas formuladas.

Por otro lado, es importante establecer las implicaciones que los resultados pueden tener en la práctica psicológica. En nuestro ejemplo, los autores proporcionan directrices y recomendaciones claras y explícitas para los profesionales aplicados en cuanto a la efectividad del tratamiento familiar cognitivo-conductual en el TOC pediátrico en diferentes escenarios.



**Figure 3.** Funnel plot de los 31 índices *d* para los grupos tratados (círculos blancos) junto con los tres valores imputados (círculos negros) aplicando el método de “recorte-y-relleno” de Duval y Tweedie (2000). Extraído de Iniesta-Sepúlveda et al. (2017).

Como cualquier estudio empírico, un meta-análisis no está exento de limitaciones. Por tanto, es importante discutir las limitaciones que pudieran afectar al alcance y generalización de los resultados meta-analíticos. Por último, un meta-análisis debe terminar la sección Discusión ofreciendo recomendaciones para la investigación futura en el campo de estudio.

### Referencias

Si las referencias de los estudios incluidos en el meta-análisis no se reportan como un apéndice o material suplementario, es aconsejable marcarlas (e.g., con un asterisco) en la sección Referencias.

### Conclusiones finales

Los meta-análisis han ganado popularidad en la mayoría de los campos científicos. La práctica basada en la evidencia es un claro ejemplo del amplio alcance de esta metodología. Las conclusiones extraídas de un meta-análisis permiten a los investigadores sintetizar la evidencia sobre un determinado tema de estudio.

### Referencias

APA Publications y Communications Board Working Group on Journal Article Reporting Standards. (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist*, *63*, 839-851. doi: 10.1037/0003-066X.63.9.839

Borenstein, M., 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. doi:https://doi.org/10.1002/jrsm.12

Borenstein, M., Hedges L. V., Higgins, J. P. T., y Rothstein, H. R. (2014). *Comprehensive Meta-Analysis* (Vers. 3.3). Englewood, NJ: Biostat.

Botella, J., y Gambara, H. (2006). Doing and reporting a meta-analysis. *International Journal of Clinical and Health Psychology*, *6*, 425-440.

Botella, J., y Sánchez-Meca, J. (2015). *Meta-análisis en ciencias sociales y de la salud*. Madrid, Spain: Síntesis.

Cafri, G., Kromrey, J. D., y Brannick, M. T. (2010). A meta-meta-analysis: Empirical review of statistical power, type I error rates, effect sizes, and model

selection of meta-analyses published in psychology. *Multivariate Behavioral Research*, *45*, 239-270. doi: 10.1080/00273171003680187

Chacón-Moscoso, S., Sanduete-Chaves, S., y Sánchez-Martín, M. (2016). The development of a checklist to enhance methodological quality in intervention programs. *Frontiers in Psychology*, *7*, 1811. doi:10.3389/fpsyg.2016.01811.

Cook, D.J., Sackett, D.L., y Spitzer, W.O. (1995). Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam consultation on meta-analysis. *Journal of Clinical Epidemiology*, *48*, 167-171.

Duval, S., y Tweedie, R. (2000). Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455-463. doi: 10.1111/j.0006-341X.2000.00455.x

Egger, M., Smith, G. D., Schneider, M., y Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629-634.

Sin embargo, hacer y reportar un meta-análisis no es tarea fácil. El propósito de este artículo era ofrecer algunas recomendaciones para el reporte de un meta-análisis. Los meta-análisis son a menudo una tarea desafiante, ya que hacer un meta-análisis supone la aplicación de métodos estadísticos complejos y/o la toma de decisiones importantes por parte del equipo meta-analítico. Las conclusiones extraídas de los meta-análisis pueden tener un impacto sustancial, y por ello, existe una importante necesidad de hacer y reportar meta-análisis con un alto nivel de calidad. La transparencia debe estar presente cuando en un meta-análisis reportamos la decisiones que hemos tomado a lo largo de todo el proceso, como una forma de alcanzar tales estándares. Además, proporcionar una descripción detallada de cada una de las fases y decisiones tomadas permite que otros investigadores puedan replicar nuestro meta-análisis, así como hacer una evaluación crítica de sus resultados.

**Agradecimientos.-** Este trabajo ha sido financiado por el Ministerio de Economía, Industria y Competitividad y los fondos FEDER (proyecto n° PSI2016-77676-P).



- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3-8. doi: <https://doi.org/10.3102/0013189X005010003>
- 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. doi: <https://doi.org/10.1037/1082-989X.3.4.486>
- Higgins, J. P. T., y Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21, 1539-1558. doi: <https://doi.org/10.1002/sim.1186>
- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., y Botella, J. (2006). Assessing heterogeneity in meta-analysis:  $Q$  statistic or  $I^2$  index? *Psychological Methods*, 11, 193-206. doi: [10.1037/1082-989X.11.2.193](https://doi.org/10.1037/1082-989X.11.2.193)
- Hutton, B., Salanti, G., Caldwell, D. M., Chaimani, A., Schmid, C. H., Cameron, C., ... y Mulrow, C. (2015). The PRISMA Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-analyses of Health Care Interventions: Checklist and Explanations. *Annals of Internal Medicine*, 162, 777-784. doi: [10.7326/M14-2385](https://doi.org/10.7326/M14-2385)
- Iniesta-Sepúlveda, M., Rosa-Alcázar, A. I., Sánchez-Meca, J., Parada-Navas, J. L., y Rosa-Alcázar, Á. (2017). Cognitive-behavioral high parental involvement treatments for pediatric obsessive-compulsive disorder: A meta-analysis. *Journal of Anxiety Disorders*, 49, 53-64. doi: [10.1016/j.janxdis.2017.03.010](https://doi.org/10.1016/j.janxdis.2017.03.010)
- Konstantopoulos, S., y Hedges, L. V. (2009). Analyzing effect sizes: Fixed-effects models. In H. Cooper, L. V. Hedges, y J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 279-293). New York, NY: Russell Sage Foundation.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P., ... y Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology*, 62, e1-e34. doi: [10.1136/bmj.b2700](https://doi.org/10.1136/bmj.b2700)
- Light, R. J., y Pillemer, D. B. (1984). *Summing up. The science of reviewing research*. Cambridge, MA: Harvard University Press.
- Lipsey, M. W. (2009). Identifying interesting variables and analysis opportunities. In H. Cooper y L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (2nd ed., pp. 147-158). New York, NY: Russell Sage Foundation.
- López-López, J. A., Marín-Martínez, F., Sánchez-Meca, J., Van den Noortgate, W., y Viechtbauer, W. (2014). Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. *British Journal of Mathematical and Statistical Psychology*, 67, 30-48. doi: [10.1111/bmsp.12002](https://doi.org/10.1111/bmsp.12002)
- López-López, J. A., Van den Noortgate, W., Tanner-Smith, E. E., Wilson, S. J., y Lipsey, M. W. (2017). Assessing meta-regression methods for examining moderator relationships with dependent effect sizes: A Monte Carlo simulation. *Research Synthesis Methods*, 8, 435-450. doi: [10.1002/jrsm.1245](https://doi.org/10.1002/jrsm.1245)
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., y Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet*, 354, 1896-1900.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., y Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6, e1000097. doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)
- Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, 11, 364-386. doi: <https://doi.org/10.1177/1094428106291059>
- Panic, N., Leoncini, E., de Belvis, G., Ricciardi, W., y Boccia, S. (2013). Evaluation of the endorsement of the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PLOS ONE*, 8(12). doi: [10.1371/journal.pone.0083138](https://doi.org/10.1371/journal.pone.0083138)
- Raudenbush, S. W. (1994). Random effects models. In H. Cooper, y L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301-321). New York: Russell Sage Foundation.
- Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, y J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 295-315). New York: Russell Sage Foundation.
- Review Manager (2014). *RevMan* (Version 5.3) [Computer software]. Copenhagen, Denmark: The Nordic Cochrane Centre, The Cochrane Collaboration.
- Rothstein, H. R., Sutton, A. J., y Borenstein, M. (Eds.) (2005). *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. New York: Wiley.
- Rubin, D. B. (1992). Meta-analysis: Literature synthesis or effect-size surface estimation? *Journal of Educational Statistics*, 17, 363-374. doi: [10.3102/10769986017004363](https://doi.org/10.3102/10769986017004363)
- Rubio-Aparicio, M., Marín-Martínez, F., Sánchez-Meca, J., y López-López, J. A. (en prensa). A methodological review of meta-analyses about the effectiveness of clinical psychology treatments. *Behavior Research Methods*. doi: [10.3758/s13428-017-0973-8](https://doi.org/10.3758/s13428-017-0973-8)
- Rubio-Aparicio, M., Sánchez-Meca, J., López-López, J. A., Marín-Martínez, F., y Botella, J. (2017). Analysis of categorical moderators in mixed-effects meta-analysis: Consequences of using pooled versus separate estimates of the residual between-studies variances. *British Journal of Mathematical and Statistical Psychology*, 70, 439-456. doi: [10.1111/bmsp.12092](https://doi.org/10.1111/bmsp.12092)
- Sánchez-Meca, J., y Botella, J. (2010). Revisión sistemática y meta-análisis: Herramientas para la práctica profesional. *Papeles del Psicólogo*, 31, 7-17.
- Sánchez-Meca, J., López-López, J. A., y López-Pina, J. A. (2013). Some recommended statistical analytic practices when reliability generalization (RG) studies are conducted. *British Journal of Mathematical and Statistical Psychology*, 66, 402-425. doi: [10.1111/j.2044-8317.2012.02057.x](https://doi.org/10.1111/j.2044-8317.2012.02057.x)
- Sánchez-Meca, J., López-Pina, J. A., Rubio-Aparicio, M., Marín-Martínez, F., Núñez-Núñez, R. M., López-García, J. J., y López-López, J. A. (2017, Julio). REGEMA: Propuesta de una guía para la realización y reporte de meta-análisis de generalización de la fiabilidad. Conferencia presentada en el XV Congreso de Metodología de las Ciencias Sociales y de la Salud, Barcelona (España).
- Sánchez-Meca, J. y Marín-Martínez, F. (1997). Homogeneity tests in meta-analysis: A Monte Carlo comparison of statistical power and Type I error. *Quality and Quantity*, 31, 385-399.
- Sánchez-Meca, J. y Marín-Martínez, F. (2010). Meta-analysis in psychological research. *International Journal of Psychological Research*, 3, 151-163.
- Sánchez-Meca, J., Marín-Martínez, F., y Chacón-MoscOSO, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, 8, 448-467. doi: [10.1037/1082-989X.8.4.448](https://doi.org/10.1037/1082-989X.8.4.448)
- 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 differences in results. *British Journal of Mathematical and Statistical Psychology*, 62, 97-128. doi: [10.1348/000711007X255327](https://doi.org/10.1348/000711007X255327)
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., ... y Bouter, L. M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, 7, 10. doi: [10.1186/1471-2288-7-10](https://doi.org/10.1186/1471-2288-7-10)
- Stewart, L. A., Clarke, M., Rovers, M., Riley, R. D., Simmonds, M., Stewart, G., y Tierney, J. F. (2015). Preferred reporting items for a systematic review and meta-analysis of individual participant data: the PRISMA-IPD statement. *Journal of the American Medical Association*, 313, 1657-1665. doi: [10.1001/jama.2015.3656](https://doi.org/10.1001/jama.2015.3656)
- Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., ... y Thacker, S. B. (2000). Meta-analysis of observational studies in epidemiology: A proposal for reporting. *Journal of the American Medical Association*, 283, 2008-2012. doi: [10.1001/jama.283.15.2008](https://doi.org/10.1001/jama.283.15.2008)
- Valentine, J. C., Cooper, H., Patall, E. A., Tyson, D., y Robinson, J. C. (2010). A method for evaluating research syntheses: The quality, conclusions, and consensus of 12 syntheses of the effects of after-school programs. *Research Synthesis Methods*, 1, 20-38. doi: [10.1002/jrsm.3](https://doi.org/10.1002/jrsm.3)
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36, 1-48. doi: [10.18637/jss.v036.i03](https://doi.org/10.18637/jss.v036.i03)

(Artículo recibido: 01-02-2018; revisado: 24-02-2018; aceptado: 02-03-2018)