

REFERENCIAS BIBLIOGRÁFICAS

- Aronson, E., et al. (1990): *Methods of Research in Social Psychology*, 2.ª ed., McGraw-Hill, Inc.
- Brown, S. R., y Melamed, L. E. (1993): «Experimental Design and Analysis», en Michael S. Lewis-Beck (ed.), *Experimental Design & Methods*, Sage Publications, Part II, págs.75-160.
- Fortin, D. R., y Greenlee, T. B. (1998): «Using a Product/Service Evaluation Frame: An Experiment on the Economic Equivalence of Product versus Service Alternatives for Message Retrieval Systems», *Journal of Business Research*, 41, págs. 205-214.
- Gourville, J. T. (1998): «Pennies-a-Day: The Effect of Temporal Reframing on Transaction Evaluation», *Journal of Consumer Research*, 24, págs. 395-408.
- Green, P. E.; Tull, D. S., y Albaum, G. (1988): *Research for Marketing Decisions*, 5.ª ed., Prentice-Hall, Englewood Cliffs, Nueva Jersey.
- Hair, J. F., Jr., et al. (1998): *Multivariate Data Analysis*, 5.ª ed., Prentice-Hall International, Nueva Jersey.
- Keppel, G. (1991): *Design and Analysis: A Researcher's Handbook*, 3.ª ed., Prentice-Hall, Englewood Cliffs, Nueva Jersey.
- Lee, H.; Lindquist, J. D., y Acito, F. (1997): «Managers Evaluation of Research Design and Its Impact on the Use of Research: An Experimental Approach», *Journal of Business Research*, 39, págs. 231-240.
- Malhotra, N. K. (1996): *Investigación de mercados. Un enfoque práctico*, 2.ª ed., Prentice-Hall Hispanoamericana.
- Spector, P. E. (1993): «Research Designs», en Michael S. Lewis-Beck (ed.), *Experimental Design & Methods*, Sage Publications, Part I, págs. 1-74.
- Sternthanl, B.; Tybout, A. M., y Calder, B. J. (1994): «Experimental Design: Generalization and Theoretical Explanation», en Richard P. Bagozzi (ed.), *Principles of Marketing Research*, Basil Blackwell, Ltd., Cambridge.

7

Metaanálisis para la investigación científica

Julio Sánchez-Meca

En una época caracterizada por el crecimiento exponencial de la producción científica, las denominadas *revisiones de la investigación* constituyen un eslabón imprescindible entre el quehacer científico del pasado y del futuro, estableciendo el punto de partida de nuevas investigaciones. Tradicionalmente, las revisiones de la investigación se han caracterizado por carecer de una pauta sistemática en la toma de decisiones a lo largo del proceso, así como por la ausencia de referentes cuantitativos en la valoración de los resultados. De ahí la denominación de *revisiones narrativas* o *cualitativas*, donde, tras el estudio de las investigaciones más relevantes con relación a un determinado tema, el revisor expone su particular visión del estado de la cuestión (Cooper, 1989; Glass, McGaw y Smith, 1981; Rosenthal, 1991).

Frente a esta práctica habitual, durante las dos últimas décadas ha surgido el *metaanálisis* como una nueva metodología cuyo propósito es dotar a las revisiones de la investigación del rigor, objetividad y sistematización necesarios para alcanzar una fructífera acumulación del conocimiento científico. Uno de los aspectos claramente diferenciadores del metaanálisis respecto de las revisiones narrativas es su carácter cuantitativo. En efecto, del mismo modo que una investigación empírica requiere del uso de técnicas de análisis estadístico de sus datos, así también el metaanálisis aplica procedimientos estadísticos específicamente diseñados para integrar los resultados de un conjunto de estudios empíricos.

El objeto de este capítulo es ofrecer una panorámica práctica del metaanálisis. Para ello, se presenta en primer lugar una conceptualización del metaanálisis y sus fases de realización. A continuación, se centra la atención en uno de los tópicos fundamentales del metaanálisis: cómo medir de forma uniforme los resultados de los estudios empíricos. Seguirá la presentación de uno de los enfoques metaanalíticos más frecuentemente aplicados en la investigación en Economía de la Empresa: el metaanálisis psicométrico. Finalmente, estas técnicas se aplican a un ejemplo real. Las ideas que se tratan en este capítulo pueden ser ampliadas en diversos textos de metaanálisis (Cooper, 1989; Cooper y Hedges, 1994; Farley y Lehmann, 1986; Glass,

McGaw y Smith, 1981; Gómez, 1987; Hedges y Olkin, 1985; Hunter y Schmidt, 1990; Hunter, Schmidt y Jackson, 1982; Rosenthal, 1991).

1. METAANÁLISIS: ANTECEDENTES Y CONCEPTUACIÓN

Aunque sus orígenes se remontan a los años treinta, no fue hasta finales de la década de los setenta cuando comenzó a cobrar popularidad y se acuñó el término *metaanálisis* para denominar a este tipo de revisión de la investigación. El metaanálisis comenzó a aplicarse en las ciencias sociales en tres campos concretos. En primer lugar, cabe destacar los trabajos de Schmidt y Hunter de 1977 sobre la generalización de la validez de los tests de selección de personal en el ámbito de la psicología del trabajo y de las organizaciones. Los otros dos campos de aplicación original del metaanálisis fueron preñonizados por Rosenthal en 1976, con sus estudios sobre los efectos de las expectativas del experimentador sobre los resultados de las investigaciones, y por Smith y Glass, que en 1977 lo aplicaron al estudio de la efectividad de los tratamientos psicológicos.

Estos esfuerzos de integración cuantitativa de la investigación han dado lugar a una nueva metodología que hoy día se conoce como metaanálisis, aunque también se utilizan otros términos, tales como «síntesis cuantitativa» (*quantitative synthesis*) y «síntesis de la investigación» (*research synthesis*). El metaanálisis se concibe como una propuesta metodológica para la integración de las investigaciones empíricas sobre un mismo tema, caracterizada por la aplicación de las mismas normas de rigor científico exigidas a los estudios primarios y, por tanto, basada en la aplicación de los métodos estadísticos (Cooper y Hedges, 1994; Glass, McGaw y Smith, 1981; Rosenthal, 1991; Sánchez-Meca y Ato, 1989). A diferencia de los estudios primarios, en un metaanálisis los datos son los resultados de los estudios; convenientemente transformados a una métrica común que permita su integración y comparación cuantitativa.

Desde el metaanálisis se aportan normas generales que deben seguirse a lo largo de todas sus fases de realización, desde la formulación del problema, pasando por la selección de los estudios primarios, la codificación de las variables, el análisis de los resultados y la extracción de conclusiones, y finalizando con las normas de publicación del mismo. El metaanalista dispone de cierta flexibilidad en la aplicación de estas normas, pero en todo momento debe ser consciente de las posibles amenazas contra la validez y fiabilidad de las decisiones adoptadas en el proceso. Al hacer explícitas todas las decisiones tomadas en el proceso de revisión de los estudios, cualquier metaanálisis puede ser valorado objetivamente, e incluso replicado, por otros investigadores desde una perspectiva estrictamente científica.

Aunque todos los metaanálisis poseen ciertas características comunes, existen varios enfoques metaanalíticos cuyas diferencias se encuentran fundamentalmente en los objetivos específicos que persiguen y en las técnicas de análisis estadístico que aplican (Bangert-Drowns, 1986). En este capítulo se describe el metaanálisis psico-

métrico desarrollado por Hunter y Schmidt, que se basa en el análisis minucioso de la variabilidad de los resultados interestudios controlando los múltiples *artefactos estadísticos* que pueden sesgar los resultados de las investigaciones, tales como el error de muestreo, la falta de fiabilidad o de validez de las medidas, el fenómeno de la restricción del rango de las variables o el efecto de la dicotomización de las mismas (Hunter y Schmidt, 1990; Hunter, Schmidt y Jackson, 1982).

2. FASES DE UN METAANÁLISIS

Desde que un investigador decide desarrollar un metaanálisis sobre una determinada temática hasta que concluye con la publicación del mismo, se suceden una serie de etapas a lo largo de las cuales el metaanalista tendrá que tomar decisiones, que deberán quedar claramente especificadas. Existe un amplio consenso en cuanto a la formulación y secuencia de dichas etapas, siendo muy similares a las propias de cualquier investigación empírica. Así pues, se pueden establecer las siguientes fases de un metaanálisis: 1) formulación del problema; 2) búsqueda de la literatura; 3) codificación de las variables; 4) análisis estadístico e interpretación, y 5) publicación del estudio.

2.1. Formulación del problema

La primera tarea del metaanalista es delimitar la relación conceptual que se pretende estudiar a través de la literatura disponible, definiendo las variables y los conceptos implicados; por ejemplo, puede ser objeto de un metaanálisis el estudiar la relación entre el compromiso organizacional y el volumen de ventas, entre el rendimiento en el trabajo y el absentismo laboral, la efectividad de la publicidad comparativa frente a la no comparativa, la validez de las entrevistas de empleo, las diferencias de género en el estrés laboral o la relación entre las cinco grandes dimensiones de la personalidad y el rendimiento laboral.

Una vez definidas conceptualmente las variables implicadas, es preciso especificar sus diversas operativizaciones, que pueden encontrarse en los estudios empíricos. Estas definiciones operativas constituyen un criterio importante para la inclusión de los estudios en el metaanálisis. Por otra parte, se debe distinguir entre las variables fundamentales de la relación (dependientes e independientes) y las variables moderadoras de dicha relación, debiéndose anticipar cuáles pueden ser las más influyentes y el sentido de tal influencia.

2.2. Búsqueda de la literatura

Tras la formulación del problema, el paso siguiente consiste en localizar los estudios empíricos que han tratado la relación conceptual. Los cada vez más sofisticados medios de búsqueda bibliográfica informatizada posibilitan la localización

sistemática de investigaciones sobre un determinado tema, a través de descriptores y palabras clave. No obstante, estos procedimientos no cubren todo el campo de posibles publicaciones, por lo que deben utilizarse de forma complementaria otras técnicas menos sistemáticas. Una de ellas consiste en realizar búsquedas descendentes, o ascendentes, a partir de las referencias de los artículos y trabajos que el metaanalista va recogiendo sobre el problema objeto de interés (Cooper, 1989). También puede contactarse con autores o grupos de investigación especialistas en el tema para recuperar trabajos difícilmente accesibles desde los sistemas de búsqueda informatizada, especialmente trabajos no publicados.

Los trabajos susceptibles de ser incluidos en un metaanálisis serán aquellos que cumplan con una serie de criterios de selección que el metaanalista debe especificar previamente. Aunque tales criterios diferirán de un metaanálisis a otro, los siguientes suelen ser muy frecuentes:

- Los trabajos deben haber sido realizados (o publicados) dentro de un determinado período temporal. La especificación de una fecha inicial para la búsqueda puede estar justificada, bien porque el problema objeto de investigación se originó en dicha fecha, bien porque los trabajos anteriores a la misma están ya obsoletos.
- Los trabajos deben ser empíricos, es decir, deben presentar los resultados de una investigación original en la que se han tomado muestras de datos, se han registrado variables, se han asignado sujetos a distintos niveles de tratamiento, etc.
- Los documentos y publicaciones de los trabajos deben incluir los datos e información estadística suficientes para que el metaanalista pueda calcular los *tamaños del efecto*.

Por muy completa y exhaustiva que sea una búsqueda bibliográfica, nunca podrá agotar todos los trabajos realizados sobre el tema. El metaanalista debe ser consciente de los sesgos y limitaciones que puede tener la base bibliográfica a la que ha tenido acceso, así como las dificultades de generalización de sus resultados. La principal amenaza contra la validez de los resultados de un metaanálisis es el problema del *sesgo de publicación*, según el cual es más probable que los estudios que presentan resultados estadísticamente significativos sean aceptados para su publicación que los estudios con resultados no significativos. Si un metaanálisis dispone sólo de trabajos publicados, entonces sus resultados pueden estar sesgados en el sentido de una sobreestimación del *efecto de interés* (Rosenthal, 1991). Algunas propuestas, como las de este autor y Orwin (1983), permiten determinar si el sesgo de publicación puede ser una seria amenaza para las conclusiones de un metaanálisis. Pero, en cualquier caso, el metaanalista debe explicitar las técnicas de búsqueda de la literatura empleadas y, en la medida en que sus medios se lo permitan, utilizar el mayor número de éstas.

2.3. Codificación de las variables

Una vez que se dispone de los estudios seleccionados para el metaanálisis, el paso siguiente consiste en definir las características de los estudios que supuestamente pueden afectar a los resultados y, en consecuencia, explicar la variabilidad interestudios. En un metaanálisis dichas características actúan como variables potencialmente moderadoras y deben ser identificadas siguiendo criterios eminentemente teóricos. La mayor parte de las variables moderadoras que pueden definirse en un metaanálisis serán específicas del mismo, mientras que algunas de ellas son comunes a todos. Se pueden clasificar las variables moderadoras en sustantivas, metodológicas y extrínsecas.

- Las *sustantivas* hacen referencia al fenómeno objeto de investigación, incluyendo la forma en que se han operativizado las variables, las características de los sujetos participantes, el contexto cultural, geográfico e histórico, etc. Por ejemplo, el metaanálisis realizado por Damanpour en 1991 sobre los determinantes de la innovación organizacional, incluyó como variables moderadoras el tipo de organización (servicios, industrial), el tipo de innovación (administrativa, técnica, radical, incremental), la etapa de adopción (iniciación, implementación) y el alcance de la innovación (bajo, alto).
- Por *metodológicas* se entiende aquellos aspectos del método y del diseño de la investigación, con independencia del fenómeno investigado. Estas características suelen ser comunes a todos los metaanálisis. Cabe mencionar características tales como el tamaño muestral, el tipo de diseño (intergrupo frente a intragrupo), la calidad del estudio o la mortalidad experimental, entre otras.
- Por último, las *extrínsecas* representan factores externos a las características sustantivas y metodológicas, tales como la fecha de publicación, la fuente de publicación (trabajos publicados frente a no publicados), el género y la disciplina del autor. En principio, estas características no tendrían por qué influir en los resultados; sin embargo, en numerosos metaanálisis se han observado efectos moderadores de algunas de ellas.

En esta fase del metaanálisis también es preciso definir el índice del tamaño del efecto que se va a utilizar para representar los resultados de cada estudio empírico en el metaanálisis. Dependiendo del diseño de los estudios que se pretenden integrar, los índices del tamaño del efecto más utilizados son la diferencia media tipificada, d , para los diseños de grupo, y el coeficiente de correlación de Pearson, r , para los diseños de correlación (Rosenthal, 1991). Es frecuente que los estudios primarios no especifiquen el tamaño del efecto, por lo que el metaanalista tendrá que calcularlo a partir de los datos estadísticos disponibles. También es muy habitual que los estudios primarios incorporen múltiples variables dependientes en sus resultados, lo que conllevará la obtención de una estimación media del tamaño del efecto global del estudio (Marín-Martínez y Sánchez-Meca, 1999).

Tanto las características moderadoras de los estudios como el cálculo de los tamaños del efecto se registran en un *protocolo*, o cuestionario de codificación, cuyas normas para su correcta cumplimentación se deben recoger en un *manual de codificación*. En este manual se detallan de la forma más clara, precisa y objetiva los criterios y escalas de valoración de las variables objeto de codificación. Dicho manual es imprescindible, ya que la codificación de buena parte de las características de los estudios puede requerir cierto grado de inferencia por falta de información en los estudios.

La principal amenaza de esta fase del metaanálisis es la falta de fiabilidad en la codificación de las características de los estudios y en los cálculos del tamaño del efecto (Cooper, 1989). Es por ello que se hace imprescindible en todo metaanálisis un estudio de la fiabilidad del proceso de codificación, que suele consistir en la codificación independiente por dos (o más) codificadores, bien de todos los estudios empíricos, bien de una muestra aleatoria de los mismos. La puesta en común de las codificaciones efectuadas permite a continuación depurar el manual de codificación, siendo el objetivo último de este proceso asegurar que cualquier otro investigador que decidiera aplicar el manual a los mismos (u otros) estudios alcanzará las mismas valoraciones.

2.4. Análisis estadístico e interpretación

A partir de los objetivos e hipótesis formuladas en la primera fase del metaanálisis y de los tamaños del efecto y variables moderadoras registrados en cada estudio primario, se aplican técnicas de análisis estadístico específicamente diseñadas para tratar los datos metaanalíticos. Éstos serían los principales aspectos que aborda el análisis estadístico de un metaanálisis:

- El estudio del tamaño del efecto medio: representaciones gráficas de la distribución de los resultados, la tendencia central de los efectos (mediante alguna media ponderada), su variabilidad y su significación estadística (Glass, McGaw y Smith, 1981; Sánchez-Meca y Marín-Martínez, 1997, 1998b).
- La búsqueda de características moderadoras de la variabilidad de los tamaños del efecto, mediante la aplicación de técnicas de análisis de varianza (para variables moderadoras cualitativas), análisis de regresión simple (para variables moderadoras cuantitativas) y análisis de regresión múltiple, para la formulación de modelos explicativos del tamaño del efecto a partir de las variables moderadoras más relevantes (Hedges y Olkin, 1985; Marín-Martínez y Sánchez-Meca, 1998; Sánchez-Meca y Marín-Martínez, 1998a).

El análisis e interpretación de los datos metaanalíticos mediante técnicas estadísticas constituye una característica clave de esta metodología, y a su correcta comprensión se dedican los epígrafes 3 y 4.

2.5. Publicación del estudio

La difusión científica de un metaanálisis precisa de las mismas normas que se exigen a los estudios primarios. El esquema básico que debe guiar la redacción de un metaanálisis es, pues: 1) introducción; 2) método; 3) resultados, y 4) discusión (Cooper, 1989; Rosenthal, 1991).

1. La *introducción* de un metaanálisis, en la misma línea que la introducción de cualquier artículo de investigación, debe justificar el interés del trabajo y sus aportaciones novedosas frente a otros estudios.
2. En la sección *método* se marcan las más pronunciadas diferencias entre un estudio primario y un metaanálisis. En este último caso no se trata de especificar las condiciones experimentales bajo las que se plantea una investigación, sino más bien de precisar las múltiples decisiones que conlleva un proceso de revisión. Así, deben especificarse las técnicas de búsqueda de los estudios, los criterios de inclusión de los estudios, la medida de los resultados, los criterios de codificación de las características de los estudios y los análisis estadísticos seleccionados para el tratamiento de los datos.
3. La sección *resultados* incluye el análisis de la distribución de los tamaños del efecto, aportando su tendencia central, variabilidad y asociación con las características de los estudios, tanto desde una perspectiva descriptiva como inferencial. Las partes de esta sección pueden estructurarse en los siguientes apartados: *a)* estudio de la fiabilidad de la codificación; *b)* análisis del sesgo de publicación; *c)* análisis descriptivo de las características de los estudios; *d)* análisis descriptivo de la distribución de los tamaños del efecto; *e)* búsqueda de variables moderadoras influyentes; y, si procede, *f)* formulación de un modelo explicativo que incluya los moderadores que mejor expliquen la variabilidad de los tamaños del efecto.
4. En la sección *discusión* se presentan los resultados más relevantes obtenidos en el metaanálisis, y se conectan con los de otros estudios empíricos, revisiones teóricas y metaanálisis anteriores relacionados con el propio. Las conclusiones que se alcancen deben ponerse en relación con las teorías del área de investigación en que se ubica el metaanálisis, planteando su confirmación, reformulación o la necesidad de nuevos enfoques teóricos.

Todas las fases expuestas para la realización de un metaanálisis son fundamentales para la obtención del producto final, hasta el punto de que un sesgo en cualquiera de ellas afectará a la validez y utilidad de las restantes.

3. LA MEDIDA DE LOS RESULTADOS

Para poder integrar los resultados de múltiples estudios o investigaciones primarias acerca de una misma temática, es preciso que vengan expresados en una métrica común que los haga directamente comparables. En ciencias sociales es muy frecuente

que, para la medición de un mismo concepto, se empleen muy diversas definiciones operativas, con diferentes indicadores y escalas de medida. Desde el metaanálisis se ha propuesto el tamaño del efecto como el indicador idóneo para representar el resultado de una investigación¹. Son múltiples los índices del tamaño del efecto que se han derivado, dependiendo del diseño de la investigación, aunque es posible transformarlos entre ellos (Glass, McGaw y Smith, 1981; Hedges y Olkin, 1985). Por regla general, los índices del tamaño del efecto que se aplican en metaanálisis pretenden responder a preguntas específicas como, por ejemplo, cuál es la fuerza de la relación entre una variable dependiente y una variable independiente. En metaanálisis son dos las «familias» de tamaños del efecto que se suelen utilizar con mayor frecuencia, la familia *d* y la familia *r*, para los diseños de grupo y los diseños correlacionales, respectivamente.²

3.1. La familia *d*

Cuando el diseño del estudio implica la asignación de sujetos a grupos o niveles de tratamiento, ya sea aleatoria (diseños experimentales) o no aleatoria (diseños cuasi-experimentales), los índices del tamaño del efecto que se suelen aplicar se corresponden con la «familia *d*», mediante los cuales se compara el nivel medio estandarizado de dos grupos (Hedges y Olkin, 1985; Rosenthal, 1991). El índice de la «familia *d*» más comúnmente utilizado es la *diferencia media tipificada*, que consiste en calcular la diferencia entre las medias de los dos grupos y dividirla por la desviación típica intragrupo:

$$g = \frac{M_E - M_C}{S} \quad (1)$$

siendo *g* la diferencia media tipificada; *M_E* y *M_C*, las medias de los grupos experimental y control; y *S*, la desviación típica intragrupo que se obtiene mediante (Hedges y Olkin, 1985):

$$S = \sqrt{\frac{(n_E - 1)S_E^2 + (n_C - 1)S_C^2}{n_E + n_C - 2}} \quad (2)$$

¹ Aunque la corriente metaanalítica preconizada por Rosenthal proponía como resultado de los estudios el nivel de probabilidad, *p*, obtenido en la prueba estadística aplicada, actualmente apenas se utiliza en metaanálisis. La razón de su desuso está en que el nivel *p* representa una información parcial del resultado de un estudio, ya que da cuenta de la existencia o no de una relación estadísticamente significativa entre variables, pero sin especificar su magnitud. Por ejemplo, un nivel *p* de 0,001, que usualmente se interpreta en los términos de una alta significación estadística para el efecto estudiado, no se debe confundir con el hallazgo de un fenómeno de alta magnitud. De hecho, a partir de un tamaño muestral suficientemente elevado, podemos detectar un efecto incluso de muy escasa magnitud con el mismo nivel *p*. Ello se debe a que el resultado de una prueba de significación estadística, el nivel *p*, está en función del tamaño muestral y de la magnitud del efecto estudiado, según la relación: *Prueba de Significación = Tamaño del Efecto × Tamaño Muestral*.

siendo *n_E* y *n_C* los tamaños muestrales de los dos grupos, y *S_E²* y *S_C²*, las varianzas insesgadas². Este índice presenta un ligero sesgo positivo como estimador de la diferencia media tipificada en la población, δ . Un estimador insesgado de δ viene dado por (Hedges y Olkin, 1985):

$$d = c(m)g \quad (3)$$

donde *c(m)* es el factor de corrección del sesgo, que se define como: $c(m) = 1 - 3/(4m - 1)$, siendo $m = n_E + n_C - 2$. Por ejemplo, sean dos grupos —experimental y control— con medias $M_E = 8,933$ y $M_C = 7,200$, varianzas $S_E^2 = 5,781$ y $S_C^2 = 3,733$ y tamaños muestrales $n_E = 15$ y $n_C = 10$. Aplicando la ecuación (2) obtenemos como desviación típica $S = 2,231$; mediante la ecuación (1) calculamos la diferencia media tipificada $g = 0,777$; y ésta se corrige aplicando la ecuación (3), alcanzando el valor $d = 0,751$.

La diferencia media tipificada puede interpretarse como una puntuación típica que establece la distancia entre el centro de las distribuciones de dos poblaciones de sujetos en una determinada variable dependiente. Suponiendo distribuciones normales, es posible traducir el valor de *d* a percentiles de la distribución normal tipificada para facilitar su interpretación. Así, para un valor $d = 0,751$, el área de probabilidad acumulada bajo dicha distribución es de 0,7734, lo que significa que el sujeto medio del grupo experimental se sitúa en el percentil 77,34 por 100 de la distribución del grupo de control (Glass, McGaw y Smith, 1981; Sánchez-Meca y Ato, 1989).

En 1988, Cohen propuso valores normativos para distintos índices del tamaño del efecto que pueden ayudar a interpretar la magnitud o fuerza de la relación manifestada por un determinado valor. Así, para la diferencia media tipificada propone como tamaños del efecto bajos, medios y altos aquellos que se sitúen en torno a 0,20, 0,50 y 0,80, respectivamente. En nuestro ejemplo, pues, un valor $d = 0,751$ puede interpretarse orientativamente como un tamaño del efecto de magnitud alta, ya que se encuentra cerca del valor 0,80.

No es habitual que un artículo de investigación incluya las puntuaciones individuales de cada sujeto, por lo que la obtención de la diferencia media tipificada suele requerir cierto esfuerzo de cálculo adicional. Si el artículo aporta directamente las medias y desviaciones típicas de los dos grupos, el cálculo de *d* es directo mediante las ecuaciones (1), (2) y (3). Sin embargo, es muy frecuente que los artículos de investigación especifiquen sólo parte de los datos estadísticos, lo cual afectará al grado de precisión en el cálculo del tamaño del efecto. Veremos a continuación los casos más frecuentes.

² Glass, McGaw y Smith propusieron dividir no por *S*, sino por la desviación típica del grupo de control, *S_C*, por el hecho de que, en ocasiones, la aplicación de un tratamiento puede alterar la variabilidad natural de los sujetos en la variable dependiente, como resultado de un efecto interactivo sujeto-tratamiento. Pero siempre que se cumpla el supuesto de homogeneidad de las varianzas entre los dos grupos, el índice más eficiente es el propuesto por Hedges, *g*, y es éste el de uso más frecuente en metaanálisis.

Un caso muy común consiste en presentar únicamente el resultado de la prueba de significación aplicada. En un diseño de dos grupos, la prueba estadística habitual es la T de significación de diferencias entre medias independientes. A partir del resultado de la prueba T , y conociendo los grados de libertad, puede obtenerse la diferencia media tipificada, g , mediante:

$$g = T \sqrt{\frac{1}{n_E} + \frac{1}{n_C}} \quad (4)$$

Así, en nuestro ejemplo, suponiendo que se ha aplicado la prueba T , se hubiera obtenido: $T(23) = 1,903$, $p(\text{bilateral}) = 0,070$. A partir de la ecuación (4), la obtención de g es directa:

$$g = (1,903) \sqrt{(1/15) + (1/10)} = 0,777$$

valor que coincide con el anteriormente calculado a partir de la ecuación (1).

En ocasiones, el artículo especifica únicamente el nivel p de probabilidad (generalmente bilateral) alcanzado por la prueba estadística aplicada sobre los datos. En este caso, la obtención de g pasa necesariamente por la consulta de las tablas de la distribución del estadístico de contraste (o el uso de un programa estadístico), mediante la cual se determina el valor del estadístico asociado a dicho nivel de probabilidad. Una vez conseguido el valor del estadístico, se obtiene g a partir de la ecuación (4).

Siguiendo con nuestro ejemplo, supóngase que la información que se conoce es que se ha aplicado la prueba T de diferencias entre medias y que $p(\text{bilateral}) = 0,070$. Sabiendo que los grados de libertad de dicha prueba son $GL = n_E + n_C - 2 = 15 + 10 - 2 = 23$, consultamos en las tablas de la distribución t de Student con $GL = 23$ y el percentil $1 - \alpha/2 = 1 - 0,07/2 = 0,965$, obteniendo: $t_{1 - \alpha/2, GL} = 0,965, 23 = 1,900$. A continuación, mediante la ecuación (4) obtenemos:

$$g = (1,900) \sqrt{(1/15) + (1/10)} = 0,776$$

un valor muy similar al obtenido anteriormente.

Otra situación también frecuente es aquella en la que sólo se especifica en el artículo el nivel de significación al que ha resultado significativa la prueba estadística. Nos estamos refiriendo a expresiones tales como $p < 0,05$. En este caso, una estimación conservadora del tamaño del efecto consiste en identificar el valor de probabilidad exacto alcanzado por la prueba (que es desconocido) con el nivel de significación especificado. Es decir, hacemos $p = 0,05$ y seguimos los pasos anteriormente desarrollados para calcular el valor aproximado de g . Si, por el contrario, la prueba estadística aplicada no ha resultado significativa, el artículo lo reflejará indicando, por ejemplo, simplemente que $p > 0,05$; en este caso, la única opción que tenemos para no desechar esta información es adoptar una estrategia conservadora, que consiste en hacer $p = .50$ y, en consecuencia, fijar el valor de $g = 0$.

Resumiendo, la precisión de los cálculos para la obtención de la diferencia media tipificada está en función de la información disponible en los artículos de investigación. Por tanto, según los casos, podremos alcanzar el valor exacto o bien sólo una aproximación más o menos conservadora. Otras derivaciones de la diferencia media tipificada para diferentes diseños experimentales y cuasi-experimentales pueden consultarse en la obra de Glass, McGaw y Smith y en Sánchez-Meca y Ato (1989).

3.2. La familia r

Cuando los estudios que se pretenden metaanalizar aplican diseños de tipo correlacional, el índice del tamaño del efecto habitualmente utilizado es el coeficiente de correlación de Pearson, r , o bien cualquiera de sus aplicaciones, como, por ejemplo, el coeficiente ϕ , ϕ (para dos variables dicotómicas), la correlación biserial-puntual, r_{bp} (para una variable dicotómica y una variable continua), la correlación de Spearman, r_s (para dos variables ordinales). Existen fórmulas de conversión de cualquiera de estos estadísticos a coeficiente de correlación de Pearson (Glass, McGaw y Smith, 1981; Rosenthal, 1991; Sánchez-Meca y Ato, 1989). Cohen (1988) propuso como valores de magnitud baja, media y alta para el coeficiente de correlación de Pearson: 0,10, 0,30 y 0,50, respectivamente.

El coeficiente de correlación de Pearson se considera el índice universal para medir la magnitud del efecto, ya que puede aplicarse no sólo a los estudios de tipo correlacional, sino también a los diseños de grupos (Rosenthal, 1991). De hecho, se han desarrollado fórmulas de transformación de la diferencia media tipificada a coeficiente de correlación, y viceversa. Estas fórmulas son muy útiles para homogeneizar los resultados de los estudios en un único índice del tamaño del efecto. Así, la transformación de la diferencia media tipificada, d , a coeficiente de correlación de Pearson, r , se obtiene mediante:

$$r = \frac{d}{\sqrt{d^2 + 4}} \quad (5)$$

Así, en nuestro ejemplo aplicamos la ecuación (5) sobre el valor $d = 0,751$ y obtenemos como coeficiente de correlación $r = 0,351$. También es posible aplicar la transformación a la inversa para obtener el valor de d a partir de r .

Aunque la diferencia media tipificada y el coeficiente de correlación de Pearson son los índices del tamaño del efecto más comunes en los metaanálisis dentro del ámbito de la investigación en Economía de la Empresa, con menor frecuencia se utilizan otros índices del tamaño del efecto, como son las pendientes y las intercepciones de modelos de regresión simple entre las variables implicadas. Por ejemplo, en su metaanálisis sobre la relación entre rentabilidad y acciones de mercado, Szymanski, Bharadwaj y Varadarajan (1993) utilizaron como índice del tamaño del efecto los coeficientes de regresión simples (*elasticities*) registrados en los estudios primarios. Este tipo de estimadores puede consultarse en Farley y Lehmann (1986).

4. EL METAANÁLISIS PSICOMÉTRICO

De todas las fases que conlleva la realización de un metaanálisis, la que le imprime mayor carácter es, sin duda, la de integración cuantitativa de los resultados de los estudios, mediante la aplicación de los métodos estadísticos desarrollados al efecto. Pero distintos enfoques metaanalíticos proponen técnicas de análisis estadístico diferentes. De los diversos enfoques metaanalíticos que actualmente existen, nos centramos en el que se ha aplicado con mayor frecuencia en el ámbito de la investigación en Economía de la Empresa: el denominado metaanálisis psicométrico propuesto por Hunter y Schmidt.

El origen de esta propuesta metaanalítica se sitúa en los trabajos de Schmidt y Hunter sobre la generalización de la validez de criterio de los tests de selección de personal. Frente a la hipótesis de la generalización de la validez, en aquella época (1977) dominaba la hipótesis de la «especificidad situacional», según la cual la validez de un test se ve notablemente afectada por la situación específica de aplicación: el entorno laboral, el tipo de trabajo o las características de los trabajadores, entre otras. Así pues, desde este enfoque teórico carecía de sentido la integración interestudios de unos coeficientes de validez necesariamente heterogéneos. Pero el enfoque metaanalítico desarrollado por Hunter y Schmidt permitió comprobar la veracidad de dicha teoría, mediante la formulación de un modelo estadístico que posibilita determinar si la variabilidad observada en los coeficientes de validez de un determinado test se debe a que dichos coeficientes difieren entre sí en función del contexto de aplicación o si, por el contrario, la variabilidad observada se puede explicar exclusivamente por el influjo de artefactos estadísticos, tales como el error de muestreo, la restricción del rango, el error de medida en las variables o la falta de validez en las variables.

Es por ello que a este enfoque metaanalítico también se le conoce como el enfoque de la «generalización de la validez» («*validity generalization*»). Aunque Hunter y Schmidt desarrollaron su enfoque en el contexto de la integración de coeficientes de validez de los tests para la selección de personal, su posterior aplicación a otros ámbitos de la investigación en las ciencias sociales y, en especial, en Psicología del Trabajo y de las Organizaciones, es lo que conllevó su más reciente denominación como «metaanálisis psicométrico». En consecuencia, la presentación de su enfoque no se va a limitar exclusivamente a la integración de coeficientes de validez, sino de coeficientes de correlación de Pearson entre dos variables cualesquiera³.

El propósito del metaanálisis es, en primer lugar, obtener una estimación de la magnitud del efecto medio de un conjunto de estudios sobre un determinado problema de investigación; es decir, una correlación media interestudios. En segundo lugar, analizar la variabilidad observada entre los coeficientes de correlación para

³ Aunque nuestra exposición del metaanálisis psicométrico se va a centrar en el coeficiente de correlación de Pearson como índice del tamaño del efecto, Hunter y Schmidt (1990) han desarrollado las fórmulas para aplicarlo sobre la diferencia media tipificada.

valorar si dicha variabilidad se debe a la intervención de variables moderadoras sobre los resultados de los estudios. Finalmente, si cabe sospechar la presencia de variables moderadoras, intentar detectar cuáles son las más relevantes.

Hunter y Schmidt sostienen que los coeficientes de correlación que se obtienen en los estudios individuales están sometidos al influjo de diversos artefactos estadísticos cuyo efecto es, generalmente, reducir su valor con relación al verdadero coeficiente de correlación en la población, ρ , produciéndose una infraestimación de dicho parámetro. Los coeficientes de correlación empíricos deben, pues, corregirse para obtener una estimación más precisa de la correlación poblacional. Estos artefactos estadísticos también afectan a la variabilidad observada entre los coeficientes de correlación individuales, provocando un incremento «artificial» de la misma. Es por ello que, antes de pasar a la búsqueda de variables moderadoras que expliquen tal variabilidad, se hace preciso comprobar si la variabilidad observada entre las correlaciones se debe al influjo de los artefactos estadísticos.

Hasta la fecha, se han descrito once artefactos estadísticos que pueden afectar a la magnitud y a la variabilidad de los coeficientes de correlación empíricos, los cuales se presentan en el cuadro 7.1. Todos los artefactos, a excepción del décimo —errores de reporte—, pueden corregirse sobre las correlaciones empíricas. De los diez que pueden corregirse, todos, a excepción del primero —el error de muestreo—, ejercen un influjo sistemático sobre las correlaciones empíricas y, en consecuencia, puede corregirse su efecto individualmente sobre cada correlación. Para ello, cada estudio empírico debe aportar los datos necesarios para poder efectuar la corrección. El error de muestreo, por el contrario, ejerce un influjo no sistemático y su corrección sólo puede hacerse sobre el conjunto de las correlaciones empíricas. Una vez corregidas las correlaciones empíricas, es posible obtener un promedio más exacto de cuál debe ser el coeficiente de correlación en la población sin las imperfecciones de los estudios empíricos.

Se presenta a continuación el procedimiento metaanalítico para la corrección de los cinco artefactos más importantes: el error de muestreo, el error de medida en las variables dependiente e independiente y la variación en el rango de las variables dependiente e independiente⁴.

4.1. Metaanálisis con la corrección del error de muestreo

Hunter y Schmidt se basan en la teoría clásica de los tests para establecer que la correlación empírica que se obtiene en un estudio individual, r_i , es igual a la correlación verdadera en la población, ρ , más el error de muestreo de la estimación, e_i , según la expresión: $r_i = \rho + e_i$.

⁴ Aunque utilizaremos los términos «variable dependiente» y «variable independiente», estos procedimientos metaanalíticos son aplicables indistintamente al caso de dos variables sin que, necesariamente, una de ellas tenga que estar supeditada a la otra.

CUADRO 7.1

Artefactos estadísticos que afectan al valor de las correlaciones empíricas

Artefacto estadístico	Descripción
1. Error de muestreo.	Las correlaciones empíricas varían entre sí debido al error de muestreo.
2. Error de medida en la variable dependiente.	La correlación empírica verá reducido su valor cuanto menor fiabilidad tenga la medida de la variable dependiente.
3. Error de medida en la variable independiente.	La correlación empírica verá reducido su valor cuanto menor fiabilidad tenga la medida de la variable independiente.
4. Dicotomización de la variable dependiente.	La correlación empírica verá reducido su valor cuando la variable dependiente es dicotomizada.
5. Dicotomización de la variable independiente.	La correlación empírica verá reducido su valor cuando la variable independiente es dicotomizada.
6. Variación del rango en la variable dependiente.	La correlación empírica verá reducido su valor si la desviación típica de la variable dependiente es inferior a la de la población.
7. Variación del rango en la variable independiente.	La correlación empírica verá reducido su valor si la desviación típica de la variable independiente es inferior a la de la población.
8. Validez de concepto imperfecta en la variable dependiente.	La correlación empírica verá reducido su valor cuanto más imperfecta sea la validez de concepto de la variable dependiente.
9. Validez de concepto imperfecta en la variable independiente.	La correlación empírica verá reducido su valor cuanto más imperfecta sea la validez de concepto de la variable independiente.
10. Errores de reporte.	Las correlaciones empíricas pueden diferir de sus correlaciones poblacionales debido a errores tipográficos, de transcripción, de codificación de datos, errores de computación, etc.
11. Injuicio de factores extraños.	La varianza de las correlaciones empíricas puede incrementarse por la ausencia de control de factores extraños, tales como la experiencia de los participantes, el tiempo de permanencia en el trabajo, etc.

El error de muestreo es, pues, el fenómeno por el que se explica el hecho de que diferentes correlaciones empíricas obtenidas en distintos estudios alcancen distintos valores, a pesar de que todas las muestras procedan de una misma población y esti-

men la misma correlación poblacional, ρ . Dado que la correlación poblacional, ρ , y el error de muestreo, e , son independientes entre sí, la varianza total observada entre las correlaciones empíricas, σ_e^2 , puede descomponerse en dos fuentes de varianza: la varianza verdadera entre las correlaciones poblacionales, σ_ρ^2 , y la varianza debida al error de muestreo, σ_e^2 , según la expresión: $\sigma_e^2 = \sigma_\rho^2 + \sigma_e^2$. Si todas las correlaciones empíricas proceden de una misma población con parámetro ρ , entonces la varianza de las correlaciones poblacionales será nula, $\sigma_\rho^2 = 0$, y, en consecuencia, toda la varianza observada será varianza debida al error de muestreo: $\sigma_e^2 = \sigma_e^2$. En este caso, el metaanalista debería concluir que la variabilidad observada entre las correlaciones empíricas de los estudios integrados es tan sólo aparente, o «artificial», ya que se debe al fenómeno del error de muestreo, no siendo necesaria la búsqueda de variables moderadoras de los estudios. Si, por el contrario, realmente las correlaciones empíricas proceden de poblaciones con correlaciones paramétricas diferentes, entonces la varianza observada, σ_e^2 , no contendrá sólo varianza debida al error de muestreo, σ_e^2 , sino que también incluirá varianza verdadera entre las correlaciones paramétricas (es decir, $\sigma_\rho^2 > 0$). En este caso, se hace preciso continuar los análisis para encontrar aquellas variables que están moderando los resultados de los estudios.

Hunter y Schmidt proponen estimar la varianza total observada en las correlaciones empíricas, σ_e^2 , calculando la varianza de los valores r_i de los estudios mediante:

$$S_r^2 = \frac{\sum N_i (r_i - \bar{r})^2}{\sum N_i} \quad (6)$$

donde N_i es el tamaño muestral de *i*ésimo estudio, r_i es la correlación empírica del *i*ésimo estudio y \bar{r} es una media ponderada de las correlaciones empíricas, que se obtiene mediante:

$$\bar{r} = \frac{\sum N_i r_i}{\sum N_i} \quad (7)$$

La varianza debida a error de muestreo, σ_e^2 , se estima mediante:

$$S_e^2 = \frac{(1 - \bar{r}^2)^2}{\bar{N} - 1} \quad (8)$$

donde \bar{N} es el tamaño muestral medio de los estudios: $\bar{N} = \sum N_i/k$, siendo k el número de estudios. Y la varianza real de las correlaciones poblacionales, σ_ρ^2 , se estima sustrayendo la varianza debida a error de muestreo, S_e^2 , de la varianza observada, S_r^2 :

$$S_\rho^2 = S_r^2 - S_e^2 \quad (9)$$

Una vez estimadas la varianza observada y la varianza debida al error de muestreo, debe comprobarse si las correlaciones empíricas son homogéneas entre sí, es decir, si la varianza observada se debe fundamentalmente al artefacto estadístico de la varianza de error o si, por el contrario, parte de la varianza observada se debe al influjo de variables moderadoras o, lo que es lo mismo, a que existe varianza verdadera entre las correlaciones empíricas. Hunter y Schmidt han propuesto dos pruebas estadísticas que permiten responder a esta cuestión. Una de ellas consiste en aplicar lo que ellos han denominado la «regla del 75%», según la cual si al menos el 75 por 100 de la varianza observada está formada por varianza de error de muestreo, entonces puede desecharse la hipótesis de que exista varianza verdadera entre las correlaciones empíricas y, en consecuencia, concluir que las correlaciones de los estudios son homogéneas. Si, por el contrario, la varianza debida al error de muestreo no llega a explicar dicho 75 por 100, entonces debe suponerse que existen variables moderadoras que están afectando a la magnitud de las correlaciones empíricas y, en ese caso, no se sostiene la hipótesis de la homogeneidad. Así pues, la regla de 75 por 100 consiste en calcular el valor:

$$\text{Regla-75\%} = (100)(S_o^2)/(S_e^2) \quad (10)$$

Otro procedimiento propuesto por Hunter y Schmidt para comprobar la hipótesis de homogeneidad de los coeficientes de correlación consiste en aplicar el estadístico Q mediante:

$$Q = (k)(S_o^2)/(S_e^2) \quad (11)$$

siendo k el número de estudios. Bajo la hipótesis de homogeneidad, el estadístico Q se distribuye según una ley χ^2 de Pearson con $k - 1$ grados de libertad. Así, siendo α el nivel de significación adoptado en el contraste, si el valor obtenido con la ecuación (11) supera al percentil $100(1 - \alpha)$ de dicha distribución, la hipótesis de homogeneidad no es sostenible y, en consecuencia, deberá procederse a la búsqueda de variables moderadoras que expliquen la heterogeneidad observada⁵.

El resultado de la prueba de homogeneidad condiciona, pues, los pasos a seguir a continuación. Si los coeficientes de correlación empíricos son homogéneos entre sí, entonces la correlación media es representativa de todos ellos y puede estimarse la correlación poblacional construyendo un intervalo de confianza mediante:

$$\bar{r} \pm |z_{\alpha/2}| S_r \quad (12)$$

⁵ Estudios de simulación han puesto de manifiesto los problemas de la regla del 75% para mantener bien ajustada la tasa de error tipo I, así como la escasa potencia estadística que ofrecen tanto la prueba Q como la regla del 75% con un número pequeño de estudios (e.g., Sagie y Koslowsky, 1993; Sánchez-Meca y Marín-Martínez, 1997). Su aplicación debe, pues, condicionarse a una base metaanalítica grande.

donde $z_{\alpha/2}$ es la puntuación de la distribución normal tipificada correspondiente al percentil $100(\alpha/2)$ y S_r es el error típico de la correlación media, que se estima por:

$$S_r = \frac{1 - \bar{r}^2}{\sqrt{N - k}} \quad (13)$$

siendo N el tamaño muestral total: $N = \sum N_i$.

Si la prueba de homogeneidad resulta significativa, el paso siguiente consistirá en la búsqueda de variables moderadoras, para lo cual los estudios se agrupan en función de alguna característica diferencial de los mismos y se repiten los análisis anteriores con cada subconjunto, es decir, las ecuaciones (6)-(13). Si las correlaciones medias de los subgrupos difieren claramente entre sí, a la vez que las varianzas dentro de los subgrupos son inferiores a la varianza del metaanálisis global, esto será evidencia de que la variable moderadora influye realmente en los resultados⁶.

4.2. Metaanálisis con la corrección de otros artefactos

Por regla general, los artefactos estadísticos que afectan al coeficiente de correlación provocan una atenuación de su verdadero valor en la población. Ése el caso, por ejemplo, del error de medida, tanto en la variable independiente como en la dependiente, así como el fenómeno de la restricción del rango en ambas variables. Por tanto, la correlación media obtenida en un metaanálisis tenderá a ser inferior respecto del verdadero coeficiente de correlación en la población asumiendo ausencia de artefactos estadísticos. La varianza de las correlaciones empíricas (o atenuadas) también se deja afectar por los artefactos estadísticos incrementándose «artificialmente». En esta sección se presentan los procedimientos metaanalíticos para corregir el efecto de la atenuación provocado por la falta de fiabilidad y por la variación del rango de las variables implicadas en el cálculo de las correlaciones empíricas.

Corrección del error de medida

La atenuación producida por el *error de medida* puede corregirse de forma individual sobre cada coeficiente de correlación empírico si se dispone de los coeficientes de fiabilidad de las variables dependiente e independiente. Siendo r_{xx} y r_{yy} los

⁶ Aunque otros enfoques aplican rutinariamente pruebas de significación estadística para detectar variables moderadoras, Hunter y Schmidt desaconsejan su uso (y también en la investigación primaria), especialmente cuando el metaanálisis en cuestión dispone de un número reducido de estudios (inferior a 30 o 40) y el tamaño muestral total no es muy elevado. Bajo estas condiciones, la baja potencia estadística de las pruebas estadísticas y la capitalización sobre el azar que supone probar variables moderadoras de forma «ciega» (sin una teoría subyacente) hacen muy cuestionable su utilidad. Sólo con un número alto de estudios estos autores consideran conveniente la aplicación de pruebas estadísticas de significación que, por razones de espacio, no se presentan en este capítulo, pero que el lector interesado puede consultar la obra de Hunter y Schmidt.

coeficientes de fiabilidad de las variables independiente y dependiente, respectivamente, de un estudio empírico, es posible definir dos factores de corrección de la atenuación, a_1 y a_2 , uno para cada variable mediante:

$$a_1 = \sqrt{r_{xx}} \quad (14)$$

$$a_2 = \sqrt{r_{yy}} \quad (15)$$

De esta forma, la correlación corregida, r_c (o desatenuada), puede estimarse dividiendo la correlación empírica, r , por el producto de los factores de corrección, a_1 y a_2 :

$$r_c = r/a_1 a_2 \quad (16)$$

Por ejemplo, en un estudio se obtiene una correlación empírica $r = 0,34$, siendo los coeficientes de fiabilidad de las dos variables: $r_{xx} = 0,64$ y $r_{yy} = 0,72$. La correlación corregida, o desatenuada, se estima calculando, primero, los factores de corrección:

$$a_1 = \sqrt{r_{xx}} = \sqrt{0,64} = 0,80 \quad \text{y} \quad a_2 = \sqrt{r_{yy}} = \sqrt{0,72} = 0,85$$

y, segundo, aplicando la ecuación (16): $r_c = r/a_1 a_2 = 0,34/(0,80)(0,85) = 0,50$. Así, el error de medida ha supuesto en este caso una reducción del 32 por 100 sobre la verdadera correlación.

Corrección de la variación del rango

Coefficientes de correlación obtenidos en diferentes estudios con diferentes desviaciones típicas en la variable independiente no son directamente comparables, ya que la variabilidad de las puntuaciones afecta a la magnitud de la correlación. Este artefacto afecta tanto a la correlación media obtenida en un metaanálisis como a su variabilidad. Así, se hace preciso unificar la variabilidad de la variable independiente tomando una desviación típica como punto de referencia. Siendo s_x la desviación típica de un determinado estudio y S_x la desviación típica de referencia, puede definirse el factor de corrección de la variación en el rango de la variable independiente, a_3 , mediante:

$$a_3 = \sqrt{u_x^2 + r^2 - u_x^2 r^2} \quad (17)$$

donde u_x es la razón entre las dos desviaciones típicas, $u_x = s_x/S_x$, y r es la correlación empírica obtenida en el estudio. Así, la correlación corregida, r_c , se obtiene dividiendo la correlación empírica, r , por el factor de corrección, a_3 :

$$r_c = r/a_3 \quad (18)$$

Por ejemplo, sea un estudio empírico que obtiene una correlación $r = 0,34$, siendo la desviación típica de la variable independiente, $s_x = 10$. Suponiendo que la desviación típica de referencia es $S_x = 25$, tenemos que $u_x = s_x/S_x = 10/25 = 0,4$, lo que supone un factor de corrección:

$$a_3 = \sqrt{u_x^2 + r^2 - u_x^2 r^2} = \sqrt{(0,4)^2 + (0,34)^2 - (0,4)^2(0,34)^2} = 0,51$$

y una correlación corregida $r_c = r/a_3 = 0,34/0,51 = 0,67$. Es decir, en este caso el hecho de utilizar en el estudio una desviación típica sensiblemente inferior a la de referencia ha supuesto una reducción del 49 por 100, casi la mitad del valor corregido.

En el ámbito de la aplicación de los tests para la selección de personal es muy frecuente el fenómeno de la restricción del rango que, como se manifiesta en el ejemplo, produce una atenuación de la verdadera correlación entre dos variables por el hecho de utilizar una desviación típica inferior a la de la población. En selección de personal, la restricción del rango suele producirse porque la correlación empírica se calcula a partir de la proporción de solicitantes que han superado una determinada puntuación en el test y, como consecuencia de ello, han sido admitidos para el puesto de trabajo. Si se conoce dicha proporción, p , y asumiendo normalidad en la distribución de la variable independiente, es posible estimar la razón entre las desviaciones típicas, u_x , sin necesidad de conocerlas, mediante:

$$u_x = \sqrt{1 - \mu_x^2 + z_{1-p} \mu_x} \quad (19)$$

donde z_{1-p} representa a la puntuación de la distribución normal tipificada correspondiente al percentil $100(1-p)$ y μ_x se obtiene mediante:

$$\mu_x = y/p \quad (20)$$

siendo y la altura en la ordenada de la distribución normal tipificada para el percentil $100(p)$. A continuación, se procedería según las ecuaciones (17) y (18).

Por ejemplo, en un estudio de selección de personal la proporción de solicitantes que superaron la puntuación criterio en el test de selección para ser admitido en el trabajo ha sido $p = 0,20$. Sobre la muestra de solicitantes admitidos se obtiene una correlación empírica $r = 0,34$. Para calcular u_x primero se obtiene μ_x consultando las tablas de la distribución normal tipificada y aplicando la ecuación (20): $\mu_x = y/p = 0,2803/0,20 = 1,4015$. A continuación, mediante la ecuación (19), y sabiendo que $z_{1-p} = z_{0,80} = 0,84$, se obtiene el valor de u_x :

$$u_x = \sqrt{1 - \mu_x^2 + z_{1-p} \mu_x} = \sqrt{1 - (1,4015)^2 + (0,84)(1,4015)} = 0,46$$

Se obtiene el factor de corrección mediante la ecuación (17):

$$a_3 = \sqrt{u_x^2 + r^2 - u_x^2 r^2} = \sqrt{(0,46)^2 + (0,34)^2 - (0,46)^2(0,34)^2} = 0,55$$

Finalmente, la correlación corregida se obtiene por la ecuación (18): $r_c = r/a_3 = 0,34/0,55 = 0,62$. Es decir, en este caso una razón de selección del 20 por 100 ha supuesto una reducción de la verdadera correlación del 45 por 100.

La variable dependiente también puede presentar variación en el rango a través de los estudios de un metaanálisis, pudiéndose corregir mediante la aplicación de las ecuaciones (17) y (18) adaptadas al caso de esta variable. En concreto, siendo $u_y = s_y/S_y$ la razón entre las desviaciones típicas de la variable dependiente en el estudio, s_y , y de referencia, S_y , el factor de corrección de la variación en el rango de esta variable se define como:

$$a_4 = \sqrt{u_y^2 + r^2 - u_y^2 r^2} \quad (21)$$

siendo la correlación corregida el resultado de aplicar:

$$r_c = r/a_4 \quad (22)$$

Finalmente, si existe variación del rango en las dos variables simultáneamente, la correlación corregida será el resultado de dividir la correlación empírica por el producto de los dos factores de corrección de este artefacto:

$$r_c = r/a_3 a_4 \quad (23)$$

Metaanálisis de las correlaciones corregidas

En el epígrafe 4.1 se ha tratado el procedimiento metaanalítico que corrige sólo el efecto del error de muestreo. Si además del error de muestreo se pretende corregir el efecto atenuante que ejercen determinados artefactos estadísticos, tales como el error de medida o la restricción del rango, es preciso que cada estudio individual aporte los datos estadísticos necesarios para calcular los factores de corrección individuales. Éstos son los pasos a seguir para realizar un metaanálisis sobre las correlaciones corregidas.

1. La información de partida que se requiere de cada estudio es: a) el coeficiente de correlación empírico (o atenuado), r_i ; b) el tamaño muestral, N_i ; y c) el factor de corrección compuesto, A_i , que se define como el producto de los factores de corrección de los diversos artefactos cuyo efecto se pretende corregir. Si, por ejemplo, se pretende corregir el efecto del error de medida en las variables dependiente e independiente y el efecto de la variación del rango de ambas variables, entonces: $A_i = a_1 a_2 a_3 a_4$, donde a_1, a_2, a_3 y a_4 , son los factores de corrección de dichos artefactos para el *i*ésimo estudio, según se definieron en las ecuaciones (14), (15), (17) y (21). El factor de corrección compuesto se define, pues, en función de los artefactos que se quieren corregir de forma individual sobre las correlaciones empíricas.

2. A continuación, se calcula para cada estudio la correlación corregida (o desatenuada), r_{ci} , y el peso de ponderación de cada estudio, w_i , que se obtienen, respectivamente, mediante:

$$r_{ci} = r_i/A_i \quad (24)$$

$$w_i = N_i A_i^2 \quad (25)$$

3. Para estimar el error de muestreo corregido de cada estudio, S_{ei}^2 , es preciso calcular previamente el error de muestreo no corregido, S_{ei}^2 , mediante:

$$S_{ei}^2 = \frac{(1 - r_i^2)^2}{N_i - 1} \quad (26)$$

siendo r_* la media ponderada de las correlaciones no corregidas (o atenuadas):

$$r_* = \frac{\sum w_i r_i}{\sum w_i} \quad (27)$$

Así, el error de muestreo corregido se estima mediante:

$$S_{ei}^2 = S_{ei}^2/A_i^2 \quad (28)$$

En este punto, se dispone en cada estudio de los siguientes datos: a) la correlación empírica (o atenuada); r_i ; b) el tamaño muestral, N_i ; c) el factor de corrección compuesto, A_i ; d) la correlación corregida (o desatenuada), r_{ci} ; y e) la varianza del error de muestreo corregida, S_{ei}^2 .

4. El paso siguiente consiste en efectuar el metaanálisis, propiamente dicho, sobre las correlaciones corregidas, lo que conlleva los siguientes pasos:
 - a) Calcular la correlación media corregida, \bar{r}_c , y ponderada por los pesos w_i :

$$\bar{r}_c = \frac{\sum w_i r_{ci}}{\sum w_i} \quad (29)$$

- b) Calcular la varianza total de las correlaciones corregidas, S_{rc}^2 :

$$S_{rc}^2 = \frac{\sum w_i (r_{ci} - \bar{r}_c)^2}{\sum w_i} \quad (30)$$

c) Calcular la varianza corregida debida al error de muestreo, S_c^2 :

$$S_c^2 = \frac{\sum w_i S_{ei}^2}{\sum w_i} \quad (31)$$

d) Calcular la varianza verdadera entre las correlaciones corregidas sustrayendo de la varianza total la varianza debida al error de muestreo:

$$S_p^2 = S_r^2 - S_c^2 \quad (32)$$

e) Una vez que ya se ha estimado la correlación media corregida, \bar{r}_c , y la varianza verdadera de dichas correlaciones, S_p^2 , el siguiente paso consiste en comprobar la hipótesis de la homogeneidad de las correlaciones corregidas mediante la regla del 75 por 100, o bien mediante la aplicación de la prueba estadística Q , que se distribuye según una distribución χ^2 con $k - 1$ grados de libertad. Las ecuaciones (10) y (11) permiten obtener estos dos estadísticos, con la precaución de utilizar las varianzas corregidas, S_c^2 y $S_{c_i}^2$, en lugar de las no corregidas.

5. Finalmente, si la hipótesis de homogeneidad de las correlaciones no se sostiene, debe procederse a la búsqueda de variables moderadoras que expliquen la variabilidad que permanece una vez controlados los artefactos. Para ello, se repiten los análisis aquí expuestos para cada subgrupo o categoría de la variable moderadora en cuestión y se comparan las correlaciones medias corregidas y sus varianzas.

Para finalizar, debe quedar claro al lector que el procedimiento aquí expuesto para metaanalizar un conjunto de correlaciones corregidas está supeditado al hecho de que se disponga en cada estudio de la información necesaria para calcular los factores de corrección. Cuando en un metaanálisis un número considerable de estudios no contiene tal información existen procedimientos alternativos que permiten definir distribuciones de los artefactos estadísticos y poder así corregir las correlaciones empíricas. Estos procedimientos pueden consultarse en la obra de Hunter y Schmidt.

5. UNA APLICACIÓN PRÁCTICA

Como ejemplo de aplicación del metaanálisis psicométrico, se presenta a continuación el metaanálisis realizado por Sadri y Robertson (1993) que tenía como objetivo estimar la magnitud de la relación entre las expectativas de autoeficacia de las personas y su rendimiento en el trabajo. Aunque se trata de una relación bien establecida en la literatura no hay acuerdo, sin embargo, respecto de cuál es la magnitud de la relación. Sadri y Robertson realizaron una búsqueda bibliográfica exhaustiva

logrando identificar un total de 26 estudios empíricos que aportaban coeficientes de correlación entre medidas de ambos conceptos. Las seis primeras columnas del cuadro 7.2 presentan los datos de dichos estudios: tipo de medida del criterio (rendimiento versus conducta), tamaño muestral, N_i , coeficiente de correlación de Pearson, r_i , coeficiente de fiabilidad del predictor (la autoeficacia), r_{xx} , y coeficiente de fiabilidad del criterio, r_{yy} .

CUADRO 7.2

Datos del metaanálisis de Sadri y Robertson (1993)^a

Estudio	Criterio	N_i	r_i	r_{xx}	r_{yy}	A_i	S_{ei}^2	N_i	r_{xx}	r_{yy}	A_i	S_{ei}^2
1	1	120	0,76	0,81	0,92	0,863	0,880	89,424	0,0066	0,0088		
2	1	278	0,29	0,60	0,85	0,714	0,406	141,780	0,0028	0,0055		
3	1	105	0,32	0,89	0,87	0,880	0,364	81,301	0,0075	0,0097		
4	1	52	-0,03	0,93	0,90	0,915	-0,033	43,524	0,0153	0,0183		
5	1	110	0,16	0,85	0,93	0,889	0,180	86,955	0,0072	0,0091		
6	1	194	0,28	0,84	0,92	0,879	0,319	149,923	0,0041	0,0052		
7	1	212	0,22	0,84	0,92	0,879	0,250	163,834	0,0037	0,0048		
8	1	111	0,37	0,84	0,92	0,879	0,421	85,781	0,0071	0,0092		
9	1	181	0,54	0,83	0,86	0,845	0,639	129,198	0,0043	0,0061		
10	1	20	0,49	0,88	0,88	0,880	0,557	15,488	0,0412	0,0532		
11	1	83	0,38	0,83	0,85	0,840	0,452	58,557	0,0095	0,0135		
12	1	97	0,20	0,83	0,85	0,840	0,238	68,434	0,0082	0,0116		
13	1	24	0,76	0,83	0,85	0,840	0,905	16,932	0,0340	0,0482		
14	1	40	0,64	0,83	0,85	0,840	0,762	28,220	0,0201	0,0284		
15	1	67	0,48	0,83	0,85	0,840	0,571	47,269	0,0119	0,0168		
16	1	60	0,71	0,83	0,85	0,840	0,845	42,330	0,0133	0,0188		
17	2	235	0,15	0,77	0,82	0,795	0,189	148,379	0,0033	0,0053		
18	2	102	0,16	0,71	0,87	0,786	0,204	63,005	0,0077	0,0125		
19	2	82	0,28	0,75	0,90	0,822	0,341	55,350	0,0097	0,0143		
20	2	62	0,41	0,75	0,90	0,822	0,499	41,850	0,0128	0,0190		
21	2	88	0,40	0,68	0,78	0,728	0,549	46,675	0,0090	0,0170		
22	2	98	0,43	0,89	0,84	0,863	0,497	73,265	0,0081	0,0108		
23	2	53	0,50	0,69	0,86	0,770	0,649	31,450	0,0151	0,0254		
24	2	53	0,45	0,90	0,93	0,915	0,492	44,361	0,0151	0,0180		
25	2	62	0,16	0,77	0,86	0,814	0,197	41,056	0,0128	0,0194		
26	2	16	0,11	0,77	0,86	0,814	0,135	10,595	0,0522	0,0788		

^a Con propósitos didácticos se han modificado algunos datos, y añadido otros, respecto de la base de datos original ofrecida en Sadri y Robertson (1993). Criterio: tipo de medida del criterio (1: rendimiento; 2: conducta). N_i : tamaño muestral del estudio. r_i : coeficiente de correlación empírico. r_{xx} y r_{yy} : coeficientes de fiabilidad de las variables independiente y dependiente. A_i : factor de corrección compuesto. r_{ci} : coeficiente de correlación corregido. w_i : peso de ponderación. S_i^2 : varianza de error de muestreo corregida.

5.1. Metaanálisis con la corrección del error de muestreo

En primer lugar, se presentan los pasos a seguir para realizar un metaanálisis en el que el único artefacto estadístico que se va a corregir es el error de muestreo, obviándose, por tanto, la información referente al error de medida:

1. Cálculo de la correlación media ponderada por el tamaño muestral, a través de la ecuación (7):

$$\bar{r} = \sum N_i r_i / \sum N_i = 879,890 / 2.605 = 0,338$$

2. Cálculo de la varianza observada, mediante la ecuación (6):

$$S_r^2 = [\sum N_i (r_i - \bar{r})^2] / \sum N_i = [\sum N_i (r_i - 0,338)^2] / 2.605 = 81,862 / 2.605 = 0,03142$$

3. Cálculo de la varianza de error de muestreo, mediante la ecuación (8):

$$S_e^2 = (1 - \bar{r}^2) / (\bar{N} - 1) = (1 - 0,338^2) / (100,19 - 1) = 0,00791$$

4. Cálculo de la varianza verdadera, mediante la ecuación (9):

$$S_p^2 = S_r^2 - S_e^2 = 0,03142 - 0,00791 = 0,02351$$

5. Comprobación del supuesto de homogeneidad de las correlaciones con la regla del 75 por 100, mediante la ecuación (10):

$$\text{Regla-75\%} = (100) S_e^2 / S_r^2 = (100)(0,00791) / (0,03142) = 25,17\%$$

6. Comprobación del supuesto de homogeneidad con la prueba Q, mediante la ecuación (11):

$$Q = (k) S_e^2 / S_p^2 = (26)(0,00791) / (0,02351) = 103,28, p < 0,001$$

7. Construcción del intervalo de confianza al 95 por 100, en torno a la correlación media, mediante las ecuaciones (12) y (13):

$$S_{\bar{r}} = (1 - \bar{r}^2) / \sqrt{N - k} = (1 - 0,338^2) / \sqrt{2.605 - 26} = 0,0174$$

$$\bar{r} \pm |z_{\alpha/2}| S_{\bar{r}} = 0,338 \pm (1,96)(0,0174) = 0,372; 0,304$$

Tanto la regla del 75% como la prueba Q de homogeneidad de las correlaciones nos conduce a la conclusión de que la variabilidad observada no se debe sólo a la

varianza del error de muestreo, sino que debe haber variables moderadoras que le afectan. Por tanto, la correlación media obtenida, 0,338, así como su intervalo de confianza (0,304; 0,372), no representa adecuadamente al conjunto de estudios.

El paso siguiente consiste, pues, en la búsqueda de variables moderadoras. Como ejemplo, se analiza a continuación la variable moderadora dicotómica «tipo de criterio», con categorías: «rendimiento» y «conducta». El cuadro 7.3 resume los cálculos. Se observa cómo la correlación media obtenida con medidas de rendimiento ($\bar{r}_1 = 0,365$) es claramente superior a la de los estudios que utilizan como criterio las conductas en el trabajo ($\bar{r}_2 = 0,281$), indicando que ésta es una característica de los estudios que está afectando a la magnitud de la correlación entre las dos variables, si bien en ambos casos la correlación entre expectativas de autoeficacia y rendimiento en el trabajo es significativa, como así lo evidencian los correspondientes intervalos de confianza en torno a la correlación media. No obstante, deben existir otras variables potencialmente moderadoras, ya que todavía queda un alto porcentaje de la varianza observada por explicar: el 80,55 por 100 en las medidas de rendimiento y el 44,29 por 100 en las conductas de trabajo.

CUADRO 7.3

Análisis de la variable moderadora tipo de criterio

Categoría: Rendimiento (N = 10)	Categoría: Conducta (N = 10)
(7) $\bar{r} = \sum N_i r_i / \sum N_i = 640,570 / 1.754 = 0,365$	(7) $\bar{r} = \sum N_i r_i / \sum N_i = 239,320 / 851 = 0,281$
(6) $S_r^2 = [\sum N_i (r_i - 0,365)^2] / 1.754 = 62,409 / 1.754 = 0,03558$	(6) $S_r^2 = [\sum N_i (r_i - 0,281)^2] / 851 = 15,411 / 851 = 0,01811$
(8) $S_e^2 = (1 - 0,365^2) / (109,62 - 1) = 0,00692$	(8) $S_e^2 = (1 - 0,281^2) / (85,1 - 1) = 0,01009$
(9) $S_p^2 = S_r^2 - S_e^2 = 0,03558 - 0,00692 = 0,02866$	(9) $S_p^2 = S_r^2 - S_e^2 = 0,01811 - 0,01009 = 0,00802$
(10) Regla - 75% = $(100)(0,00692) / (0,03558) = 19,45\%$	(10) Regla - 75% = $(100)(0,01009) / (0,01811) = 55,71\%$
(11) $Q = (16)(0,03558) / (0,00692) = 82,26, p < 0,001$	(11) $Q = (10)(0,01811) / (0,01009) = 17,95, p < 0,05$
(13) $S_{\bar{r}} = (1 - \bar{r}^2) / \sqrt{N - k} = (1 - 0,365^2) / \sqrt{1.754 - 16} = 0,0208$	(13) $S_{\bar{r}} = (1 - \bar{r}^2) / \sqrt{N - k} = (1 - 0,281^2) / \sqrt{851 - 10} = 0,0317$
(12) $\bar{r} \pm z_{\alpha/2} S_{\bar{r}} = 0,365 \pm (1,96)(0,0208) = 0,406; 0,324$	(12) $\bar{r} \pm z_{\alpha/2} S_{\bar{r}} = 0,281 \pm (1,96)(0,0317) = 0,343; 0,219$

5.2. Metaanálisis de las correlaciones corregidas

Si además del error de muestreo interesa aprovechar la información sobre los coeficientes de fiabilidad de las dos variables para corregir el influjo atenuante del error de medida, se seguirán estos pasos:

1. Los datos preliminares necesarios son la correlación empírica, r_i , el tamaño muestral, N_i , y el factor de corrección compuesto, A_i , que en este caso se definirá como el producto de los factores de corrección del error de medida en las variables dependiente e independiente: $A_i = a_1 a_2$. Los factores a_1 y a_2 se obtienen mediante las ecuaciones (14) y (15), respectivamente. La séptima columna del cuadro 7.2 muestra el valor alcanzado por A_i en cada estudio.
2. A continuación, se calculan las correlaciones corregidas, r_{ci} , mediante la ecuación (24), y los pesos de ponderación, w_i , mediante la ecuación (25). Las columnas octava y novena del cuadro 7.2 muestran los resultados de estos cálculos.
3. Con objeto de estimar la varianza del error de muestreo corregida en cada estudio, se calcula primero la correlación media de las correlaciones no corregidas mediante la ecuación (27), obteniendo el valor: $r_+ = \sum w_i r_i / \sum w_i = 613,019 / 1.804,936 = 0,3396$. A continuación, se estima la varianza de error de muestreo de cada estudio mediante la ecuación (26), alcanzando los valores que figuran en la décima columna del cuadro 7.2. Y, por último, se obtiene la varianza de error de muestreo corregida para cada estudio mediante la ecuación (28), con los resultados presentados en la última columna de dicho cuadro.
4. El paso siguiente consiste en realizar los cálculos metaanalíticos propiamente dichos:

- a) Cálculo de la correlación media corregida mediante la ecuación (29):

$$\bar{r}_c = \sum w_i r_{ci} / \sum w_i = 732,946 / 1.804,936 = 0,406$$

- b) Cálculo de la varianza total de las correlaciones corregidas mediante la ecuación (30):

$$S_c^2 = \sum w_i (r_{ci} - \bar{r}_c)^2 / \sum w_i = \sum w_i (r_{ci} - 0,406)^2 / \sum w_i = 81,428 / 1.804,936 = 0,04511$$

- c) Cálculo de la varianza de error de muestreo corregida mediante la ecuación (31):

$$S_{e_c}^2 = \sum w_i S_{e_i}^2 / \sum w_i = 20,691 / 1.804,936 = 0,01146$$

- d) Cálculo de la varianza verdadera de las correlaciones corregidas mediante la ecuación (32):

$$S_p^2 = S_c^2 - S_{e_c}^2 = 0,04511 - 0,01146 = 0,03365$$

- e) Comprobación de la hipótesis de homogeneidad de las correlaciones corregidas. Para ello, se aplica la regla del 75% según la ecuación (10), pero adaptada a su uso con las varianzas de las correlaciones corregidas: Regla-75% = $(100)(S_{e_c}^2)/(S_c^2) = (100)(0,01146)/(0,04511) = 25,4\%$. También puede aplicarse la prueba Q de homogeneidad a través de la ecuación (11):

$$Q = (k)(S_c^2)/(S_{e_c}^2) = (26)(0,04511)/(0,01146) = 102,34, p < 0,001$$

Así pues, mientras que en el metaanálisis con corrección sólo del error de muestreo la correlación media era de 0,338, al corregir los efectos del error de medida en las dos variables, la correlación media se eleva a 0,406. No obstante, las pruebas de homogeneidad continúan resultando significativas, lo que indica que, una vez controlados estos tres artefactos estadísticos, todavía resta casi un 75 por 100 de la varianza observada por explicar. El análisis debería, pues, continuar poniendo a prueba el influjo de variables moderadoras, para lo cual se repetirían los análisis realizados con cada subgrupo o categoría de la variable moderadora.

REFERENCIAS BIBLIOGRÁFICAS

- Bangert-Drowns, R. L. (1986): «Review of developments in meta-analytic method», *Psychological Bulletin*, 99, págs. 388-399.
- Cohen, J. (1988): *Statistical power analysis for the behavioral sciences*, 2.ª ed., Erlbaum, Hillsdale, Nueva Jersey.
- Cooper, H. M. (1989): *Integrating research: A guide for literature reviews*, 2ª ed., Sage, Beverly Hills.
- Cooper, H. M., y Hedges, L.V. (eds.) (1994): *The handbook of research synthesis*, Sage, Nueva York.
- Damanpour, F. (1991): «Organizational innovation: A meta-analysis of effects of determinants and moderators», *Academy of Management Journal*, 34, págs. 555-590.
- Farley, J. U., y Lehmann, D. (1986): *Meta-analysis in marketing: Generalization of response models*, Lexington Books, Lexington.
- Glass, G. V.; McGaw, B., y Smith, M. L. (1981): *Meta-analysis in social research*, Sage, Beverly Hills.
- Gómez, J. (1987): *Meta-análisis*, PPU, Barcelona.
- Hedges, L. V., y Olkin, I. (1985): *Statistical methods for meta-analysis*, Academic Press, Orlando.
- Hunter, J. E., y Schmidt, F. L. (1990): *Methods of meta-analysis: Correcting error and bias in research findings*, Sage, Beverly Hills.
- Hunter, J. E.; Schmidt, F. L., y Jackson, G. B. (1982): *Meta-analysis: Cumulating research findings across studies*, Sage, Beverly Hills.
- Marín-Martínez, F., y Sánchez-Meca, J. (1998): «Testing dichotomous moderators in meta-analysis», *Journal of Experimental Education*, 67, págs. 69-81.
- Marín-Martínez, F., y Sánchez-Meca, J. (1999): «Averaging dependent effect sizes in meta-analysis: A cautionary note about procedures», *Spanish Journal of Psychology*, 2, págs. 32-38.
- Martínez Arias, M. R. (1995): *Psicometría: Teoría de los tests psicológicos y educativos*, Síntesis, Madrid.
- Orwin, R. G. (1983): «A fail-safe N for effect size in meta-analysis», *Journal of Educational Statistics*, 8, págs. 157-159.
- Rosenthal, R. (1976): *Experimenter effects in behavioral research*, Irvington, Nueva York.
- Rosenthal, R. (1991): *Meta-analytic procedures for social research* (rev. ed.), Sage, Newbury Park.
- Sadri, G., y Robertson, I. T. (1993): «Self-efficacy and work-related behaviour: A review and meta-analysis», *Applied Psychology: An International Review*, 42, págs. 139-152.
- Sagie, A., y Koslowski, M. (1993): «Detecting moderators with meta-analysis: An evaluation and comparison of techniques», *Personnel Psychology*, 46, págs. 629-640.
- 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 (coords.), *Tratado de psicología general. I: Historia, teoría y método*, Alhambra, Madrid, págs. 617-669.
- 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 & Quantity*, 31, págs. 385-399.
- Sánchez-Meca, J., y Marín-Martínez, F. (1998a): «Testing continuous moderators in meta-analysis: A comparison of procedures», *British Journal of Mathematical & Statistical Psychology*, 51, páginas 311-326.
- Sánchez-Meca, J., y Marín-Martínez, F. (1998b): «Weighting by inverse-variance or by sample size in meta-analysis: A simulation study», *Educational & Psychological Measurement*, 58, págs. 211-220.

- Schmidt, F. L., y Hunter, J. E. (1977): «Development of a general solution to the problem of validity generalization», *Journal of Applied Psychology*, 62, págs. 529-540.
- Smith, M. L., y Glass, G. V. (1977): «Meta-analysis of psychotherapy outcome studies», *American Psychologist*, 32, págs. 752-760.
- Szymanski, D. M.; Bahradwaj, S. G., y Varadarajan, P. R. (1993): «An analysis of the market shareprofitability relationship», *Journal of Marketing*, 57, págs. 1-18.

Metodología para la investigación en marketing y dirección de empresas

Francisco José Sarabia Sánchez
(Coordinador)



PIRÁMIDE

COLECCIÓN «ECONOMÍA Y EMPRESA»

Director:
Miguel Santesmases Mestre
Catedrático de la Universidad de Alcalá de Henares

Diseño de cubierta: C. Carabina

Realización de cubierta: Anaf Miguel

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeren, plagiaren, distribuyeren o comunicaren públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier otro medio, sin la preceptiva autorización.

© Francisco José Sarabia Sánchez
© EDICIONES PIRÁMIDE, S. A., 1999
Juan Ignacio Luca de Tena, 15. 28027 Madrid
Teléfono: 91 393 89 89. Fax: 91 742 36 61
Depósito legal: M. 28.914-1999
ISBN: 84-368-1342-1
Printed in Spain
Impreso en Anzos, S. L.
Fuenlabrada (Madrid)

*A nuestros padres. A ellos
debemos en gran medida lo que
somos y a dónde hemos llegado.
Para ellos nuestro amor y
reconocimiento por todos los días
que han pasado sin sus hijos.*

Relación de autores

Enrique Bigné Alcañiz

Catedrático de Comercialización e Investigación de Mercados de la Universidad Jaime I.

José Joaquín Céspedes Lorente

Profesor titular del Área de Organización de Empresas de la Universidad de Almería.

Ángel Fernández Nogales

Profesor titular de Comercialización e Investigación de Mercados de la Universidad Autónoma de Madrid.

Idelfonso Grande Esteban

Profesor titular de Comercialización e Investigación de Mercados de la Universidad Pública de Navarra.

Jesús Gutiérrez Cillán

Catedrático de Comercialización e Investigación de Mercados de la Universidad de Valladolid.

Teodoro Luque Martínez

Profesor titular de Comercialización e Investigación de Mercados de la Universidad de Granada.

Waldo Pérez Aguiar

Profesor titular de E. U. de Organización de Empresas de la Universidad Politécnica de Madrid.

Ignacio Rodríguez del Bosque

Catedrático de Comercialización e Investigación de Mercados de la Universidad de Cantabria.

Ana Isabel Rodríguez Escudero

Profesora titular de Comercialización e Investigación de Mercados en el Departamento de Economía y Administración de Empresas de la Universidad de Valladolid.

Salvador Ruiz de Maya

Profesor titular de Comercialización e Investigación de Mercados en la Universidad de Murcia.

Manuel Sánchez Pérez

Profesor titular de Comercialización e Investigación de Mercados de la Universidad de Almería.

Julio Sánchez Meca

Profesor titular de Metodología de las Ciencias del Comportamiento de la Universidad de Murcia.

Francisco José Sarabia Sánchez

Profesor titular de Comercialización e Investigación de Mercados de la Universidad Miguel Hernández.

Rodolfo Vázquez Casielles

Catedrático de Comercialización e Investigación de Mercados de la Universidad de Oviedo.

Índice

Presentación.....	17
1. Ciencia y método científico (Ana I. Rodríguez y Jesús Gutiérrez)...	21
1. La ciencia como fenómeno.....	21
1.1. Conocimiento científico y sociedad.....	21
1.2. Conocimiento científico versus conocimiento vulgar.....	22
2. Concepto de ciencia.....	25
3. Los elementos de la construcción científica.....	30
3.1. Los conceptos.....	30
3.2. Los modelos.....	32
3.3. Las hipótesis.....	33
3.4. Las leyes científicas.....	35
3.5. Las teorías.....	37
4. Tipos de métodos científicos.....	38
5. Orientaciones sobre la metodología científica.....	40
2. La investigación científica (Jesús Gutiérrez y Ana I. Rodríguez).....	51
1. Ámbitos de la actividad tecnocientífica.....	51
2. La investigación científica: proceso y tipología. Referencia al caso de las tesis doctorales.....	54
3. La investigación descriptiva.....	57
4. La investigación exploratoria.....	59
5. La investigación explicativa.....	64
6. Una reflexión final respecto de las tesis doctorales.....	66

3. Definición del tema a investigar (<i>Francisco J. Sarabia</i>)	69
1. El rol del investigador	69
1.1. Rol externo del científico	69
1.2. Rol interno del científico	71
1.3. Papeles específicos del investigador	72
1.4. El papel del investigador ante la información	73
1.5. El cientificismo	75
2. Determinación del problema a investigar	77
3. Problemas e hipótesis	80
3.1. El problema objeto de investigación	80
3.2. Dos herramientas para fijar el problema objeto de investigación	82
3.3. La hipótesis: concepto y reglas básicas	84
3.4. Tipos de hipótesis	86
3.5. Las hipótesis mal formuladas	91
4. Fuentes de Información para la Investigación. Las bases de datos (<i>Ildefonso Grande</i>)	95
1. Introducción a la importancia de la información secundaria	95
2. Clases de bases de datos	96
3. ¿Por qué se crean bases de datos y se emplea información secundaria en las investigaciones?	97
4. Estructura de las bases de datos	99
5. Bases de datos internas	102
5.1. Creación de una base de datos interna	103
6. Bases de datos externas	109
6.1. Bases de datos en España	109
6.2. Evaluación de las bases de datos externas. Causas y criterios	109
5. Fuentes de información para la investigación. La red Internet y las bases de datos electrónicas (<i>José J. Céspedes</i>)	125
1. Internet como fuente de información	125
2. Internet y la investigación científica	127
3. Fuentes de información económica en Internet	129
4. Bases de datos de contenido empresarial	134
4.1. Bases de datos no españolas	135
4.2. Bases de datos españolas	139
4.3. Fuentes bibliográficas	141
5. Revistas académicas, revistas electrónicas y archivos de <i>Working-Papers</i> en Internet	144
6. Un ejemplo práctico	145
6. La experimentación (<i>Salvador Ruiz</i>)	151
1. Naturaleza de la experimentación	152
2. Tipos de experimentación	153

3. La experimentación como proceso	157
4. Los sujetos de la muestra	159
5. Los problemas del control en la experimentación y la validez	160
6. Las variables independientes	162
7. Las variables dependientes	163
8. Los diseños experimentales y el análisis de los datos	165
9. Una aplicación real del diseño experimental	170
7. Metaanálisis para la investigación científica (<i>Jullo Sánchez</i>)	173
1. Metaanálisis: antecedentes y concepción	174
2. Fases de un metaanálisis	175
2.1. Formulación del problema	175
2.2. Búsqueda de la literatura	175
2.3. Codificación de variables	177
2.4. Análisis estadístico e interpretación	178
2.5. Publicación del estudio	179
3. La medida de los resultados	179
3.1. La familia <i>d</i>	180
3.2. La familia <i>r</i>	183
4. El metaanálisis psicométrico	184
4.1. Metaanálisis con la corrección del error de muestreo	185
4.2. Metaanálisis con la corrección de otros artefactos	189
5. Una aplicación práctica	194
5.1. Metaanálisis con la corrección del error de muestreo	194
5.2. Metaanálisis de las correcciones corregidas	198
8. La Investigación cualitativa (<i>Ángel Fernández</i>)	201
1. La investigación cualitativa en el área empresarial y comercial	201
2. Características metodológicas de las técnicas cualitativas	202
2.1. Planteamiento general de los estudios cualitativos	203
2.2. Técnicas principales	204
2.3. Muestreo y trabajo de campo	205
2.4. Análisis de resultados e informe	206
3. La entrevista en profundidad	207
3.1. Características principales	207
3.2. Tipos de entrevistas en profundidad	207
3.3. Desarrollo de la entrevista	208
4. La reunión de grupo	208
4.1. Características principales	208
4.2. El grupo de enfoque	209
4.3. La tormenta de ideas o <i>brainstorming</i>	210
5. Las técnicas proyectivas	211
5.1. Características principales	211
5.2. Tipos de técnicas proyectivas	211

6.	Las técnicas de observación	213
6.1.	Características principales	213
6.2.	Procedimientos de observación	215
6.3.	Tipos de observación	215
6.4.	Principales aplicaciones de la observación	216
7.	Los estudios ómnibus cualitativos	217
8.	Un caso real de aplicación de técnicas cualitativas	218
8.1.	Resultados	221
9.	El estudio de casos (Waldo Pérez)	225
1.	Introducción	225
2.	Fundamentos del estudio de casos	226
3.	Diseño e implantación del estudio de casos	229
3.1.	El objetivo de la investigación	230
3.2.	El marco teórico	230
3.3.	La unidad y el nivel de análisis	230
3.4.	La selección de casos	231
3.5.	El caso piloto	232
3.6.	El protocolo de cada caso	233
3.7.	Informe general: resultados y conclusiones	237
4.	La generalización de los resultados	237
5.	El método del caso aplicado al análisis de la acumulación de activos tecnológicos	239
6.	Las conclusiones	240
10.	El análisis de contenido (Enrique Bigné)	253
1.	Introducción	253
2.	Concepto de análisis de contenido	255
3.	Características del análisis de contenido	257
4.	Formas de análisis	258
5.	Fases del estudio del contenido	260
6.	Aplicación práctica	263
11.	Encuestas para la investigación (Teodoro Luque)	271
1.	Introducción	271
2.	Modalidades de encuestas	273
3.	Encuestas autoadministradas	274
3.1.	Encuesta postal	274
3.2.	Otras encuestas autoadministradas impresas	277
3.3.	Encuesta por fax	278
3.4.	Encuesta por correo electrónico	279
3.5.	Encuestas mediante discos enviados por correo	280
3.6.	CAPÍ (Computer-Assisted Personal Interviewing)	280

4.	Encuestas a distancia con entrevistador	281
4.1.	Encuesta telefónica	282
4.2.	Encuesta telefónica grabada	284
4.3.	CATI (Computer Assisted Telephone Interviewing)	285
5.	Encuesta con entrevistador cara a cara	287
5.1.	Encuesta personal a domicilio	287
5.2.	Otras modalidades de la encuesta personal	289
6.	Procedimientos de entrevista de carácter permanente	290
6.1.	Los paneles	290
6.2.	Ómnibus	298
7.	Aplicaciones y comparación entre los diferentes tipos de encuesta	301
12.	La construcción del cuestionario (Ignacio Rodríguez)	305
1.	Definición de la información necesaria para lograr los objetivos de la investigación	305
2.	La selección del modelo de entrevista a realizar	306
3.	Definición de la información a recoger en el cuestionario	310
4.	Tipos de información y preguntas a realizar	313
5.	¿Cómo hay que preguntar?	316
6.	El orden de las preguntas	319
7.	La presentación del cuestionario	327
8.	Comprobación o verificación del cuestionario	328
13.	Construcción de escalas de medida (Francisco J. Sarabia)	331
1.	Definición de medición y proceso para generar medidas	332
2.	Concepto y variables: definición y tipos	334
3.	Valor numérico y escalas naturales	337
4.	Escalas compuestas: concepto y tipos	340
5.	Técnicas para formar escalas compuestas	344
5.1.	Diferencial semántico (DS)	344
5.2.	Formato Stapel	348
5.3.	Suma constante	349
5.4.	Técnica de comparación-par de Thurstone	350
5.5.	Técnica aditiva de Likert	352
6.	La decisión del tipo de escalamiento	353
7.	Factores de calidad de una escala	355
14.	Validez y fiabilidad de escalas (Manuel Sánchez y Francisco J. Sarabia)	361
1.	Fuentes de variación y errores de medición	362
2.	Análisis de la fiabilidad	364
2.1.	Medición de la fiabilidad	365
2.2.	El coeficiente α de Cronbach (1951)	368

2.3. Otros métodos para determinar la fiabilidad	372
2.4. Influencia de la fiabilidad en la predicción de los valores verdaderos	374
2.5. Comparación entre valores de α	375
3. Análisis de la validez	377
3.1. Validez de contenido	380
3.2. Validez de construcción o de concepto	381
3.3. Validez predictiva o relacionada con criterios	384
3.4. Medición de la validez. La matriz multiconcepto-multimétodo	385
4. Relación entre validez y fiabilidad	388
15. Evaluación de la investigación científica (Rodolfo Vázquez)	393
1. Evaluación de artículos y su publicación en revistas internacionales	393
2. Alternativas y criterios de evaluación de artículos en revistas españolas	405
3. Criterios de evaluación para participar en congresos	410
4. Evaluación de proyectos de investigación y realización de tesis doctorales	413
5. Dirección para futuras líneas de investigación	417
Anexo estadístico	421
Glosario de términos	429

Presentación

En 1998 se realizó un Taller de Metodología organizado por ACEDE en Arnedillo (La Rioja) al que asistí como un oyente más. Lo que viví y compartí con otros compañeros me ayudó a cristalizar la idea de realizar un libro sobre aspectos metodológicos de la investigación científica que fuese aplicable a mi área de trabajo (Marketing) y a la de Dirección y Organización de Empresas. En sí, la idea no era nueva, ya que imparto un curso de doctorado sobre metodología científica en Marketing y sabía que la falta de un manual o texto específico ha preocupado a todos los que trabajamos en estas áreas.

Con la idea puesta sobre el papel, contacté con dieciséis compañeros y catorce me respondieron positivamente. Nos pusimos a la tarea y el resultado es esta obra, que, pese a tener un título que pudiera parecer ostentoso, ofrece los aspectos metodológicos más relevantes que deben ser conocidos por quienes nos dedicamos a enseñar y a trabajar en la investigación. Es evidente que una obra de este tipo no puede —ni debe— contemplar todos los posibles elementos metodológicos. Algunos aspectos como la modelización, la cuasi-experimentación, las fuentes clásicas de información, algunas metodologías cualitativas y todos los aspectos estadísticos (muestras, análisis, etc.) han sido obviados. Su inclusión no sólo haría que la obra perdiese manejabilidad, sino que se convertiría en un texto casi enciclopédico por su contenido y extensión.

La obra consta de quince capítulos agrupados en siete bloques: iniciación al trabajo científico, fuentes de información, metodologías o formatos cuantitativos, metodologías o formatos cualitativos, encuestas y cuestionarios, medición y escalas de medida y evaluación final del trabajo. La estructura de la obra es clásica, ya que se inicia con los aspectos de la ciencia y del método científico y termina con aquellos relacionados con la evaluación del trabajo ya realizado. Sin embargo, el contenido trata de ser innovador, pues se ha orientado para que no sólo se sepa «qué hacer» sino también «cómo hacer lo que hay que hacer».