

Cuaderno de Investigación en la Educación, número 2, octubre 1990
Centro de Investigaciones Educativas
Facultad de Educación, Río Piedras
Universidad de Puerto Rico, Recinto de Río Piedras

LA SINTESIS DE INVESTIGACIONES A TRAVES DEL METAANALISIS

José N. Caraballo
Universidad de Puerto Rico
Recinto de Río Piedras
Facultad de Educación

En la educación tanto como en las ciencias sociales y naturales, las revisiones de literatura se llevan a cabo con uno de los siguientes propósitos: para obtener un entendimiento más completo posible de un problema de investigación, para obtener conclusiones generales de un grupo de investigaciones, y para identificar áreas que requieran de investigación. Al revisar la literatura el investigador tiene que llevar a cabo una síntesis de la información presentada en los distintos artículos. En el caso en que se desea evaluar la evidencia a favor o en contra de una hipótesis, el investigador, como paso preliminar, debe pasar juicio sobre la calidad de las investigaciones para así excluir de la síntesis aquellas cuyos resultados puedan ser dudosos. Una vez hecho esto, se decide a favor o en contra de la hipótesis basado en el número de investigaciones que la apoye o la refute.

Existen varios problemas asociados a la forma tradicional de revisar la literatura. En primer lugar se encuentra el hecho de que rara vez la colección de artículos que se sintetizan constituyen una muestra representativa de la población de artículos sobre el tema en cuestión. Esto trae como consecuencia la posibilidad de que se generalicen resultados de un grupo de estudios cuyas características no son representativas de la "población" de estudios realizados sobre el tema. En segundo lugar, se puede demostrar estadísticamente que el utilizar la proporción de artículos a favor o en contra como criterio tiende hacia la retención de la hipótesis nula aún cuando no sea cierta (Hedges y Olkin, 1985). Por último, el hecho de que en un grupo de artículos la hipótesis sea retenida o rechazada no implica que los resultados de las investigaciones sean consistentes. Por ejemplo, aún cuando en varias investigaciones se rechace la hipótesis nula, es posible que hallan diferencias significativas entre los resultados.

Meta-análisis

Los problemas presentados en la sección anterior requieren de alternativas para la síntesis de investigaciones. El meta-análisis es un método cuantitativo que pretende aminorar los problemas mencionados. En la técnica de meta-análisis los resultados de los estudios se expresan como un índice que representa el grado de relación entre las variables o la magnitud del efecto que produce un tratamiento experimental. Una vez que se han obtenido estos índices, se combinan para así llegar a conclusiones generales. Las técnicas meta-analíticas pueden ser utilizadas con una variedad de metodologías y

diseños de investigación. El requisito principal es que las investigaciones generen resultados que se puedan cuantificar.

El problema de sintetizar investigaciones sobre problemas o hipótesis similares comenzó a ser estudiado por varios autores en la década de los años 30. Desde los inicios aparecieron dos corrientes distintas. Tippet (1930), Fisher (1932) y Pearson (1933) desarrollaron técnicas analíticas en las que se combinaban los resultados de las pruebas de significación. En esta corriente se combinaban los diferentes valores de probabilidad p de los distintos estudios para obtener un valor p global. Si éste era menor que el nivel de significación, digamos 0.05, entonces se concluía que el tratamiento experimental o la relación entre variables era significativa. De lo contrario se concluía que el tratamiento experimental no era efectivo, o que no había relación entre las variables bajo estudio.

Por otro lado, Cochran (1937) y Yates y presentados en la tabla siguiente: Cochran (1938) se preocuparon por obtener, no solo una decisión global sobre una hipótesis, sino Tabla 1: Datos para el cálculo de la magnitud del estimar la magnitud del efecto producido por el efecto-tratamiento experimental o la magnitud de la relación entre las variables. Aunque esta segunda corriente también data de la década del 30, no fue hasta 1976 que Gene V. Glass la desarrolló y aplicó con éxito en la educación y otras ciencias sociales. A Glass se le atribuye el nombre meta-análisis a esta metodología. Debido a que todavía es reciente, aún hoy día se están refinando y desarrollando técnicas de análisis. Las técnicas de análisis que se presentarán más adelante son debidas en su mayoría a Hedges (1985) y Hedges y Olkin (1985).

Cálculo de estimados de efectos

Los ejemplos y técnicas que se presentan en este artículo se basan en el estimador de la magnitud del efecto producido por un tratamiento experimental según lo definió Glass (1976); este, conocido como el Estimador de Glass, se define como:

$$(1) \quad d = \frac{YE - YC}{SC}$$

donde YE representa el promedio en la variable dependiente obtenido por el grupo que recibe el tratamiento experimental, Yc representa el promedio en la variable dependiente obtenido por el grupo que sirve de control o comparación y Sc es la desviación estándar del grupo control o el promedio ponderado de las desviaciones estándar de los dos grupos.

Como ejemplo del cálculo del estimado de la magnitud del efecto, considere los datos presentados en la tabla siguiente:

Tabla 1: Datos para el cálculo de la magnitud del efecto.

	Media	d.e.
Control	8.6	3.20
Experimental	11.4	2.67

Si utilizamos la desviación estándar del grupo control como referencia, el cálculo de d se lleva a cabo como sigue:

$$\begin{aligned} D &= \frac{Ye - Yc}{SC} \\ &= \frac{11.4 - 8.6}{3.20} \\ &= 2.8 \end{aligned}$$

$$\frac{3.2}{4} = 0.875$$

El valor de **d**, 0.875, es la medida de la magnitud del efecto. Mientras mayor es el valor absoluto de este número, mayor es el efecto que produce el tratamiento. Un valor de **d** cercano a 0.2 se considera un efecto bajo; un valor de **d** cercano a 0.5 se considera un efecto mediano; y un valor cercano a 0.8 se considera un efecto alto.

Para interpretar el índice hay que notar que el mismo es una razón entre la diferencia entre los promedios de los grupos dividida entre la desviación estándar del grupo que sirve como control. Por lo tanto, el índice es una medida de cuantas desviaciones del grupo control hay entre los promedios de los grupos control y experimental. El promedio del grupo experimental se encuentra a 0.875 desviaciones estándar sobre el promedio del grupo control. Utilizando una tabla de la curva normal se puede demostrar que la posición del promedio del grupo experimental coincide con la percentila 81 del grupo control. Por lo tanto, se puede decir que el efecto del tratamiento experimental es tal que si se aplica a un sujeto promedio, su percentila aumentará de 50 a 81.

Ejemplos de Meta-análisis

Antes de considerar los aspectos metodológicos y analíticos presentaremos dos ejemplos de meta-análisis publicados en la literatura. El primero de ellos es un meta-análisis de la efectividad de la enseñanza por medio de computadoras en el nivel universitario llevado a cabo por Kulik, Kulik y Cohen (1980). El mismo tuvo como propósito determinar el efecto de la instrucción por computadoras sobre las siguientes variables: aprovechamiento, tiempo de aprendizaje, bajas en cursos, actitudes hacia la enseñanza asistida por la computadora y las actitudes hacia la materia enseñada.

La muestra de artículos sobre el tema fue obtenida al examinar varios índices a través del sistema DIALOG. Estos índices fueron: Compendex, Comprehensive Dissertation Abstracts, ERIC, Inspec, Psychological Abstracts, Scisearch y Social Scisearch. En adición, se examinaron las bibliografías de los artículos encontrados. La búsqueda inicial dio como resultado más de quinientos artículos. Los títulos y resúmenes de los artículos fueron examinados para descartar aquellos que no se relacionaban directamente con el problema en cuestión. Como consecuencia de esto, el número de artículos se redujo a 180. Los 180 artículos fueron leídos para determinar si tenían la información necesaria para incluirlos en el análisis. Solo 59 artículos cumplieron este requisito y fueron los que se utilizaron para la síntesis.

Tabla 2. Resultado del estudio sobre la efectividad de la enseñanza asistida por la computadora

Variable dependiente	N	d
Aprovechamiento	54	0.2
Tiempo de estudio	8	0.67
Bajas	13	n.s.
Actitudes CAI	11	0.24
Actitudes materia	7	0.18

En la tabla anterior N es el número de estudios utilizados en la síntesis para cada variable de interés y des el promedio del estimado del efecto encontrado en los estudios. Como se puede ver en esta tabla, la enseñanza asistida por la computadora tiene un efecto

moderado sobre el tiempo de estudio, un efecto bajo sobre el aprovechamiento y las actitudes, y un efecto no significativo sobre el número de bajas en los cursos.

Como segundo ejemplo de meta-análisis se presenta un estudio realizado por Guskey y Gates (1986) sobre la efectividad del método de "mastery learning" en los niveles elemental y secundario. Las preguntas que se investigaron fueron las siguientes: ¿Cuán efectivo es el "mastery learning"?, ¿Depende la efectividad del nivel escolar en que se utiliza?, y ¿Depende la efectividad de la materia escolar enseñada?

Para recopilar los artículos sobre el tema los investigadores consultaron los índices Dissertation Abstracts International, ERIC y Psychological Abstracts. En adición, consultaron el libro *Mastery Learning: A Comprehensive Bibliography* (Ilymel, 1982) y un artículo sobre el tema publicado por Block y Burnes (1976). El número inicial de títulos encontrados fue de 1000. Al examinar los títulos y sus resúmenes, el número se redujo a 234, de los cuales solo se pudieron localizar 144. Al evaluar los artículos el número se redujo a 38, veintisiete de los cuales eran estudios realizados en la escuela elemental y secundaria.

Aspectos metodológicos

La metodología a seguirse en un meta-análisis es similar a la que se sigue en los estudios primarios y se puede conceptualizar en cinco etapas:

1. Formulación del problema
2. Recopilación de datos
3. Evaluación de datos
4. Análisis e interpretación de datos
5. Presentación de resultados

Algunas de las diferencias metodológicas hacen que sea más difícil de llevar a cabo que una investigación primaria. En particular, el meta-análisis debe llevarse a cabo siguiendo un procedimiento riguroso para evitar que los resultados se vicien.

Formulación del problema –Como en toda investigación, antes de comenzar a recopilar datos es necesario que el problema bajo estudio esté bien definido. Esto incluye redactar las preguntas de investigación claramente y definir constructos tales como tratamientos, controles y variables dependientes. Debe hacerse notar que, como se pretende integrar los resultados de una colección de estudios primarios, cabe la posibilidad de encontrar entre ellos definiciones alternas de las variables y los controles. La definición clara de las variables y controles servirá en la etapa de evaluación de datos como criterio de inclusión de los estudios. Como un ejemplo, considere que en estudios sobre aprovechamiento éste se ha medido en formas variadas: mediante pruebas construidas por los maestros o los investigadores, mediante pruebas estandarizadas y mediante el ex índice académico. En un meta-análisis sobre este tema, ¿se deben incluir todas estas definiciones operacionales distintas?

En cuanto a los controles, considere un ejemplo de investigación médica. Para determinar la efectividad de una droga experimental, ésta se administra a un grupo de pacientes, mientras que a otro grupo de pacientes no se le administra ningún tratamiento, o se le administra un placebo. La magnitud relativa de la efectividad del tratamiento puede depender del tipo de control que se utilice para hacer la comparación.

En adición a la definición de variables, se debe decidir el modo de inquirir. El meta-análisis puede ser llevado a cabo para poner a prueba una o más hipótesis establecidas a priori, o puede ser llevado a cabo para generar hipótesis nuevas que

deberían estudiarse en una investigación primaria. Los ejemplos que se presentaron en la sección anterior fueron llevados a cabo para poner a prueba hipótesis. En el estudio de Guskey y Gates (1986), sin embargo, es posible tratar de explicar el porqué de la gran variación obtenida en los efectos mediante alguna variable interventora. De encontrarse alguna posible explicación, ésta deberá ponerse a prueba posteriormente en un estudio primario diseñado para ese propósito.

Un último aspecto a considerarse consiste en decidir qué información relevante debe ser extraída de los artículos para ser codificada. La Tabla 4 presenta algunas de las variables y características que pueden ser relevantes. Entre más exhaustiva sea la lista de características a ser codificadas, mayor cantidad de hipótesis explicativas se podrán poner a prueba.

Tabla 4. Variables codificables

1. Información de referencia
2. Características de la muestra
3. Características de los tratamientos
4. Características de los controles
5. Características de las personas que administran los tratamientos
6. Características del diseño de investigación
7. Tipos de variables independientes

Recopilación de datos - Esta etapa es una de las más críticas y difíciles de llevar a cabo. Consiste en obtener una muestra de los artículos publicados sobre el tema. Existen varias formas para encontrar títulos de artículos: examinar las bibliografías publicadas en revisiones de literatura, preguntar a los especialistas en el área las referencias de estudios importantes, consultar índices tales como ERIC, Psychological Abstracts, y otros, consultar índices de citas, y examinar manualmente las revistas importantes.

En la medida en que esta búsqueda sea completa, mayor será la posibilidad de generalizar los resultados del meta-análisis. En el caso de que el número de artículos sea extremadamente grande, se puede hacer uso de técnicas de muestreo para analizar un subconjunto de ellos.

El mayor problema que ocurre al seleccionar la muestra de estudios es que muchos no se publican si no se obtienen resultados significativos. Por esta razón, si solo se seleccionan artículos publicados en revistas de prestigio, se obtendrá una muestra que estará viciada hacia resultados significativos. Otro problema relacionado es que muchos autores que obtienen resultados no significativos tienden a no incluir toda la información necesaria para hacer el cálculo de los efectos.

Evaluación de datos - Una de las críticas iniciales al meta-análisis, según fue desarrollado por Glass, es que se recomendaba la inclusión de cualquier estudio sin pasar juicio de su calidad (Glass, McGaw y Smith, 1981). Los críticos argumentan que los resultados no pueden ser muy confiables si la calidad de los estudios incluidos en la síntesis es baja. Slavin (1986) ha propuesto que la evidencia presentada en los distintos informes sea evaluada y que solo se incluya en el análisis la mejor evidencia. Claramente, si no se toma en cuenta la calidad puede haber problemas, pero, por otro lado, no hay criterios de inclusión que sean objetivos.

Entre las consideraciones que deben tomarse en cuenta para la inclusión de artículos se encuentra examinar los procedimientos seguidos en cada uno y determinar hasta que punto son adecuados. En adición, se puede analizar la fortaleza ante la amenaza a la validez interna y externa (Campbell y Stanley, 1974). Si se codifica los estudios de acuerdo a la metodología y fortaleza, se hace posible determinar analíticamente su efecto sobre los resultados. Por lo tanto, se sugiere que los criterios de inclusión no sean demasiado rígidos y que se codifique toda la información que sea relevante en cuanto a la calidad del estudio.

Análisis de datos - El análisis de datos en meta-análisis tiene tres etapas fundamentales: cálculo del estimado de la magnitud del efecto, combinación de los estimados para obtener un estimado global, y análisis de la homogeneidad de los estudios. El análisis de homogeneidad tiene como propósito determinar si hay consistencia en los resultados de los distintos estudios. En el caso en que no haya homogeneidad, el estimado global del efecto pierde sentido y se debe proseguir el análisis hasta encontrar una variable que explique la falta de consistencia en los resultados. En un estudio sobre diferencias en habilidad espacial debido al sexo se encontró una falta de homogeneidad que fue explicada como producto de la utilización de dos tipos distintos de instrumentos de medición. En este caso es necesario calcular la magnitud del efecto para cada uno de los instrumentos. Hasta ahora la discusión se ha centrado en el uso del estimador de Glass, sin embargo, el meta-análisis puede ser llevado a cabo con otros índices.

Conclusión

El meta-análisis es una técnica reciente que ha ido ganando aceptación rápidamente. En sus inicios la técnica fue criticada en cuatro aspectos fundamentales (Wolf, 1986): 1. Se combinan resultados de estudios que utilizan distintas metodologías e instrumentos de medición, 2. Se mezclan resultados de estudios de buena calidad con estudios de calidad cuestionable, 3. Como hay una tendencia a publicar estudios con resultados significativos, el meta-análisis tenderá hacia resultados significativos también, y 4. En ocasiones se calculan varios estimados de un mismo artículo y esto tiende a inflar el estimado global de la magnitud del efecto.

Las primeras dos críticas, aunque válidas, pueden ser investigadas analíticamente si se codifica la metodología, instrumentos y otras características de los estudios así como también su calidad. De esta manera, se puede comprobar si de hecho los resultados de los estudios están influenciados por estas variables. En el caso de que así sea, se pueden descartar del análisis aquellos estudios de baja calidad, o informar un índice global para cada nivel de la variable en cuestión.

La tercera crítica fue discutida anteriormente y solo vale la pena mencionar aquí que el problema no es intrínseco del meta-análisis. Cualquier revisión de literatura tradicional está también sujeta a los vicios de publicación. El último punto se refiere a un problema real que se puede reducir si, para cada estudio en donde se puede calcular más de un índice se calcula un índice promedio.

El mayor problema del meta-análisis es que muchos investigadores no incluyen toda la información necesaria para hacer los estimados. Esta es un área que esperamos mejore en el futuro a medida que los editores de revistas tomen conciencia y exijan la inclusión de toda la información relevante. Bajo estas circunstancias, el meta-análisis será claramente una mejor alternativa que las revisiones de literatura tradicionales para la síntesis de estudios en

donde los resultados se puedan expresar en forma cuantitativa.

REFERENCIAS

Campbell, D.T. y J.S. Stanley (1963). *Experimental and Quasi-experimental Designs in Education*. Chicago: Rand McNally.

Carlberg, C.G., D.W. Johnson, R. Johnson, G. Maruyama, K. Kavale, C.C. Kulik, R.S. Lysakowski, S.W. Pflaum y H.J. Walberg (1984). Meta-analysis in education: How it has been used? *Educational Researcher*, 13(8), 6-15.

Cochran, W.G. (1937). Problems arising in the analysis of a series of similar experiments. *Journal of the Royal Statistical Society (suplemento)*, 4, 102-118.

Cooper, H.M. y R. Rosenthal (1980). Statistical vs. traditional procedures for summarizing research findings. *Psychological Bulletin*, 87, 442-449.

Fisher, R.A. (1932). *Statistical methods for research workers* (4ta. ed.). London: Oliver and Boyd.

Gianconia, R.M. y L.V. Hedges (1982). Identifying features of effective open education. *Review of Educational Research*, 52(4), 579-602.

Glass, G.V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3-8.

Glass, G.V., B. McGaw y M.L. Smith (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Guskey, T.R. y S.L. Gates (1986). Synthesis of research on the effects of mastery learning in elementary and secondary classrooms. *Educational Leadership*, 43(8), 73-80.

Hedges, L.V. (1984). Estimation of effect size under nonrandom sampling: The effects of censoring studies yielding statistically insignificant mean differences. *Journal of Educational Statistics*. 9(1), 61-85.

Hedges, L.V. (1985). Issues in meta-analysis. *Review of Research in Education*, 13, 353-396.

Hedges, L.V., J.A. Shymansky y G. Woodworth (1989). *Modern methods of meta-analysis*. Washington, DC: National Science Teachers Association.

Hedges, L.V. e I. Olkin (1985). *Statistical methods for meta-analysis*. Orlando, Fla.: Academic Press.

Kulik, J.A., C.C. Kulik y P.A. Cohen (1980). Effectiveness of computer-based college teaching: A meta-analysis of findings. *Review of Educational Research*, 50, 525-544.

Pearson, K. (1933). On a method for determining whether a sample of given size n supposed to be drawn from a parent population having known probability integral has probability drawn at random. *Biometrika*, 25, 379-410.

Slavin, R.E. (1984). Meta-analysis in education: How it has been used? *Educational Researcher*, 13, 6-15.

Slavin, R.E. (1984). A rejoinder to Carlberg et. al. *Educational Researcher*, 13(8), 24-27.

Slavin, R.E. (1986). Best-evidence synthesis: An alternative to meta-analytical and traditional research. *Educational Researcher*, 15(9), 5-11.

Slavin, R.E. (1987). Best-evidence synthesis: Why less is more. *Educational Researcher*, 16(4), 15-16.

Tippet, L.C.H. (1931). *The method of statistics*. London: Williams and Norgate.

Walberg, H.J. (1984). Meta-analysis in education: A reply to Slavin. *Educational Researcher*, 13(8), 16-23.

Wolf, F.M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Beverly Hills, CA: Sage.

Yates, F. y W.G. Cochran (1938). The analysis of groups of experiments. *Journal of Agricultural Science*, 28, 556-580.