

PRINCIPIOS DEL PARADIGMA CUANTITATIVO EN LA INVESTIGACIÓN EDUCATIVA

Walter Salazar Rojas
Escuela de Educación Física y Deportes
Universidad de Costa Rica, San José, Costa Rica
E-mail: wsalazar@cariari.ucr.ac.cr

Resumen

Salazar Rojas, W. (2002)- Principios del paradigma cuantitativo en la investigación educativa. **Revista de Ciencias del Ejercicio y la Salud**, 2 (1), 61-71. Este artículo discute las bases históricas y filosóficas del paradigma cuantitativo en la investigación educativa. Se analiza el concepto numérico de la realidad en el marco de la filosofía griega tomando como referencia a Pitágoras y Platón. Posteriormente, se discute el concepto de ciencia e investigación científica. Se define la variable, la taxonomía y los niveles de medición. Siguiendo los lineamientos de Campbell y Stanley se analizan los conceptos del diseño experimental, de validez interna y externa, las amenazas que diversos factores le plantean a las mismas, y cuáles son los elementos de control que el método científico y los diseños de investigación tienen para contrarrestar estas amenazas: aleatorización, pretest y grupo control. Finalmente se presenta uno de los avances más relevantes de la metodología cuantitativa como es el meta-análisis, técnica para resumir y analizar resultados de grupos de investigaciones. Se enumeran sus características y los pasos que hay que realizar para completar un meta-análisis. **PALABRAS CLAVES:** metodología cuantitativa, diseño experimental, validez interna, validez externa, variable, nivel de medición, meta-análisis.

*Mide lo mesurable y trata de hacer
medible lo que todavía no lo es.*
Galileo Galilei

PRINCIPIOS HISTORICOS Y FILOSOFICOS

Mencionan Campbell y Stanley (1973) que en el desarrollo de la metodología de la investigación, dos libros que hicieron un aporte significativo fueron el de McCalll (1923) "How to experiment in education" (Cómo experimentar en educación) y dos años después el de Fisher (1925) "Statistical methods for research workers" (Métodos estadísticos para investigadores). El primero de ellos centrado en los procedimientos para obtener datos correctos y adecuados que luego iban a ser objeto de los análisis estadísticos correspondientes, y el segundo desarrollando las técnicas estadísticas inferenciales del análisis de varianza.

Sobre todo el libro de Fisher vino a sentar las bases de lo que luego en las ciencias sociales se ha conocido como el paradigma cuantitativo en la investigación y el cual se ha llegado a pensar que tiene una historia relativamente reciente, pero que en realidad sus orígenes pueden encontrarse en el mismo desarrollo de la civilización occidental en la Grecia clásica.

Hacia el año 580 A.C., nació en Samos el fundador de la matemática griega, Pitágoras, el cual enseñaba que la naturaleza está regida por números, es más, estos son el lenguaje de la misma (Bronowski, 1979). Dos de sus aportes más conocidos tienen que ver con la relación entre la matemática y la armonía musical, y la prueba del teorema de

Pitágoras, que relaciona la geometría con la aritmética.

Pitágoras descubrió que una cuerda tensa al vibrar produce una nota grave. Si luego se divide la cuerda en segmentos iguales: dos, tres o cuatro segmentos, y se pone el nodo o punto fijo de la cuerda en uno de estos puntos divisorios y se pone a vibrar la cuerda, se producen sonidos armónicos. Si el nodo no está en uno de esos puntos, los sonidos no son armónicos. Este descubrimiento de la relación entre los acordes musicales y la escala matemática los llevó a inferir que los números constituyen los elementos del universo: “Los principios de los números son los elementos de todos los seres” (Guedj, 1998).

Aparte de esta relación entre los sonidos y las matemáticas, Pitágoras también logró demostrar otra relación entre la geometría y los números. Se sabía ya en su tiempo de la existencia del teorema que lleva su nombre, el cual indica que en un triángulo rectángulo el cuadrado de la hipotenusa es igual a la suma de los cuadrados de los dos lados adyacentes. El fue el primero en demostrar matemáticamente que esa equivalencia era correcta a través de una prueba de naturaleza geométrica.

Posteriormente, las doctrinas pitagóricas fueron incorporadas a la filosofía griega por otro pensador– Sócrates – del cual tenemos evidencia en los escritos de su discípulo, Platón.

En el libro VII de La República, encontramos un diálogo entre Sócrates y Glaucón, donde el primero le pregunta al segundo: “¿Cuál es la ciencia que atrae el al alma de lo que deviene hacia lo que es ?” Sócrates responde que se trata de una ciencia que se extiende a todo, que sirve para todas las artes, para todas las operaciones del espíritu y para todas las ciencias: la ciencia de los números (Guedj, 1998):

S- ...debemos acudir a una ciencia que se aplique a todas.

G- ¿A cuál?

S- Pues a la que resulta tan común por el uso que hacen de ellas las artes, los discursos y las ciencias. Con esta habrá que contar entre las primeras.

G- Sigo formulándote la pregunta- dijo

S- Es la que enseña lo que es uno, dos y tres...En resumen trata del número y del cálculo. Porque ¿no es verdad que todo arte y toda ciencia se ven obligadas a participar de ella?

G- Desde luego – asintió

(...)

S- Me parece que es ella (la ciencia de los números) una de las que buscamos como conducente por naturaleza al conocimiento puro...

(...)

S- No cabe duda de que la ciencia del cálculo y la aritmética se ocupan por entero del número.

G- Desde luego

S- Una y otra, pues, parece que conducen hacia la verdad.

G- Son perfectamente aptas para ello.

S- He aquí, según parece, que tenemos ya dos de las ciencias que buscamos. Ambas son necesarias de todo punto al guerrero y al filósofo, al primero para la mejor ordenación de los ejércitos, y al segundo para que emerja del mundo perecedero hacia la esencia de las cosas, si es que se precia de hombre calculador.

G- Así es- afirmó

(...)

G- ¿Qué intención te guía entonces? –preguntó

S- ...porque es lo cierto que esa ciencia conduce al alma hacia lo alto y la obliga a razonar sobre los números...

G – Verdad es lo que afirmas

(...)

S – Queda, pues, adoptada como la primera de las ciencias.

(Platón, 1973)

Esta conceptualización de que el hombre debe obedecer el “mandato de la recta razón” , la cual tiene un fundamento aritmético y geométrico, ha permeado el desarrollo de la ciencia y civilización occidental, y muy recientemente, el filósofo francés Michel Serres (2002) ha manifestado su apreciación por el aporte griego al declarar:

“Salida de Grecia y difundida, que yo sepa, sin imperialismo y derramamiento de sangre, la geometría enseñó por primera vez a los hombres verdades universales. Qué maravilla! A pesar de las diferencias que se dice tienen las culturas, a pesar de lo resistente de la barrera de las lenguas, a pesar de la mala fe que pueda presentar un sujeto..., nadie puede negar que tres más tres son seis y que el cuadrado de la hipotenusa es igual a la suma de los cuadrados de los otros dos lados, en el caso de un triángulo rectángulo. De esta manera, desde Pitágoras y el sexto siglo antes de Cristo, sin armas ni dinero, sin dogma ni poder, se difundió la ciencia rigurosa... De cierta manera, la matemática funciona en todo lado como lengua universal, y eso tres veces: se aplica a todas las cosas del mundo, pone de acuerdo a todos los hombres por las necesidades de su demostración, certifica que el mensaje emitido hace miles de años sea el mismo que transmitimos a las generaciones futuras: por eso a menudo me pregunto si no existe otra historia que la de las matemáticas”

LA CIENCIA Y SU METODO

El fin de la ciencia

De acuerdo a Braithwaite (1996,p.1), el fin de la ciencia “consiste en establecer leyes generales sobre el comportamiento de eventos empíricos u objetos en los que la ciencia en cuestión está interesada, para así permitimos conectar nuestro conocimiento de eventos conocidos por separado y hacer predicciones confiables de eventos aún desconocidos”.

Tenemos además una que la ciencia utiliza un procedimiento por medio del cual accede al conocimiento, y el cual es la investigación científica, definida como “una investigación sistemática, controlada, empírica, amoral, pública y crítica de fenómenos naturales. Se guía por la teoría y las hipótesis sobre las presuntas relaciones

entre esos fenómenos” (Kerlinger y Lee, 2001, p.13).

Características de la investigación científica

En enfoque que emplea la ciencia para resolver problemas tiene cinco características de acuerdo a Tuckman (1978):

1. Sistemática. Luego de definir un problema, donde se identifican las variables, se procede a diseñar la investigación que va a examinar la relación entre las variables. Se procede luego a recolectar los datos, que luego de analizarse van a determinar la evaluación del problema y las hipótesis.
2. Lógica. El examen de los procedimientos usados en el proceso le permite al investigador evaluar las conclusiones derivadas.
3. Empírica. Los datos recolectados son los que permiten alcanzar conclusiones.
4. Reductiva. La investigación utiliza muchos eventos individuales o datos, y los usa para establecer relaciones más generales.
5. Repetible. El proceso de investigación se registra minuciosamente de forma tal que otros investigadores puedan repetir los mismos pasos para probar las conclusiones obtenidas.

La variable: definición, taxonomía y medición

Una variable es algo que varía (Kerlinger, 1986). Es un símbolo que representa un concepto y al que se le pueden asignar numerales o categorías.

Las variables o conceptos son abstracciones formadas a partir de una generalización de casos particulares (Shapiro y Hardy, 1985). Pueden definirse a través de dos procedimientos. Una definición conceptual es aquella que utiliza otros conceptos o variables para definir el concepto o variable que es objeto de estudio. Además están las definiciones operacionales que son aquellas donde se especifica cómo se va a hacer la medición del concepto.

La clasificación más utilizada de las variables es la que las divide en variables independientes y dependientes. La variable independiente en la investigación experimental es aquella que es el antecedente o causa, mientras que la dependiente como su nombre lo indica es la que es el consecuente o el efecto. En la investigación no experimental esta clasificación no tiene tanta aplicación, aunque puede utilizarse una clasificación en variables predictoras y variables criterio.

En la medición de las variables se pueden distinguir cuatro niveles: nominal, ordinal, intervalos y razones.

La escala nominal es aquella donde el sujeto o objeto se clasifica en una categoría específica. Estas categorías deben llenar dos requisitos: (a) ser exhaustivas, es decir, un objeto no puede quedarse sin ser clasificado en una categoría, y (b) ser mutuamente excluyentes. Un sujeto no puede ser clasificado en más de una categoría. En la medición nominal se realiza un conteo de las frecuencias en cada categoría y se utiliza estadística no paramétrica para su análisis.

La escala ordinal se utiliza para ordenar los sujetos o objetos de acuerdo a alguna característica o variable. Este ordenamiento se hace de menor a mayor o viceversa. La diferencia de un sujeto a otro no es comparable con otras diferencias entre sujetos. También utiliza estadística no paramétrica en el análisis.

En la escala de intervalos se le asignan numerales a los sujetos de acuerdo a la cantidad de la variable que posean. El

cero en la escala de intervalos es relativo. Las variables medidas en una escala de intervalos son variables continuas, por lo que se puede aplicar la estadística paramétrica como el promedio, desviación estándar, análisis de varianza, regresión, etc.

Las escalas de razones representan el nivel más alto de medición, y son aquellas que tienen un cero absoluto, es decir, el cero representa la ausencia de la cualidad que se está midiendo. Estas escalas son también susceptibles de análisis con la estadística paramétrica.

Diseños de investigación y el control de la varianza

De acuerdo a Kerlinger y Lee (2001), el plan y la estructura para contestar las preguntas de la investigación se denomina el diseño de investigación, el cual tiene dos objetivos principales: (a) contestar el o los problemas, y (b) controlar la varianza.

Este control de la varianza se refiere a las tres fuentes principales de varianza: la verdadera, la de error y la extraña.

La varianza sistemática es la que es producida por los tratamientos. En una investigación típicamente lo que se busca es que la varianza verdadera sea maximizada. Para esto el investigador debe buscar aplicar tratamientos intensivos, y que las diferencias entre los tratamientos sean grandes.

La varianza de error es la varianza entre los puntajes de los sujetos. Esta varianza puede disminuirse de dos maneras: (a) utilizando muestras que sean más homogéneas entre sí, y (b) utilizando instrumentos de medición que tengan una alta confiabilidad.

La varianza extraña es la que es producida por variables potencialmente independientes que van a afectar a la variable dependiente, pero que no han sido incluidas en el estudio. Para tratar con esta varianza extraña hay varias estrategias que pueden seguirse: (a) utilizar la asignación aleatoria de sujetos a los grupos experimentales (en

los diseños entre-sujetos), y la asignación aleatoria de los órdenes de tratamientos (en los diseños intra-sujetos), (b) convertir la variable extraña potencial en una “constante”, al escoger sujetos en base a un nivel de la variable extraña, (c) incluir la variable extraña como una variable más independiente dentro del estudio, y de esta manera estudiar el efecto directo que esta puede producir sobre la variable dependiente.

Los diseños de investigación siguiendo la formulación clásica hecha por Campbell y Stanley (1973) están compuestos de los siguientes elementos y su simbología:

1. Tratamiento: **X**
2. Medición: **O**
3. Aleatorización: **R**

La combinación de estos elementos permite la elaboración de diseños que varían en su nivel de complejidad, empezando por los diseños pre-experimentales, luego los diseños cuasi-experimentales, y terminando en los diseños experimentales.

Es el grado de control experimental, determinado por los elementos presentes en el mismo el que va a determinar el tipo de diseño que se establezca.

Existen tres elementos dentro del diseño que van a influir sobre el nivel de control de la varianza extraña y de error: El primero es la aleatorización la cual tiene que ver con (a) la selección aleatoria de la muestra a partir de la población, (b) la asignación aleatoria de los sujetos a los grupos, (c) la asignación aleatoria de los tratamientos a los grupos, y (d) la asignación aleatoria del orden de tratamientos.

El segundo elemento es la presencia de un grupo control, el cual es el que no recibe ninguno de los niveles de tratamiento que se están examinando. Este grupo tiene la función de detectar la presencia de variables extrañas que puedan estar afectando a la variable dependiente y confundiendo el efecto de la variable independiente.

El tercer elemento es la realización de un pretest, el cual va a determinar la equivalencia de los grupos en el momento inicial, antes de que el tratamiento sea aplicado a los grupos. Este pretest también recibe el nombre de línea base.

Si un diseño de investigación no tiene aleatorización, pretest ni grupo control, es un diseño de lo más deficiente en su control, y se clasifica dentro de los diseños pre-experimentales de Campbell y Stanley (1973). Por el contrario, los diseños que incluyen los tres elementos de control, se clasifican como diseños experimentales que es donde se puede cumplir con el principio enunciado por Kerlinger (1986) de maximizar la varianza verdadera, minimizar la varianza de error y controlar la varianza extraña.

Validez interna y externa

Campbell y Stanley (1973) establecieron la importancia de las dos formas de validez de un estudio, que posteriormente fueron ampliadas por Cook y Campbell (1979). La validez interna se refiere a la certeza que tiene el investigador de que los cambios observados en la variable dependiente se deban al efecto de la variante independiente, mientras que la validez externa se relaciona con la posibilidad de generalización de los resultados: ¿a qué sujetos, situaciones o tratamientos pueden generalizarse los resultados?.

Si se quiere obtener control de la validez interna, el investigador debe eliminar la posibilidad de que haya hipótesis rivales que puedan explicar los resultados obtenidos, para esto debe emplear procedimientos para aumentar el control experimental, los cuales por lo general tienen un efecto negativo sobre la validez externa. Así que normalmente el investigador debe escoger entre maximizar la validez interna y disminuir la externa o viceversa, si quiere tener la máxima validez externa debe sacrificar en buena medida los controles de

validez interna para tener las mayores posibilidades de generalización.

Campbell y Stanley (1973) identificaron ocho amenazas a la validez interna de un estudio:

1. Historia: los eventos que ocurren durante el experimento y que no son parte de el tratamiento.
2. Maduración: los procesos internos de los sujetos que operan como resultado del paso del tiempo.
3. Medición: los efectos de un test en las aplicaciones siguientes del mismo test.
4. Instrumentación: cambios en la calibración de un instrumento. Incluye la falta de acuerdo entre administradores del test.
5. Regresión estadística: el hecho de que grupos de sujetos escogidos porque tenían puntajes extremos en el pretest, en mediciones posteriores sus puntajes ya no son tan extremos.
6. Sesgo de selección: identificación de los grupos de comparación no por forma aleatoria.
7. Mortalidad experimental: pérdida de los sujetos de los grupos de comparación por razones no aleatorias.
8. Interacción entre maduración y selección: donde el pasaje del tiempo puede afectar más a un grupo que a otro.

Otra amenaza a la validez interna fue identificada por Rosenthal (1966) y la denominó la “expectativa” que se refiere a que el investigador puede esperar o anticipar que ciertos sujetos tendrán una mejor (o peor) actuación o rendimiento que otros. Típicamente este efecto es inconsciente de parte de los investigadores. Para controlarlo se utilizan los diseños de doble ciego.

Igualmente, Campbell y Stanley (1973) identificaron cuatro amenazas a la validez externa de una investigación:

1. Efectos reactivos o interactivos de la medición: el hecho de que el pretest puede hacer al sujeto más consciente o sensitivo del tratamiento que va a recibir. El resultado es de que el tratamiento no es tan efectivo sin la presencia del pretest.
2. Interacción entre el sesgo de selección y el tratamiento experimental: cuando un grupo se selecciona en base a una característica, puede resultar que el tratamiento sólo funcione en grupos que posean esa característica.
3. Efectos reactivos de las situaciones experimentales: el hecho de que tratamientos que son efectivos en situaciones muy restrictivas (laboratorios) pueden no serlo en situaciones que no lo son.
4. Interferencia de tratamientos múltiples: cuando los sujetos reciben más de un tratamiento, puede ser que los efectos de un tratamiento previo influyan en uno posterior.

AVANCES RECIENTES EN EL PARADIGMA CUANTITATIVO: EL META-ANÁLISIS

Con el continuo incremento de estudios en el área de la educación, también aumenta la necesidad de realizar buenas revisiones de literatura. Hasta finales de los años 70, las revisiones narrativas eran la única fuente para resumir la información disponible. Esta situación empezó a cambiar con la publicación por Glass (1976, 1977) sobre una metodología denominada “meta-análisis”, la cual es básicamente una técnica cuantitativa de resumir resultados de investigaciones de muchos estudios.

Los propósitos de la información que se presenta a continuación son dos: (a) presentar la información histórica del desarrollo del meta-análisis dentro de las ciencias de la estadística, y (b) describir los

pasos comunes que se siguen al desarrollar un meta-análisis.

a. El desarrollo de las técnicas meta-analíticas.

Cuando Glass (1976, 1977) desarrolló el término “meta-análisis”, mucha gente pensó que la metodología involucrada era también algo nuevo. Sin embargo, la teoría estadística subyacente no era nueva, habiendo sido desarrollada durante la década de los 30. Como lo discuten Hedges y Olkin (1982), se habían desarrollado dos enfoques distintos para combinar los resultados de investigaciones. Ambos enfoques se habían derivado de una metodología que originalmente se desarrolló para resumir y sintetizar datos de la literatura en agricultura. Un enfoque tenía que ver con combinar los valores de las probabilidades de los tests estadísticos, mientras que el otro combinaba los índices de las magnitudes de los tratamientos.

El enfoque de combinar los valores de las probabilidades fué desarrollado por Tippet (1931) y luego refinado por Fisher (1932) y Pearson (1933). En este enfoque, las probabilidades exactas de los tests de significancia (valores p) se combinan para probar el efecto global (Kerlinger, 1986, p.101). Una debilidad de este enfoque es que no puede brindar un estimado de la magnitud del efecto del tratamiento. El segundo enfoque, combina los estimados de la magnitud del efecto del tratamiento, y fue desarrollado por Cochran (Cochran, 1937; Yates y Cochran, 1938). En este enfoque, se obtiene un estimado del efecto promedio del tratamiento y de su variabilidad.

Más recientemente, se han desarrollado una serie de enfoques alternativos a partir de los originales. Bangert-Drowns (1986) clasifica cinco formas de métodos meta-analíticos: (1) meta-análisis Glassiano (Glass, 1976; Glass, McGaw y Smith, 1981); (2) meta-análisis de efectos del estudio (Bangert-Drowns, Kulik y Kulik, 1984); (3) método de probabilidades

combinadas (Cooper, 1979, 1982; Rosenthal, 1978); (4) combinación aproximada de datos con tests de homogeneidad (Hedges, 1982; Hedges y Olkin, 1985), y (5) combinación aproximada de datos con corrección del error de muestreo (Hunter, Schmidt y Jackson, 1983).

Antes de que Glass introdujera el concepto de meta-análisis, la revisión narrativa era la herramienta usada para integrar un cuerpo de estudios en un tópico particular. Un meta-análisis intenta extraer información relevante de un grupo de estudios. Sin embargo, hay diferencias marcadas en cuanto al grado de objetividad entre un revisor narrativo y un revisor meta-analítico.

Un revisor narrativo empieza examinando las debilidades de cada estudio y sus fortalezas. A pesar de que no hay reglas específicas que un revisor narrativo pueda seguir para evaluar la “calidad” de un estudio, un enfoque popular es utilizar la categorización de Campbell y Stanley (1973), la cual clasifica los diseños en pre-experimentales, cuasi-experimentales y experimentales. La mayor parte del tiempo, los estudios se excluyen de análisis posteriores basados en decisiones subjetivas. Al mismo tiempo, no hay consenso entre los revisores narrativos acerca de cómo “pesar” la evidencia de los estudios que ellos consideran buenos.

Un revisor meta-analítico por el contrario, sigue un proceso objetivo y público, donde hay una secuencia de pasos claramente definida de forma tal que cualquier persona puede repetirlos. Toda la información relevante acerca de los diferentes pasos se incluye en el reporte final. Ejemplos de estos pasos son las decisiones tomadas acerca de: (a) codificación de las variables (independientes y dependientes), (b) cálculo del tamaño del efecto, (c) la corrección del tamaño de la muestra y de sesgo, (d) aplicar un factor de peso, y (e) la escogencia del análisis estadístico apropiado.

La falta de consenso acerca de asuntos metodológicos en las revisiones narrativas de debe a que el proceso se ve más como un “arte” (donde se requiere una destreza de parte del revisor) que una ciencia. Un artista no sigue unas reglas especificadas, pero usa su estilo particular (una característica propia y subjetiva) para producir una revisión. Un revisor meta-analítico por el contrario, sigue las reglas del método científico, donde la objetividad y la verificación de los pasos metodológicos son aspectos claves. Como lo afirma Kerlinger (1986, p.6), “los científicos no aceptan las afirmaciones como ciertas, a pesar de que la evidencia en principio parezca promisoría. Ellos insisten en probarlos. Ellos también insisten en que los procedimientos de prueba estén abiertos a la inspección pública”.

De acuerdo a Halliwell y Gauvin (1982) hay tres ventajas en utilizar un meta-análisis en comparación a las narraciones narrativas. La primera es de que por usar un sistema objetivo de codificación de las diferentes características y resultados de los estudios, el revisor meta-analítico tiene menos probabilidad de combinar incorrectamente los resultados de varios estudios en comparación al revisor narrativo. Segunda, de que a través del uso de técnicas estadísticas (estadística descriptiva e inferencial) para resumir y analizar la información, el revisor tendrá una visión más precisa del patrón de resultados y de cómo las características de los estudios (variables independientes) van a influir en el resultado final. Y finalmente, de que a través de una búsqueda extensiva de la literatura, y a través de una codificación objetiva de las características y resultados, el revisor puede establecer una bibliografía práctica y un banco de datos del tópico que se revisa.

Debido a que el desarrollo de las técnicas meta-analíticas es relativamente joven, todavía no hay una única forma aceptada para realizar un meta-análisis. En el área de la educación y de las ciencias sociales en general, los dos métodos más frecuentes son el de Glass y el de Hedges y

Olkin, o una combinación de ambos. Ambos derivan el tamaño del efecto (una medida de la magnitud del efecto del tratamiento) como la unidad central de análisis. Un tamaño del efecto es una medida de la magnitud del efecto de un tratamiento particular sobre una variable dependiente, el cual toma en cuenta la variabilidad dentro del estudio del cual es calculado. Los tamaños del efecto pueden compararse entre estudios porque han sido previamente normalizados.

b. Pasos para realizar un meta-análisis.

Glass y otros (1981) y Thomas y French (1986) especificaron los procedimientos a seguir cuando se realiza un meta-análisis. Tales procedimientos deben describirse explícitamente en el documento escrito de forma tal que cualquier persona puede repetirlos. Esto es análogo a la sección de metodología en una investigación experimental, donde los pasos se describen detalladamente para que puedan ser seguidos en otro estudio. Un meta-análisis genérico incluiría los siguientes pasos:

1. Identificación del problema
2. Búsqueda de la literatura
3. Codificación de las variables moderadoras
4. Cálculo del tamaño del efecto
5. Corrección por sesgo del tamaño de la muestra
6. Análisis estadístico
7. Publicación

Identificación del problema

Las áreas aptas para una revisión son aquellas que tienen una cantidad suficiente de estudios (al menos 20 estudios) pero que por lo menos a primera vista no tienen un consenso en cuanto a los resultados globales. Como una síntesis cuantitativa, un meta-análisis puede realizarse para proveer respuestas a preguntas específicas, por ejemplo ¿La enseñanza individualizada produce mejores resultados que la enseñanza

en grupo?. Las preguntas que típicamente han sido objeto de revisiones narrativas y que han producido resultados conflictivos pueden analizarse a través de este método cuantitativo. El meta-análisis también puede usarse para determinar la magnitud del efecto de un tratamiento particular. Las revisiones narrativas por su naturaleza no pueden dar al investigador interesado tal información.

Búsqueda de la literatura

La búsqueda de literatura debe ser extensiva. Debe buscarse toda la información disponible en (a) bases de datos en computadora, (b) búsquedas manuales de artículos, tesis, disertaciones y resúmenes; (c) buscando en las listas de las referencias de cada artículo obtenido, y (d) usando los índices de citas (Science Citation Index) para buscar artículos y autores claves. Las diferentes fuentes que se utilizaron deben especificarse en el documento. Es posible también codificar la fuente de los estudios usados en el meta-análisis y por lo tanto utilizarlos como variables moderadoras en el análisis estadístico. Tal análisis puede permitir que se obtengan conclusiones en relación a si el tipo de publicación tiene un efecto en la variable dependiente.

Codificación de las variables moderadoras

Las variables moderadoras que puedan en potencia modificar el efecto sobre la variable dependiente se codifican ya sea categórica o cuantitativamente. Tales variables se derivan típicamente de la literatura y a través del conocimiento de preguntas de interés. Ejemplos de variables moderadoras pueden incluir (a) característica del sujeto (edad, sexo, características demográficas, socioeconómicas, nivel educativo, etc.), (b) características del estudio (validez interna del estudio, tamaño de la muestra, tipo de asignación de los sujetos, tipo de diseño, etc), y (c) tipo de variable dependiente (test utilizado).

Algo importante en muchos estudios es determinar posibles mecanismos que puedan explicar porqué un tratamiento tiene efectos específicos. Por eso es relevante que a la hora de codificar variables moderadoras, debe incluirse aquellas características que puedan tener relación con posibles mecanismos.

Cálculo del tamaño del efecto

Hay dos procedimientos para el cálculo del tamaño del efecto. Glass (1977) lo calcula obteniendo la diferencia entre el promedio del grupo experimental y el promedio del grupo control y luego dividiendo esta diferencia por la desviación estándar del grupo control. Hedges y Olkin (1985) lo calculan utilizando la misma diferencia entre los promedios del grupo experimental y del control y luego dividiendo esta diferencia por la desviación estándar combinada de los dos grupos.

Corrección por sesgo del tamaño de la muestra

Los tamaños del efecto están sesgados positivamente en las muestras pequeñas. Por lo tanto, los tamaños del efecto deben corregirse en su sesgo antes de proseguir con el análisis. Si no se hace esta corrección, el tamaño del efecto promedio puede sobrestimar el efecto correcto.

Debido a que el tamaño de la muestra de un estudio particular puede influir en el tamaño del efecto, los estudios con más sujetos deben tener más peso que los estudios con muestras pequeñas. Para controlar por este efecto, se le aplica un factor de pesaje al tamaño del efecto obtenido. Los tamaños de efecto obtenidos de estudios con muestras grandes se pesan con factores mayores que los que se obtienen de muestras pequeñas.

Análisis estadístico

Se pueden utilizar las técnicas estadísticas inferenciales (análisis de varianza, r de Pearson, análisis de regresión simple o múltiple) para analizar la relación de las características o variables moderadoras con el tamaño del efecto (variable dependiente) (Glass, 1977). Hedges y Olkin también usan técnicas específicas para probar el tamaño del efecto. Primero, se realiza un test de homogeneidad para examinar el grupo de tamaños del efecto. Si el test no es significativo, es decir, si los tamaños del efecto son homogéneos, no se realiza ningún análisis de las variables moderadoras. Si el test es significativo, indica que hay una variabilidad confiable en el grupo de tamaños del efecto, y procede entonces una búsqueda de variables moderadoras. Para hacer esto se realiza una técnica de regresión (Hedges y Olkin, 1983) donde se examina el efecto de las variables predictoras en el tamaño del efecto.

Publicación

El paso final es reportar los resultados del meta-análisis en una revista científica o en un libro. En este momento, la revisión por los pares académicos toma lugar. Con un meta-análisis, la revisión de los pares debe ser más exigente que con un artículo sencillo, ya que los resultados van a tener una mayor influencia en la comunidad científica, y pueden dirigir la investigación futura en un área específica, tanto para desarrollarla como para detenerla.

REFERENCIAS

- Bangert-Drowns, R.L. (1986). Review of developments in meta-analytic method. *Psychological Bulletin*, 99, 389-399.
- Bangert-Drowns, R.L.; Kulik, J.A., y Kulik, C.L. (1984). *The influence of study features on outcomes of educational research*. Estudio presentado en la reunión anual #92 del American Psychological Association en Toronto.
- Braiwthwaite, R. (1996). *Scientific explanation*. Hendon, VA: Books International.
- Bronowski, J. (1979). *El ascenso del hombre*. Bogotá: Fondo Educativo Interamericano.
- Campbell, D.T. y Stanley, J.C. (1973). *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires: Amorrortu.
- Cochran, W.G. (1937). Problems arising in the analysis of a series of similar experiments. *Journal of the Royal Statistical Society*, 4, 102-118.
- Cook, T.D. y Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
- Cooper, H.M. (1979). Statistically combining independent studies: a meta-analysis of sex differences in conformity research. *Journal of Personality and Social Psychology*, 37, 131-146.
- Cooper, H.M. (1982). Scientific guidelines for conducting integrative research. *Review of Educational Research*, 52, 291-302.
- Fisher, R.A. (1925). *Statistical methods for research workers*. London: Oliver y Boyd.
- Fisher, R.A. (1932). *Statistical methods for research workers (4th ed.)*. London; Oliver y Boyd.
- Glass, G.V. (1976). Primary, secondary, and meta-analysis research. *Educational Researcher*, 5, 3-8.
- Glass, G.V. (1977). Integrating findings; The meta-analysis of research. *Review of Research in Education*, 5, 351-379.
- Glass, G.V.; McGaw, B., y Smith, M. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Guedj, D. (1998). *El imperio de las cifras y los números*. Barcelona: Ediciones B, S.A.
- Halliwell, W. y Gauvin, L. (1982). Integrating and interpreting research findings: a challenge to sport psychologists. En J.T. Partington, T. Orlick y J.H. Salmela (Eds), *Sport in perspective*, Ottawa: Coaching Association of Canada.
- Hedges, L.V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, 92, 490-499.
- Hedges, L.V. y Olkin, I. (1982). Analyses, reanalyses, and meta-analysis. *Contemporary Education Review*, 1, 157-165.
- Hedges, L.V. y Olkin, I. (1983). Regression models in research synthesis. *American Statistician*, 37, 137-140.
- Hedges, L.V. y Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.
- Hunter, J.E.; Schmidt, F.L. y Jackson, G.B. (1983). *Meta-analysis: cumulating research findings across studies*. Beverly Hills, CA: Sage.
- Kerlinger, F.N. (1986). *Foundations of behavioral research*. New York: Holt, Rinehart y Winston.
- Kerlinger, F.N. y Lee, H.B. (2001). *Investigación del comportamiento: métodos de investigación en ciencias sociales*. México: McGraw-Hill.

- McCall, W.A. (1923). *How to experiment in education*. Nueva York: Macmillan.
- Pearson, K. (1933). On a method of determining whether a sample size n supposed to have been drawn from a parent population having a known probability has probably been drawn at random. *Biometrika*, 25, 379-410.
- Platón (1973) *Obras completas*. Madrid: Aguilar.
- Rosenthal, R. (1966). *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts.
- Rosenthal, R. (1978). Combining results of independent studies. *Psychological Bulletin*, 85, 185-193.
- Serres, M. (2002). *Lo universal contra un particularismo llamado mundialización*. San José, periódico La Nación, 29 de setiembre del 2002, suplemento Ancora, pp. 1-3.
- Shapiro, G.M. y Hardy, H. (1985). *A research primer for the social behavioral sciences*. Orlando: Academic Press.
- Thomas, J.R., y French, K.E. (1986). The use of meta-analysis in exercise and sport: a tutorial. *Research Quarterly for Exercise and Sport*, 57, 196-204.
- Tippet, L.H.C. (1931). *The methods of statistics*. London: William y Norgate.
- Tuckman, B.W. (1978). *Conducting Educational Research*. New York: Harcourt, Brace y Jovanovich.
- Yates, F., y Cochran, W.G. (1938). The analysis of groups of experiments. *Journal of Agricultural Science*, 28, 556-580.