

DIFICULTADES Y ERRORES METODOLOGICOS EN LA INVESTIGACION EDUCATIVA

por Dra. M.^a PILAR COLAS BRAVO

Las investigaciones educativas han sido muy numerosas en los últimos años, y de forma especial en los países anglosajones. Este acopio de información no ha llevado aparejado una evolución de la teoría educativa y de la práctica escolar. La disparidad de resultados y la dificultad de aunarlos para la creación de postulados ha potenciado la reflexión sobre aspectos metodológicos de su desarrollo, resaltándose dificultades y errores más habituales, a la vez que se proponen modelos alternativos que pretenden subsanar dichos fallos y adecuarse mejor a las peculiaridades de las situaciones educativas.

Uno de los mayores handicaps de la investigación educativa deriva del carácter cualitativo de los fenómenos que se estudian. Los acontecimientos y acciones externos no son siempre los más adecuados para comprender o explicar un hecho educativo, ya que a veces por debajo de esas manifestaciones externas laten intereses, valores, ideologías, etc., que son los factores realmente importantes. Por otro lado el carácter inmaterial de estas expresiones hace extremadamente difícil su observación, medición y cuantificación. Esta consideración ha llevado a enfoques metodológicos más cualitativos, que actualmente coexisten junto a enfoques más tradicionales.

Nuestra reflexión no discurrirá en torno a los presupuestos epistemológicos que sustentan las corrientes de investigación educativa, ni tampoco en torno a la metodología de los estudios cualitativos, ya que, hoy por hoy, no representan la generalidad de las investigaciones educativas. Nuestro propósito es hacer un análisis de las principales dificultades y errores metodológicos que presentan las investigaciones educativas hoy día publicadas, que en su gran mayoría responden a presupuestos epistemológicos y metodológicos positivistas.

Nos referiremos siguiendo la propuesta de Dunkin y Biddle¹ a cuatro aspectos fundamentales; *problemas de medida, muestreo, diseño e interpretación estadística*.

El problema de medida quizá sea uno de los de mayor complejidad, incluyendo numerosos aspectos interrelacionados. Cooms, Dawes y Tuersky² señalan cuatro

¹ DUNKIN, M. y BIDDLE, B. (1974): *The study of teaching* (New York, Holt, Rinehart and Winston, Inc.).

² COOMBS, C.; DAWES, R. y TUERSKY, A. (1970): *Mathematical Psychology: an elementary introduction* (Englewood Cliffs, New Jersey, Prentice-Hall), *op. cit.* por ALVIRA, F. y OTROS (1981): *Los dos métodos de las ciencias sociales* (Madrid, Centro de Investigaciones Sociológicas).

problemas básicos en la teoría de la medida: 1) problema de la representación (se da cuando se transforma una realidad a un sistema formal numérico), 2) problema de la unicidad o grado de libertad que se dispone para asignar los números a los objetos y 3) problema de significación —tipo de inferencia que se produce al utilizar una escala de medida particular—. Todos estos problemas llevan a cuestionar la validez de las medidas en las investigaciones educativas y en las ciencias sociales en general.

Nosotros concretaremos esta temática en tres apartados: a) *selección de la unidad de análisis*; b) *perioricidad de la medida* y c) *instrumentos de medida*.

a) El averiguar qué variables son representativas de los fenómenos educativos es uno de los objetivos y retos que tienen planteados las ciencias de la educación. La decisión del investigador de qué aspectos estudiar viene determinada, la mayoría de las veces, por la disponibilidad de teorías o datos previos sobre variables de enseñanza e instrumentos de recogida de datos. Las investigaciones y sus resultados van aportando mayor precisión en este terreno, remodelando o incorporando nuevas propuestas de variables. A modo de ejemplo citaremos las nuevas orientaciones de considerar variables contextuales que tradicionalmente se tenían olvidadas en los estudios educativos³.

Una cuestión unida a la anterior es la referida a la amplitud de los rasgos de las variables a estudiar. Pequeñas unidades pueden resultar no significativas en la interpretación de un fenómeno; aspectos más amplios pueden impedir una mínima objetividad y validez en la recogida de datos. La relevancia de estas unidades explicativas es una cuestión que queda todavía sin resolver. La realización de estudios con variables de distinto nivel de generalidad nos puede proporcionar información más precisa sobre su representatividad, y puede ser una vía de enorme interés para llegar a establecer unas orientaciones mínimas que garanticen resultados más fructíferos en futuros trabajos.

En un estudio⁴ realizado con variables de dos niveles de amplitud se confirmaba que la variable más amplia (métodos didácticos) resultaba más significativa que la variable más específica (conductas docentes). El tipo y nivel de la variable significativa variaba en función de la variable dependiente (objetivos educativos). En determinados objetivos únicamente la variable específica (conductas docentes) se manifestaba representativa, en cambio para otros objetivos la variable amplia (métodos docentes) era la única digna de consideración.

b) La perioricidad de la recogida de datos es otra cuestión enormemente criticada en las reflexiones metodológicas⁵. La mayoría de trabajos basan sus conclusio-

³ BERLINER, D. (1976): Impediments to the study of teacher effectiveness, *Journal of Teacher Education*, 27:1, pp. 1-13. BORICH, C. y FENTON, K. (eds.) (1977): *The appraisal of teaching: concepts and process* (Addison Wesley, Reading Mass).

⁴ COLAS, P. (1984): Análisis experimental de métodos y técnicas didácticos interaccionados con algunas características de los alumnos. Tesis Doctoral Inédita. Universidad de Sevilla.

⁵ BROPHY, J. (1979): Teacher Behavior and its effects, *Journal of Educational Psychology*, 71:6; pp. 733-750.

nes en observaciones de pocas horas o en períodos de tiempo relativamente cortos, no contando apenas con estudios longitudinales o de medidas amplias de tiempo. Este hecho condiciona la validez y fiabilidad de las conclusiones y la selección de las unidades de análisis. La escasa duración de la recogida de datos determina el tipo de variable a estudiar. La selección de tiempos reducidos se basa en la aceptación de dos supuestos: a) las variables o aspectos frecuentes son constantes y representativas de los fenómenos que se estudian, b) el período de investigación es representativo de la totalidad de la situación.

Algunos argumentos ponen en tela de juicio estos supuestos⁶. No siempre la frecuencia de conductas es indicativo de su importancia en relación con la variable que se trata. Ciertos sucesos esporádicos pueden ser más explicativos que los habituales. En cuanto a la estabilidad de variables ciertos autores⁷ han tratado de comprobar este hecho, recogiendo trabajos al respecto. Sus aportaciones tienden a presentar índices negativos de estabilidad, aumentando en variables de alta inferencia. Sin embargo estos autores coinciden en señalar que la estabilidad no es sinónimo de validez.

En este apartado también habría que hacer mención de la temporalidad con la que se recogen observaciones relativas a la variable dependiente. La información procedente de distintos períodos de tiempo puede ser útil a la hora de comprender la influencia de las situaciones educativas en el comportamiento de los alumnos. En este sentido el trabajo anteriormente mencionado (4) nos indica la diferencia de resultados que se obtienen en distintas mediciones, manifestándose una mayor efectividad del tratamiento educativo a más largo plazo

c) En el apartado de *instrumentación* dos aspectos son de interés considerar; contenido y forma. El contenido va íntimamente unido a las variables que se han de medir. Determinadas facetas son más fáciles de expresarse en medidas externas, mientras otras resultan de gran dificultad su medición, encontrándose simultáneamente problemas en su validación, al no contar con datos externos que las corroboren. De ahí que haya una tendencia generalizada a medir aspectos más externos, y a disponer de un mayor número de instrumentos para éstos que para aspectos menos tangibles.

En cuanto al formato del instrumento haremos un breve análisis de la problemática que encierran las técnicas de recogida de datos más usuales.

1. *Los Test Estandarizados*, a pesar de ser instrumentos sumamente fiables y poseer un contenido curricular suficientemente válido como para poder predecir rendimientos académicos futuros, poseen algunos inconvenientes; no reflejan situaciones concretas que se pretenden evaluar y no se ajustan a ciertos contenidos. Por otro lado, los test estandarizados están a su vez sumamente correlacionados con los test de inteligencia e influenciados por otras variables escolares (antecedentes fa-

⁶ SHAVELSON, R. y DEMPSEY, N. (1976): Generability of Measures of Teaching Behavior, *Review of Educational Research*, 46: 4, pp. 553-611.

⁷ *Ibidem*.

miliares, cultura, etc.), lo que hace difícil la interpretación de informes realizados con estos instrumentos, como es el caso del informe Coleman, tan ampliamente citado en la bibliografía educativa⁸.

La utilización de estos test requiere una previa tamización por parte del investigador; eliminación de items que se correlacionen altamente con la inteligencia, factores culturales, etc., y la incorporación de otros que estén relacionados con la propia situación a medir. Esta tamización escasamente ocurre en los casos en que se utiliza este tipo de test.

2. Ante este tipo de dificultades algunos autores como Popham⁹ se inclinan por diseñar *test especiales para unidades de aprendizaje*, de acuerdo con los contenidos y objetivos de la unidad a tratar. Este tipo de pruebas permite recopilar datos acerca de diversas dimensiones del aprendizaje y son mucho más representativos de los rendimientos específicos de la situación que se analiza. Estos tests recogen períodos cortos de aprendizaje, pero desgraciadamente no han logrado apreciar el rendimiento escolar a largo plazo. La puesta en marcha de estos instrumentos exige una serie de pasos previos que aseguren su validez y fiabilidad. No es muy usual que en las investigaciones que actualmente se publican en revistas especializadas se contemple la fase de validación y fiabilidad de éstos. La falta de validez de los instrumentos utilizados invalida los resultados obtenidos.

3. *Las técnicas de observación* posee también limitaciones. La observación informal, si bien puede aportar aspectos importantes a la hora de seleccionar variables de estudio, resulta de una enorme subjetividad en la recogida de información y de alto nivel de inferencia, no permitiendo establecer relaciones entre variables.

La observación sistemática, una de las técnicas más frecuentemente utilizadas, plantea algunos inconvenientes¹⁰. Los sistemas de categorías suponen una interpretación pre-específica de los elementos esenciales a observar, recogiendo aspectos básicamente cuantitativos de los fenómenos o conductas. Esta forma de observación no recoge procesos dinámicos, ni posibles variabilidades en la conducta a observar. A la observación participativa, planteada como alternativa en esta cuestión se le acusa de subjetividad en la recogida y presentación de resultados.

4. *Los cuestionarios o escalas de actitudes* tampoco propician una valoración precisa. Tanto la selección de preguntas realizadas por el investigador como las escalas de respuesta ofrecidas suponen una interpretación del observador de la situación a observar.

A la imperfección de los instrumentos de medida se une la escasa atención a su validación y fiabilidad. En la mayoría de informes de investigación no aparece discusión alguna acerca del grado de validez y fiabilidad del instrumento utilizado, ex-

⁸ BERLINER, D. (1976), *op. cit.*

⁹ POPHAM, W. (1971), Performance Test of Teaching Proficiency: Rationale, Development and Validation, *American Educational Research Journal*, 8, pp. 105-117.

¹⁰ KYRIACOU, C. y NEWSON, G. (1982), Teacher Effectiveness: a consideration of research problems, *Educational Review*, 34: 1; pp. 3-12.

cepto en aquellos casos en que el tema de estudio sea un problema metodológico de validación de test, escalas, etc. Muy pocas veces ocurre que cuando se recurre a la observación se mencione el sistema de entrenamiento ofrecido a los observadores antes de la recogida de datos, y menos aún que se planteen estudios para comprobar la eficacia de sistemas de entrenamiento¹¹.

Problemas de muestreo. En este apartado nos referiremos al muestreo en diferentes vertientes: *selección de ejemplos representativos, selección de unidades de tiempo y selección de los sujetos.*

La principal dificultad en los estudios sobre educación estriba en la adecuada selección de situaciones educativas que sean representativas de los fenómenos que se estudian, es decir, definir un universo del que poder extraer un ejemplo medio representativo del universo. Los ejemplos seleccionados son a lo sumo significativos para una limitada población y para unas limitadas variables.

Las unidades de tiempo consideradas en las investigaciones presentan enormes oscilaciones, variando entre períodos mínimos a registros más amplios. No tenemos hoy día desgraciadamente información suficiente sobre la periodicidad necesaria para que los datos sean válidos.

El principio básico para conseguir la generalización de resultados es obtener una muestra representativa de la población de estudio. Esta premisa queda relegada en la mayoría de los casos debido a la falta de colaboración de los sujetos afectados. Ello obliga, en algunos casos, a trabajar con muestras voluntarias que inevitablemente conducen a datos sesgados, y otras, a proceder con sujetos de reserva, precaución que por otra parte apenas se contempla en las investigaciones educativas. De todas las publicaciones sobre esta temática muy pocas se plantean la necesidad de una selección adecuada de la muestra.

Problemas de diseño. Cualquier diseño utilizado, por muy elaborado que sea, supone la aceptación de supuestos simplificativos. En los estudios no experimentales, debido a la ausencia de la aleatorización y manipulación de los tratamientos, no se pueden extraer conclusiones causales entre estas variables. Por otro lado el no control de otras variables incidentes, hace que deba contemplarse la posible incidencia de otras causas no medidas referente a la variable dependiente, y por tanto los resultados sean limitados.

Los diseños experimentales asumen reducciones situacionales que desvirtúan el poder de generalización. Si bien se obtiene la hipótesis de causalidad debido a la asignación aleatoria de los sujetos experimentales a los tratamientos, que a su vez sólo es posible porque el experimentador manipula los tratamientos, se olvidan aspectos relevantes e importantes en su medición, tales como las variables contextuales. Estos diseños presentan una enorme dificultad a la hora de controlar estas variables por medio del diseño y de las técnicas estadísticas utilizadas. También el control de los tratamientos es problemático en la mayoría de los estudios.

¹¹ KILLIAN, J. (1982): A comparison of Training Methods for Systematic Observation, *Journal of Classroom Interaction*, 18:1, pp. 19-22.

Problemas de interpretación estadística. Consideraremos en este apartado dos aspectos; selección y determinación del tratamiento estadístico, e inducción e interpretación de resultados.

La utilización de un determinado estadístico exige ciertas asunciones acerca de los datos. Cuanto más complejo es el estadístico más condiciones exige su aplicación. Las investigaciones educativas tienden a usar estadísticos sofisticados, no cumpliendo con sus condiciones. Es muy frecuente observar que el análisis de varianza (estadístico de uso muy común en educación) se utiliza sin comprobar previamente las tres condiciones que exige; normalidad de las variables dependientes, igualdad de varianzas entre los distintos grupos de tratamiento e independencia de los factores con los que se trabaja o variables que se miden¹².

Otro error muy común es usar estadísticos que no son apropiados a los datos. Por ejemplo; el uso del coeficiente de correlación presupone que las variables están relacionadas conjuntamente de forma lineal; sin embargo se utilizan en variables de enseñanza que no están relacionadas de esa forma. En estos casos los resultados están oscurecidos o desvirtuados por la selección errónea del estadístico, hecho que se acentúa en los procesos que implican un alto nivel de inferencia, tales como el análisis de factores y el análisis de covarianza.

En la inducción e interpretación de resultados vamos a considerar algunos aspectos; el primero se relaciona con la pregunta de si los resultados obtenidos son suficientemente representativos para tomarse en cuenta y poder generalizarse a toda la población. En este sentido hemos de anotar que desgraciadamente en las investigaciones educativas no existe una preocupación excesiva por la aleatorización de la muestra, ni por la elección estadística de un N representativo, hecho que a su vez se agrava al no conocer la distribución teórica de los datos.

Otro problema que se plantea es la dificultad de establecer el grado o nivel para considerar relación entre variables. El investigador induce la información y no analiza si su grado de asociación es superior al nivel de significación estadístico establecido para sus datos. Su significación dependerá del tamaño de la muestra y de la elección del N, tanto como del grado de relación establecido. La mayoría de investigadores olvidan interpretar sus resultados en términos de su significación teórica o práctica, y generalizan sus resultados a todas las situaciones y a todos los sujetos.

Y para terminar aduciremos a la falta de adecuación entre la interpretación y el diseño de investigación utilizado. En los diseños no experimentales no se puede establecer una relación causa-efecto como a veces se plantea.

Los problemas metodológicos no se reducen a los anteriormente mencionados. Nuestra sucinta exposición del tema ha pretendido que el lector tome conciencia de las dificultades y errores más usuales que se dan en las investigaciones educativas, a fin de que se aprecien los datos ofrecidos en su justa medida, no cayendo en el dogmatismo científico.

¹² RUIZ, L. (1977), Métodos estadísticos de investigación. Introducción al análisis de Varianza (Madrid, Instituto Nacional de Estadística).

Los aspectos enumerados pueden servir de referencia a la hora de reflexionar sobre el desarrollo metodológico de los informes publicados. Pueden ser útiles también para perfeccionar futuros diseños de investigación, alentando a plantear temas de estudio que puedan resolver algunos de los problemas metodológicos que tienen planteados las ciencias de la educación.

Dra. M.^a PILAR COLAS BRAVO
Profesora de Pedagogía Experimental
Facultad de Filosofía y CC de la Educación
Gonzalo de Bilbao, 5 y 7
SEVILLA