

Investigación y evaluación. Implicaciones y efectos Algunas reflexiones metodológicas sobre investigación y evaluación educativa

JOSÉ MANUEL GARCÍA RAMOS

Catedrático de Métodos de Investigación y Evaluación Educativas
Universidad Complutense de Madrid
Director del Centro Universitario Francisco de Vitoria

RESUMEN

El presente artículo describe algunas reflexiones metodológicas sobre Investigación y Evaluación educativa que el autor considera relevantes, al analizar diversos puntos de vista que sobre esta problemática existen en la actualidad.

Tras analizar seriamente los fines de la investigación científica, los rasgos característicos que definen hoy a la ciencia, y algunos principios importantes que también hoy caracterizan a la epistemología (filosofía de la ciencia) actual, el autor analiza dos conceptos clave: Causalidad y Empirismo, así como otras cuestiones sobre método y paradigmas en Investigación y Evaluación Educativas.

Posteriormente y tras poner en relación, medida y evaluación e investigación e innovación educativas, define su concepto de Métodos de Investigación y Evaluación Educativa, y su propuesta de superación del debate epistemológico apoyándose en algunas propuestas personales que reflejan la importancia de trabajar con teorías más ricas y más complejas, que combinen variables de todo tipo, sobre una estrategia básicamente hipotético-deductiva.

Finalmente, el artículo ofrece algunas conclusiones sobre investigación en Evaluación.

ABSTRACT

Analyzing the different points of view that actually exist on research and educational evaluation, the author describes some methodological reflection on this article that he considers relevant.

Once the goals of scientific research, the specific characteristics that today define the science and some important principles that today characterize the actual Epistemology (science philosophy) have been analyzed, the author analyzes two key concepts: Causality and Empirism as well as other subjects such as method and paradigm in research and educational evaluation.

Afterwards and once measurement and evaluation and educational innovation have been related, he defines his concept of Methods of research an educational evaluation and this proposal of one pass the epistemological based on some personal proposal that reflect the importance of working with theories more rich and complex that combine all kind of variables based on a hipotetic-deductive strategy.

Finally the article offers some conclusions on research on evaluation.

Introducción

Es preciso reconocer que asistimos a unos momentos en los cuales se están multiplicando los puntos de vista sobre la metodología de la investigación y Evaluación educativa. Son muchos los temas metodológicos sobre los que hay y se están formando opiniones con repercusiones teóricas y prácticas.

Por ello se hace preciso situar la investigación educativa (y por ende, a la evaluación) en relación con la Ciencia, la Tecnología, la Pedagogía, las distintas disciplinas pedagógicas y la misma práctica educativa, desde un nivel conceptual y epistemológico, para pasar después a analizar las consecuencias de tales relaciones. Estas reflexiones conceptuales nos ayudarán también a definir con mayor precisión, nuestro concepto de Investigación y Evaluación, así como su ámbito de aplicación, en el campo educativo.

1. Los fines de la Investigación Científica

Los clásicos niveles de la investigación educativa reflejan de alguna manera los principales fines: Básica, Aplicada, Activa y Experimentación. La primera vía la considero sustantiva, vital, desde el reconocimiento de que su ausencia en cualquier campo, deja sin fundamento la pretensión de cientifidad de cualquier rama del saber y por supuesto, sin ella se hace difícil pensar en la existencia de los otros niveles.

Ahora bien, ello no implica quitar o restar importancia a los otros dos niveles de investigación. En los últimos años hemos asistido a un sano esfuerzo por acercar la investigación a la praxis educativa y ello es también importante, sin por ello olvidar la importancia de la investigación básica. **La investigación pedagógica aplicada** y **la investigación activa**, se dirigen fundamental y específicamente a la solución de problemas o a la explicación de fenómenos y situaciones educativas. Cuanto más próximos a la acción concreta y contextualizada, más lejanos y ajenos a la búsqueda de teorías generales.

Todo ello nos lleva directamente al debate epistemológico sobre la Ciencia, las Ciencias y la Pedagogía como Ciencia y también, al debate so-

bre «las metodologías de la investigación» y «la controversia de paradigmas».

2. Ciencia

El término ciencia es hoy considerado un término equívoco, ambiguo, difícil de caracterizar, independientemente de cual sea su objeto. Las razones de ello son varias.

Sin detenernos a analizarlas, en la actualidad suelen definirse **como rasgos más característicos de la ciencia, los siguientes** (Hdez. Pina, 1994):

1. **La objetividad.** Se trata de una característica del método científico que se consigue a través de la **contrastación intersubjetiva**.
2. **La actitud crítica.** Si algo caracteriza al científico contemporáneo es su actitud crítica. Esta actitud falsacionista es central en los planteamientos metodológicos.
3. **La facticidad.** Son muchos los autores que sostienen que la base empírica y la experiencia son los aspectos que mejor distinguen el conocimiento científico del no científico.
4. **La racionalidad.** Esta característica nos permite sistematizar de forma coherente enunciados contrastables y fundamentar una teoría sobre la realidad que la sustenta
5. **La contrastación.** Contrastar una teoría es someterla a prueba de la experiencia. Se trata de un proceso que supone conjugar lo racional y lo fáctico.
6. **La autonomía.** La actividad científica posee un carácter primordialmente interno aunque ningún investigador carece de elementos «a priori»; es decir, **todos llevamos implícita una cosmovisión subyacente** que orienta de alguna manera nuestras posiciones teóricas o conceptuales. En este sentido la actividad científica es interna, pero no exenta de planteamientos previos procedentes de esta cosmovisión.
7. **Lenguaje claro y preciso.** Esta característica posibilita la comunicabilidad de los contenidos, facilitando además el carácter autocorrectivo del proceso científico.
8. **La sistematización...** tanto en el procedimiento como en la organización de los contenidos y conocimientos.

9. **La progresión...** capacidad que la ciencia tiene para resolver problemas.

Como podemos apreciar, el concepto ciencia ha ido evolucionando hasta llegar a entenderse como una realidad compleja cuyos componentes más destacables serían:

- Tratarse de un tipo de conocimiento más riguroso que el ordinario.
- Ser una actividad que comporta un método propio.
- Poseer un lenguaje específico.
- Aparecer como una realidad dinámica de carácter autocorrector que busca incrementar los niveles de verosimilitud.

3. Algunos principios de la Filosofía de la Ciencia en la actualidad

Según los especialistas sobre el tema... **los principios más importantes que caracterizan a la epistemología actual, son los siguientes** (6 principios):

1. *Rechazo del Inductivismo* (comentamos más adelante).
2. *El problema de la verosimilitud de las teorías*. No es posible confirmar las teorías.
3. *Determinación de las teorías por la lógica* (por el método). **Nunca se puede separar totalmente, conocimiento obtenido de método utilizado.**
4. *Condicionantes sociales de la ciencia*. Importancia de las características psicológicas y sociológicas de las comunidades científicas en el progreso de la ciencia, incluso por encima de los componentes puramente lógicos de dicho proceso.
5. *Relativismo en la demarcación de la ciencia*. La tradicional discusión entre racionalistas y relativistas, en torno a qué es ciencia y qué no lo es, parece superada tras la afirmación de Kuhn, según la cual «no existe ninguna norma superior a la aprobación de la comunidad científica».
6. *Incomparabilidad de las teorías*. Si la teoría y las observaciones o hechos son inseparables, parece claro que los resultados y consecuencias de cada investigación no son separables de su contexto teórico.

Sobre la base de estos principios, se establecen unas condiciones para el uso de **cualquier aproximación o para la valoración de las aproximaciones paradigmáticas que actualmente se postulan.**

La filosofía de la ciencia actual, parece no admitir, que existan criterios absolutos de demarcación del concepto de ciencia, y tampoco la existencia de conocimientos científicos como conocimientos seguros adquiridos gracias a la observación y a la experimentación y, en consecuencia, que exista un procedimiento inferencial que nos permita derivar teorías científicas de manera absolutamente fiable. En este punto incidiremos posteriormente.

En este sentido, se afirma que todos los conocimientos deberán ser juzgados por sus propios méritos, por las soluciones que aportan dentro de un contexto social (y de una comunidad científica).

La epistemología actual parece orientarnos a aceptar abiertamente la legitimidad de las distintas metodologías de investigación educativa, evitando todo género de exclusivismo o complejos de «superioridad» o «inferioridad».

En todo caso y como investigadores, cualquiera que sea nuestra metodología prevalente o preferente, hacemos lo correcto al exigirnos unos a otros, mutuamente, rigor y calidad en los planteamientos, desarrollo y conclusiones de nuestras investigaciones.

4. Dos conceptos: Causalidad y Empirismo

Este concepto, muy consolidado en la filosofía tradicional y en muchos enfoques actuales de la investigación científica, parte de la idea según la cual «nada ocurre sin causa», todo efecto es consecuencia de una causa inmediata, pudiéndose así ordenar una sucesión de efectos y causas. Es evidente que el pensamiento causal ha jugado y juega un papel importante en la investigación científica; por otra parte, el conocimiento de las causas de unos efectos, puede ayudar a dirigir la acción educativa a las mismas, evitando o produciendo efectos indeseados o deseados.

Sin embargo y especialmente en este siglo, se ha sometido a crítica tal concepto, aplicado a la realidad; incluso desde la propia Física, donde podemos ver la evolución, desde la certeza absoluta de la mecánica de Laplace hasta las actuales concepciones evolutivas (catástrofe-caos) de un Prigogine, pasando por las aproximaciones probabilísticas (principio de indeterminación de Heisenberg). Por otra parte, las restricciones y condiciones exigidas para el cumplimiento de la causalidad entre variables, son tan fuertes, que la misma posibilidad de establecerse parece difícil.

Sin embargo y pese a todo ello, **la causalidad está implícita, cuando no explícita, en la mayoría de los postulados.**

Con la aparición de los modelos causales y de los modelos de análisis de estructuras de covarianza, el concepto de causalidad es utilizado en la vía no-experimental. En esta vía el asunto es mucho más complejo y delicado, **pudiéndose establecer «causalidad»** por analogía, en diseños o estudios donde ni hay manipulación de variables, ni azar posible, **a través de la Teoría.**

Finalmente y para evitar la «crítica fácil» al concepto de causalidad, conviene distinguir entre dos conceptos de causalidad de diferente calado: **un concepto ingenuo y simplista** y **un concepto más sólido y complejo.** En este sentido, **conviene superar el primero**, en virtud del cual una sola causa induce un solo efecto mediante un proceso susceptible de ser teóricamente representado mediante una estructura funcional sencilla y manejable, percatándose de que en muchos casos —la mayoría en nuestras Ciencias Sociales y en nuestra Ciencia Pedagógica— **lo que se da es un «enjambre» o conjunto de con-causas generadoras de «enjambres» de efectos, ligados entre sí**, mediante procesos que no admiten una representación teórica en términos de estructuras funcionales sencillas.

Una cosa será admitir este segundo concepto de causalidad y no el primero, por ingenuo y otra bien distinta concluir por el primero que la causalidad, en cuanto posibilidad de intelección racional de los fenómenos, no existe. **Evidentemente, defendemos la necesidad de usar el concepto de causalidad, correctamente, pero sin complejos de ningún tipo y con toda la extensión y fuerza con que se es posible.**

El término «empirismo», como el ya citado de «ciencia», es hoy un término polisémico, con muchos matices y que requiere muchas cualificaciones.

En nuestro campo y desde nuestro quehacer de investigadores, partimos de la **necesidad de derivar y contrastar nuestros conocimientos con la realidad. El dato empírico es importante**, sea como base del contraste de una afirmación o para la búsqueda de acuerdo de interpretaciones. Ciertamente, se excluyen quienes sólo hacen interpretaciones en sus aproximaciones a la realidad, sin aportar datos, ya que eliminan la posibilidad de una contrastación externa, **la posibilidad de una comparación inter-subjetiva, una condición para mí ineludible**, en el contexto de nuestra disciplina.

5. El método y la cuestión de los Paradigmas en la Investigación y Evaluación Educativas

De lo anterior se deriva que, si bien la objetividad y la neutralidad no deben exigirse, por inalcanzables, **si debe pedirse en toda investigación rigor y sistematicidad**, que en la mayoría de los casos vienen dados por la preocupación por la **validez interna o lógica de nuestras investigaciones**, preocupación que debe plasmarse siempre en la fase de evaluación de la consistencia interna de nuestras teorías y también por el **respeto al método** en la evaluación de la consistencia interna entre la teoría y los hechos.

La discusión, polémica o debate, especialmente candente en los últimos años, en torno a los Paradigmas y metodologías, tiene mucho que ver con el problema del método, aunque parte quizás de un nivel de abstracción superior y estrechamente vinculado a cuestiones epistemológicas, de filosofía de la ciencia.

Las raíces del debate son antiguas. En la actualidad, hay que señalar que este debate ha generado mucha confusión. No vamos a reproducir —en esta exposición— los términos e historia de este debate. Ha sido y para algunos todavía es un «debate candente».

Por otro lado, la controversia va más allá todavía, pues afecta a quienes piensan que las distintas aproximaciones son compatibles, frente a los que opinan que no es posible la integración.

Por otro lado, **el problema es en parte diferente al hablar de investigación y al hablar de evaluación de programas-investigación evaluativa**, en donde los objetivos son distintos, no buscándose tanto construir teoría, conocimiento más o menos generalizable, sino más bien obtener información para valorar y tomar de decisiones (de mejora, de cambio, de transformación).

La complejidad del debate y del problema y las situaciones de los investigadores ante él hacen difícil enfocararlo en unas pocas páginas. No obstante, conviene distinguir brevemente dos niveles en esta discusión, **Planteamientos epistemológicos; Metodológico** (métodos y diseños), y **Técnico**.

Este primer nivel, trata de recoger los distintas posiciones, enfoques o paradigmas que dirigen la investigación y la evaluación. Claro que, bajo las distintas definiciones de investigación y de evaluación, por lo general subyacen presupuestos filosóficos distintos acerca de epistemología y valor. Dichos presupuestos orientan y definen la naturaleza de la evaluación y de la investigación que se realiza. Dependiendo de ellos, de cada posición, se utilizarán unos métodos de investigación y unos modelos de evaluación, así

como distintos diseños/planteamientos de investigación-evaluación, como herramientas preferentes de trabajo. La literatura reciente está llena de argumentos acerca de la naturaleza del conocimiento.

Sin duda, que la investigación y la evaluación, especialmente ésta, aunque ambas mucho, se ven fuertemente afectadas por estas posiciones. De hecho, lo más o menos aceptable de una investigación concreta o de una investigación evaluativa particular, depende, a menudo, de la óptica que toma el que la valora.

Básicamente y en un primer intento clasificatorio, pueden distinguirse tres posiciones generales: a) naturalista-cualitativo; b) experimental-cuantitativo-racional y c) ecléctica. No nos detenemos a explicitarlos en este momento. El criterio de verdad, es básico para distinguir entre las posiciones extremas del continuo (en este nivel).

Un segundo problema es si: ¿Pueden integrarse? ¿Son complementarios? ¿Son compatibles?

Hay tesis para todos los gustos, desde las Tesis de la oposición o **incompatibilidad epistemológica**, hasta la tesis de la **complementariedad**, pasando por la tesis de la **unidad epistemológica** con diversidad metódica aceptada. Muchos otros autores han defendido la complementariedad, pero desde posiciones y planteamientos muy distintos y dispares.

Mi posición personal en este punto es la de procurar evitar el caer en la renuncia a la construcción de una teoría sistemática y ello sobre la siguiente tesis: en un proceso racional-objetivo no tiene porqué rechazarse el que muchas variables que han de ser integradas en la teoría no sean ni cuantificables, ni siquiera fácilmente representables mediante categorías analíticas sencillas; pero esto no tiene porqué conducirnos a posiciones en la que parece renunciarse a establecer una teoría sistemática; por el contrario, a lo que debe conducirnos es a un **amplio esfuerzo de producción conceptual y analítica, progresivamente conducente al establecimiento de teorías más ricas y más complejas** (veremos este punto más adelante).

En este sentido, no conviene olvidar que una de las mayores acusaciones que se hacen a la investigación educativa, es la de que no se apoya en la teoría.

Lo mismo ocurre con la evaluación educativa: su justificación no siempre está clara y sus objetivos muchas veces no se precisan.

En este nivel las consecuencias de los planteamientos epistemológicos son evidentes.

Conviene, en todo caso, no confundir el debate paradigmático con el debate metodológico cuantitativo-cualitativo; aunque es verdad, que para entender cómo se llegó al segundo debate haya que tener en cuenta el pri-

mero. Así, por ejemplo, positivismo y cuantitativismo no son la misma cosa, no conviene identificar paradigma con metodología.

En este nivel, mantenemos nuestra posición anterior, a la que conviene añadir un matiz importante: probablemente, **la fecundación progresiva de la teoría** —dentro del marco en que nos movemos, esto es, la búsqueda de una teoría compleja progresiva— **exige la aportación de elementos que hoy quedan objetivamente fuera del horizonte conceptual perceptivo de las diferentes posiciones teóricas o argumentales propias de nuestro ámbito educativo**, elementos estos que muy probablemente deban ser tomados de una amplia reflexión, por ejemplo, antropológica.

El tema de los métodos en el ámbito de la **Evaluación Educativa**, presenta algunas variaciones importantes respecto al ámbito de la investigación educativa. Aún cuando la **metodología de la evaluación de programas** proviene claramente de la metodología de investigación, ha tenido que ir haciendo adaptaciones sucesivas para resolver problemas específicos que han ido dando lugar a métodos propios que, siendo básicamente los mismos, en su proceso, al de la investigación aplicada, introducen algunas variantes o particularidades.

En la actualidad, **la evaluación de programas** se caracteriza, por la falta de acuerdo en lo que es, en cómo debe hacerse y para qué debe hacerse la propia evaluación. De hecho, podríamos volver a agrupar los grandes enfoques metodológicos alternativos, de nuevo, en dos principales: a) Esquema Tradicional-objetivista y básicamente **externo**, y b) Modelo Participativo-subjetivista-crítico y básicamente **interno**; alternativas que nos devuelven a la discusión originaria que ya hemos comentado.

En este ámbito de la Evaluación de Programas, investigación-evaluativa, se recogen, dentro de sus diseños, técnicas y herramientas al uso, todas las posibilidades que ofrece la investigación clásica en CC. Sociales, más algunas otras posibilidades, más o menos novedosas y las procedentes de la corriente cualitativa, que han ido surgiendo en la corta historia, todavía, de este área de la investigación, o, si se prefiere, en este área evaluativa. En todo caso, **no debe olvidarse** que una característica de la investigación evaluativa, que la distingue de la mayoría de las demás clases de investigación, es la de **que se realiza en un marco de acción**, alguna otra cosa, además de la investigación, está sucediendo, **hay un programa en funcionamiento, o un centro en marcha,...** **La investigación** es aquí un apéndice, un algo que «también está presente», algo que **tiene**, sin embargo, **prioridad secundaria**. Con frecuencia, los investigadores proponemos **cambiar el orden de prioridades y con algo de razón, pero ello puede ser un grave error.**

En todo caso, **el investigador evaluativo**, el evaluador de programas, **debe conocer mucho sobre investigación**, conocer lo que se encuentra en los textos acerca de métodos de investigación y luego tiene que aprender a aplicar tal conocimiento a un ámbito que a menudo resulta inhóspito para áreas importantes de su conocimiento. «Si luego no sale de su libro de texto, corre el peligro de hacer un trabajo que nada tenga que ver con lo que se le encargó, de «chocar» con el personal del programa o del centro, con el que trabaja y de ver desperdiciados los resultados del estudio, si es que llega a completar su trabajo» (Weis, 1975).

6. La medida en investigación y evaluación educativa

Un capítulo esencial, que conviene mencionar es el de los instrumentos de medida y más importante aún, el de la **medida misma**, como base y elemento esencial e inseparable de los procesos de investigación y evaluación educativos.

Medir es justamente el procedimiento para definir, obtener y ofrecer datos, información útil que posibilite el análisis para la obtención de conclusiones en todo proceso de investigación, así como el juicio o valoración posterior en que toda evaluación consiste.

También en el ámbito de la medida ha tenido un fuerte eco el debate cuantitativo-cualitativo. Sin volver al debate, hay que reconocer que la preocupación por los datos y su calidad, en estos últimos años y en ambas corrientes metodológicas, es grande. Mientras en la **perspectiva cuantitativa** se siguen dos direcciones: a) mejoras en las teorías y métodos de aplicación; y b) profundización en las teorías formales; **en la perspectiva cualitativa** se busca suplir la imperfección de las aproximaciones a la realidad e ir aumentando su validez, corrigiendo fuentes de error, por la reflexión y análisis cuidadoso de las condiciones.

En mi opinión, **el concepto de medida**, aceptando las clásicas definiciones de distintos autores... debe ligarse a un concepto más amplio que el puramente operacional **debiendo vincularse siempre a una Teoría Sustantiva o sistema de proposiciones referidas a un conjunto de constructos**, que sirven para describir, explicar y precedir algunos aspectos limitados de su dominio.

La definición de Lord y Novick (1968), es una magnífica síntesis de lo anterior: «la medida es un procedimiento para la asignación de números —puntuaciones, medidas— a propiedades especificadas de unidades experimentales, de modo que las caractericen y preserven las relaciones especí-

ficadas en el dominio comportamental» (p. 17). De ahí que la medida comience siempre con un procedimiento para identificar y relacionar los elementos del mundo real con elementos o constructos del sistema lógico, por medio de una definición semántica precisa; una vez hecho esto, puede procederse a la medición.

La medición educativa se enfrenta, en este final de siglo, con grandes problemas teóricos y prácticos, derivados todos ellos de las dificultades de construir instrumentos adecuados a los constructos que se definen, muchos de los cuales presentan graves dificultades de delimitación. De aquí, que se siga diciendo, que el principal problema que ha de enfrentar un instrumento de medida, es el de la validez de constructo.

7. La Evaluación educativa y sus ámbitos

Con relación al Concepto de Evaluación y sus ámbitos, en primer lugar, conviene señalar la evolución que ha sufrido el mismo Concepto de Evaluación a lo largo de los últimos cincuenta años. Ha evolucionado rápida y profundamente, desde la conceptualización de Tyler hasta nuestros días. Inicialmente estaba restringido —en el ámbito educativo— por la referencia al aprendizaje-rendimiento de los alumnos.

Actualmente y como consecuencia de su desarrollo histórico, el concepto de Evaluación se ha ampliado enormemente. Toda actividad educativa intencional es susceptible de Evaluación y por tanto es Evaluación Educativa.

De hecho, prácticamente todas las definiciones que sobre el término se han dado pueden englobarse en alguno de los dos conceptos siguientes. Definiciones como las de Bloom (1975), Gronlund (1985), etc..., van en la línea de la evaluación de productos educativos individuales (Assesment). Definiciones como las de Phi Delta Kappa, 1971, Stufflebean y otros (1971 y 84), Berk (1981);... van en la línea de la evaluación de programas (Evaluation). Otras definiciones tratan de abarcar ambos conceptos (Pophan, 1980; De la Orden, 1982; Pérez Juste, 1986); incluso se puede apreciar dicha orientación en Scriven, 1967 y otros muchos autores.

La tendencia actual es considerar por separado las dos facetas del término evaluación y ello, fundamentalmente, porque implican procesos y metodologías bastante distintos. En este sentido, **la diferencia básica** entre ambos conceptos parece estar **en la unidad-objetivo de análisis**: un individuo o sujeto único («assesment») o bien, un programa o modalidad de intervención («evaluation»).

8. Investigación e innovación educativa

Respecto a la relación entre Investigación e Innovación educativa lo primero que conviene apuntar es que ambos conceptos están ligados, se apoyan y justifican entre sí. La investigación educativa, en último término, se legitima por su compromiso con la praxis educativa, por el compromiso con el perfeccionamiento de la acción educativa y por tanto, por la repercusión que tiene en la optimización del aprendizaje a través de la innovación. Por otro lado, conviene que toda innovación se apoye en algún tipo de resultados de investigación; por lo menos, toda innovación que pretende permanecer y consolidarse como modo sistemático de intervención educativa perfecta.

Sin embargo, de todos es conocida la poca repercusión que, hasta el presente, ha tenido la investigación en la práctica educativa. Este tema es y ha sido objeto de abundante literatura en los últimos años. No vamos a reproducir aquí, de nuevo, la discusión sobre las causas que han provocado esta situación y sus consecuencias sobre ambos conceptos y sobre el debate cuantitativo-cualitativo.

Lo que sí parece claro, es que se buscan formas de superar esta situación y sobre todo se observa una creciente preocupación por conectar adecuadamente Investigación e Innovación, Investigación y cambio educativo, Innovación y Evaluación, etc... **Sin duda, la vía metodológica es esencial**, pues a través de la mejora y adecuación de los métodos de investigación podremos conectar adecuadamente investigación e innovación.

9. Concepto de Métodos de Investigación y Evaluación Educativa

Antes de aproximarnos un poco más al concepto de Métodos de Investigación y Evaluación, me parece necesario hacer algunas consideraciones más, sobre las características peculiares del conocimiento pedagógico, para enmarcar mejor esta aproximación.

Una de las cuestiones que surge siempre en nuestro ámbito, es la consideración de si la Pedagogía es ciencia o no lo es. En parte ya hemos contestado a esta cuestión anteriormente.

Ciertamente, el quehacer pedagógico-educativo es un arte. **La cuestión es, si ¿además es una ciencia?** Tal como definimos la ciencia anteriormente y de forma amplia **la Pedagogía es también una ciencia**, en cuanto que utiliza métodos que se caracterizan por tener valor heurístico y ser acción reflexiva, rigurosa, sistemática, organizada y disciplinada. La Pedago-

gía tiene modos de conocimiento rigurosos y metódicos con el fin de descubrir leyes en el objeto de estudio y de expresar los conocimientos adquiridos sistemáticamente. Es decir, tiene objeto, rigor, método, leyes y teorías sistemáticas.

Ciertamente, lo que más cuesta aceptar en Pedagogía, es que tenga leyes; pero el problema se reduce a la cuestión de a qué nivel de generalidad-universalidad es necesario que se formulen enunciados de la Pedagogía para que puedan considerarse leyes científicas. Citando a Sherwood (1969), quien presenta los distintos tipos de generalización aceptables en la Ciencia, puede afirmarse que la Pedagogía está consiguiendo leyes y teorías progresivamente más generales-universales.

Por otro lado y en cuanto que los conocimientos pedagógicos se pretenden tengan un alto grado de practicidad, al permitir a más corto o más largo plazo, su aplicación a la mejora educativa, puede ser considerada una Ciencia Tecnológica. Se trata pues, de la mejora del sistema educativo, de la educación en general, como criterio de validez (cuestión esta no exenta de valores de enjuiciamiento).

Lo que habrá que evitar, es que al enfatizar el carácter práctico del saber pedagógico, descuidemos y olvidemos que los fenómenos educativos, objeto de estudio de la Pedagogía, por complejos y multivariados que sean, son susceptibles de conocimiento sistemático integrable en teorías.

Tradicionalmente, el concepto de Pedagogía Experimental ha servido para denominar un campo de trabajo intelectual y profesional en torno a la metodología científico-experimental, empírico-científica, y empírico-experimental, por usar diversas denominaciones que se han utilizado para definir la metodología propia de la disciplina. En este sentido, lo que ha caracterizado a nuestra disciplina ha sido su «adjetivo» (experimental) definiéndose históricamente como una modalidad de conocimiento científico que se fundamenta en la experimentación y observación científica.

Numerosos autores al elaborar una definición de la Pedagogía Experimental, hacen referencia a algunas de sus características básicas: **ser una disciplina metodológica, un enfoque en el estudio de la educación, una disciplina aplicada, dirigida a la resolución técnica de problemas que plantea la realidad educativa.**

Si tuviéramos que avanzar en una definición propia, diríamos que la Pedagogía Experimental, los **Métodos de Investigación y Evaluación, se caracterizan, por las siguientes notas:**

- a) **Finalidad científica.**
- b) **Rigor en sus procedimientos.**

- c) **Énfasis en el Método.**
- d) **Carácter Teórico-Práctico.**
- e) **Perspectiva o enfoque de los problemas educativos.**
- f) **Metodología empírico-experimental;** metodologías de base empírica y/o experimental que buscan el contraste científico de la realidad.

Ahora bien, aceptando todo lo anterior, no es menos cierto que el gran drama que las Ciencias Sociales (Economía, Psicología, etc...) tienen planteado en este momento y la Pedagogía no es una excepción (en las Ciencias Naturales y en la Física por ejemplo, también ocurre aunque en menor medida) **es el haber descubierto que la mayoría de las variables explicativas de los fenómenos con que trabajamos, son variables «cualitativas»** (entendiendo por ellas las que comúnmente denominamos Nominales y especialmente, Ordinales) y no tanto variables «cuantitativas» (susceptibles de ser medidas cardinalmente, lo que comúnmente denominamos como de Intervalo o de Razón).

Ello nos obliga a definir Teorías Explicativas cada vez más generales, más complejas, más potentes en el sentido popperiano, en las cuales estén presentes ambos tipos de variables. Las primeras, no cuantificables estrictamente, son la mayoría en nuestro ámbito educativo. Aún aceptando la medida de variables como Rendimiento, aptitudes y algunas otras, al nivel de Cuasi-intervalo («cuantificables»), la mayoría siguen siendo ordinales y/o nominales. Estas variables son, sin embargo, susceptibles de algún tipo de medida, por supuesto son susceptibles de ser definidas en términos lógicos y sobre todo, presentan alto poder lógico-explicativo de muchos fenómenos educativos.

En este sentido, entiendo que nuestra disciplina debe conducirnos a trabajar con Teorías Tipo 2 (T2), según el esquema siguiente:

- a) T1: Sólo maneja categorías analíticas cuantificables (k), por tanto, hipótesis con sólo variables k, que a través del proceso inferencial desembocan en Teoremas K, que son susceptibles de contraste empírico.
- b) T2: maneja todo tipo de categorías analíticas que puedan ser acuñadas para caracterizar **razonablemente** un grupo de fenómenos; algunas serían **K** y otras no, serían **q** (cualitativas). Por tanto, tendríamos Hipótesis K e hipótesis Q, orgánicamente ligadas, que a través del proceso inferencial desembocan en Teoremas que impli-

can «relaciones funcionales» o estructurales que son contrastables empíricamente en términos de k y con implicaciones k , sobre condiciones definidas por las categorías q . Ciertamente, en todo caso habremos de definir y justificar razonablemente, qué variables consideramos k y cuáles q .

Por otro lado, y asumiendo la ampliación que en los últimos años se ha producido en la disciplina que con la denominación de «Métodos de Investigación» en Educación ha incorporado otros métodos de investigación además del enfoque tradicional y el concepto de ciencia que hemos defendido anteriormente, parece lógico afirmar que nuestra disciplina en la actualidad se caracterice por su **matiz metodológico**, que abarcaría así las distintas aproximaciones empírico-experimentales, cuyos objetivos son: la descripción, explicación, predicción y control, de los fenómenos educativos.

Una definición que podría ayudarnos a resumir todo lo dicho hasta aquí, nos llevaría a decir que entendemos nuestra disciplina, **como el estudio de los métodos y metodologías, técnicas y procedimientos utilizados y utilizables, para obtener un conocimiento y una explicación científica de los fenómenos educativos, así como para mejorar la realidad educativa, mediante una aproximación técnica y rigurosa a los problemas educativos y sociales, que tal realidad plantea.**

Por otra parte y sobre la base de todo lo dicho hasta aquí, me gustaría **concluir con algunas propuestas personales sobre cómo superar el actual debate metodológico**. Para ello se hace necesaria la propuesta de una metodología de investigación **razonable**, que yo apoyaría en mi propuesta anterior de trabajar con Teorías Tipo 2 (ya desarrollado anteriormente). Ello significa que:

- a) **Toda Teoría debe proponer relaciones causales**, sea dicha teoría cualitativa o cuantitativa. En este sentido, **toda Teoría** es, por definición, **hipotético-deductiva**.

Conviene recordar aquí que la Inducción no puede aceptarse, pues los enunciados observacionales están cargados de teoría, destruyéndose así la tesis de que la Ciencia comienza con la observación. En esta línea, la llamada «vía inductiva» que prefiere que la «teoría emerja de los datos» no parece aceptable. Como sabemos **los datos no existen independientemente de la estructura teórica de alguien. Algo es erigido en dato por alguien**, que lo hace a partir de alguna clase de estructura teórica (buena o mala; consciente o in-

consciente; explícita o implícita) que le permite reconocer como dato algo que percibe.

Como ejemplo, basta decir que yo no puedo distinguir entre un «esquizofrénico» y un paciente con «delirio paranoico». Sin embargo, un buen psiquiatra sí que puede. De la misma manera que no soy capaz de distinguir si se ha estropeado el cárter o una bujía, por el ruido del motor del coche. Un buen mecánico puede hacerlo. Y es que **para mí**, los síntomas del enfermo o el ruido del motor **no son datos**, son hechos o fenómenos que no reconozco como datos mientras no tenga una estructura teórica mínima que lo permita. No puedo o no soy capaz de discernir o de aceptar como significativo un dato que no puedo identificar como tal.

Ahora bien, esto no quiere decir que la percepción desde una teoría inicial (T_0) no pueda plantearme un campo de extensión conceptual o tensión problemática con respecto a T_0 . En ese caso, no es la mera percepción de algo distinto —sin más— lo que hace que esa tensión problemática desaparezca, sino la **reelaboración hipotético-deductiva a partir de T_0** ; ésta es la que me conduce a una reflexión a partir de un nuevo dato.

Partimos así de la existencia siempre de una Teoría Fundamental, entendida como **conjunto, frecuentemente no bien articulado, de principios epistemológicos y sustantivos (teóricos) que tenemos por ciertos y a través de los cuales tematizamos nuestra observación.**

Ciertamente, **la realidad es objetiva**, pero no lo es la visión de la misma desde nuestra subjetividad. **La visión de la realidad está siempre mediada por la subjetividad.** De hecho, todos podemos constatar que sobre una misma realidad uno ve cosas que otros, simplemente, no ven y viceversa.

Parte sustantiva de esta propuesta personal de superación del debate metodológico es pues, el reconocimiento de que **necesitamos Teorías**, debiendo dirigir nuestro esfuerzo principal a proponerlas y después, contrastarlas con la realidad. Estas teorías, como ya se dijo, manejarán variables cuantitativas (k) y cualitativas (q) que permitan ligarlas orgánicamente y que a través del proceso inferencial (hipotético - deductivo) desemboquen en el establecimiento de relaciones funcionales o estructurales que sean contrastables empíricamente.

Ciertamente, nuestra mayor dificultad es que trabajamos —en la mayoría de las ocasiones— con variables cualitativas. Ello nos obligará a un gran esfuerzo de producción teórica, conceptual y analítica, progresivamente conducente a teorías más ricas y más complejas.

En cierto sentido, puede decirse que la mayoría de nuestras Teorías, son o serán cualitativas y deberán constituirse por la vía hipotético-deductiva en continua interacción con la realidad que pretendemos explicar, comprender y mejorar. Lo mismo ocurre en otras C.C. Sociales, como por ejemplo la Economía, donde todas las teorías actuales trabajan con variables cualitativas y proceden del trabajo desarrollado por vía hipotético-deductiva.

- b) Junto a lo anterior, repetimos, **ello no implica reducir las variables con las que trabajamos a variables únicamente cuantitativas (k)**, tal como comentamos al hablar de Teorías tipo 1 y tipo 2. Si lo hiciéramos, estaríamos dejando fuera del campo explicativo de la teoría fenómenos en los que intervienen variables «q» (la mayoría, como hemos dicho). No admitiríamos así una posición — como ésta — sistemáticamente reduccionista. El clásico ejemplo del «farol y la moneda» puede iluminar lo que aquí defendemos¹. De ahí que el «enfoque cuantitativo» no sea aceptable si implica una reducción como la descrita.
- c) Ahora bien, **ello no significa** —como en ocasiones se observa— que debamos **identificar gratuitamente estructura hipotético-deductiva y estrategia reduccionista** (ya descrita). Los que así lo hacen, creen o postulan que la superación del reduccionismo también

¹ En este clásico ejemplo, muy usado de todas las Ciencias Sociales, una persona pierde una moneda al cruzar la calle. Al llegar al otro lado se da cuenta de ello. En la calle sólo hay un farol, que ilumina una parte pequeña de la calle. La persona busca la moneda sólo en la zona iluminada por la farola y así lo hace durante horas. Llega un amigo y le pregunta que qué hace. El «buscador» de la moneda dice que la busca desde hace horas y que no la encuentra. El amigo le invita a buscarla por toda la calle, negándose el «buscador», pues no está iluminada por el farol. La moraleja es muy sencilla: habrá que «arrancar la farola, sin desconectarla de la fuente de electricidad, si hace falta» o buscar un «grupo electrógeno» o una linterna potente, para buscar la moneda, sin reducir la búsqueda a la zona iluminada por el farol. *Habrà, pues, que acomodar el método al objeto de estudio, que es la realidad objetiva, evitando reduccionismos como el señalado. Nuestras teorías, progresivamente, deberán tratar de explicar toda la realidad, para ello, frecuentemente, deberán trabajar con variables «q» y variables «k».*

implica el abandono de estructuras teóricas hipotético — deductivas, cayendo así en una postura difícilmente compatible con el estudio racional y sistemático de los fenómenos. **En la medida en que «investigación o enfoque cualitativo» signifique el no establecimiento de relaciones causales, racionalmente discernibles, no nos parece aceptable.**

Lo que propugnamos, pues, es **un enfoque progresivo** en el que, asumiendo los elementos positivos de ambos enfoques, **se vaya a la construcción de modelos explicativos que combinen variables «q» y «k»**, que establezcan relaciones causales entre ellas, lo que permitirá una explicación (intelección y comprensión, son, en este sentido, sinónimos de explicación) que naturalmente debe poder estar abierta a varios tipos de contrastación, en algunos aspectos o dimensiones cuantitativas... y en otras cualitativas.

La tarea no es fácil, ciertamente, pero el objetivo justifica el esfuerzo que entre todos debemos realizar.

10. Investigación sobre evaluación. Algunas conclusiones

En la actualidad, pues, nos encontramos en una época caracterizada por una visión más equilibrada de las posibilidades de la investigación educativa. Un período «de evaluación crítica», en el cual se intenta evaluar la calidad de la investigación pedagógica y su contribución al progreso de la educación. Una época caracterizada por una gran preocupación metodológica, una época de «reflexión» sobre lo conseguido y lo no conseguido, intentando definir:

- a) Cómo usar mejor los métodos y técnicas existentes.
- b) Qué áreas o problemas educativos han sido y no han sido suficientemente abordados desde la metodología de investigación.
- c) Nuevos modos de adaptarse al estudio de los problemas.
- d) Nuevas formas de definir problemas educativos, susceptibles de ser investigados desde distintas alternativas metodológicas.

De la misma manera, podríamos afirmar que la vía metodológica y de Investigación es esencial para la Evaluación Educativa.

En primer lugar, no debemos olvidar que Investigación y Evaluación son dos cosas distintas, que ciertamente comparten muchos elementos de

proceso y método, pero con un objetivo diverso: la evaluación, conduce a decisiones; por el contrario, la investigación conduce a conclusiones. En la Evaluación de programas o su correlato metodológico, la Investigación-Evaluativa, la investigación es secundaria en prioridad; lo prioritario es valorar el programa, centro,... como fuente de toma de decisiones.

Con ello se quiere indicar también que la evaluación debe adaptarse al contexto y al objeto evaluado (programas, centros, etc...) y perturbar lo menos posible las operaciones y dinámicas del mismo. Ello no significa que no se produzcan interferencias en ocasiones. Lógico, entre otras cosas es preciso recoger datos, pero habrá que procurar, aquí especialmente, recabar sólo los necesarios (en ocasiones, el investigador o el evaluador solicitan más información de la que utilizarán después).

En segundo lugar, debemos tener en cuenta algunos otros problemas que habrá que cuidar en la investigación evaluativa son:

- a) tendencia del programa o del centro, a cambiar, mientras está siendo evaluado;
- b) las relaciones entre evaluadores y el personal del programa del centro; y
- c) el hecho de que el programa se halla inserto en un sistema de organización o el centro educativo en un sistema socio-político más amplio y de que la naturaleza de dicho sistema condicionará, de alguna manera, los resultados. **La metodología de investigación y el sentido común pueden ser buenas herramientas** para ayudar a minimizar esta problemática.

En tercer lugar y dada la importancia que concedemos a la Evaluación que, no lo olvidemos, tiene más influencia decisiva y directa sobre la calidad de la educación, contar con un claro **sistema evaluativo es básico**. Si la evaluación se constituye en motor y palanca del aprendizaje y de la mejora de la calidad, es porque “determina en gran medida – al prescribir los objetivos de la educación— las características de la enseñanza y del aprendizaje, lo que los alumnos aprenden y como lo aprenden, lo que los profesores enseñan y cómo lo enseñan, los contenidos y los métodos; en otras palabras, el proceso y el producto de la educación” (Orden, 1982).

De aquí la importancia de cuidar la coherencia entre los objetivos de un programa o centro educativo (las metas) y la características del esquema evaluador.

No es baladí la preocupación de las responsables de los programas o centro educativos, por la evaluación. No evaluar los programas o hacerlo in-

tuitivamente conlleva riesgos excesivos. Evaluar de cualquier forma, es poner todos los medios para dirigir la enseñanza y el aprendizaje hacia metas desconocidas y seguramente no deseadas. En todo caso, es necesaria la voluntad "política" de evaluar y no sólo eso, junto a ella es preciso definir: decisiones a tomar (uso de los datos evaluativos), criterios de valoración y objetivos del programa.

Surge así uno de los problemas más peliagudos: *la elección de criterios de evaluación*. En los primeros años de la todavía corta historia de la evaluación de programas, la evaluación prestó poca atención al rol de los valores. Scriven (1983) sugiere que ello fue debido a que los evaluadores ingenuamente creían que sus actividades podían y debían ser "libres de valores". Pero muchos evaluadores aprendieron de la experiencia, que era imposible, en el mundo político de la programación social y educativa, tomar decisiones sin tener valores que fueran sobresalientes y quizás más claramente cuando los criterios evaluativos fueran seleccionados y justificados. Demasiados evaluadores e investigadores se han formado y entrenado en la creencia de que los valores no son parte de la ciencia.

En cuarto lugar, una pregunta básica que siempre debemos hacernos, al plantearnos una evaluación de programas, es la de ésta *está justificada siempre*. Como afirma Weis (1975), "por herético que parezca, la respuesta es no. La evaluación, en su calidad de investigación aplicada, está consagrada al principio de utilidad". Si no tiene ninguna influencia en las decisiones, resultará una tarea vana. La misma autora señala cuatro clases de circunstancias en las cuales no vale la pena proceder a realizar una evaluación (pp. 23-24):

- a) Cuando no hay nada que preguntar acerca del programa. Se está realizando y las decisiones sobre su futuro, o no se van tomar o ya se han tomado.
- b) Cuando el programa carece de una orientación clara. El personal del programa improvisa sus actividades de un día para otro, basándose en pocas reflexiones y menos principios y el programa cambia y se va transformando, "vagabundea" y se orienta como puede. Hay poca razón para llamar a esto "programa".
- c) Cuando las personas que deberían saberlo, no pueden ponerse de acuerdo en qué es lo que están tratando de realizar en el programa. Si existen grandes discrepancias en las metas pensadas, la evaluación carece de fundamento sobre el cual levantarse.
- d) Cuando no hay suficiente dinero, ni personal lo bastante cualificado para resolver la evaluación. Esta es una tarea exigente, que requiere tiempo, dinero, imaginación, tenacidad y destreza.

Aunque algunas personas pudieran argumentar que en circunstancias como las anteriores, “mejor es hacer algo que nada”, la experiencia demuestra que el estudio que se realiza así, rara vez arroja resultados o “luces”, sino más bien acaba convirtiéndose en “papel mojado” y termina siendo arrinconado o arrojado a la papelera.

Es, pues, esencial plantearse *el porqué* y *el para qué* de la evaluación de un programa. A veces uno puede observar que la decisión positiva sobre la realización de una evaluación particular, obedece a razones poco legítimas, tales como por ejemplo, algunas de las señaladas por Weis (1975):

- a) Aplazar una decisión. En vez de enviar la “patata caliente” de la decisión sobre un programa, a una comisión y esperar su informe, se encarga un estudio evaluativo que requiere aún más tiempo.
- b) Eludir responsabilidades. En ocasiones grupos o personas responsables de la organización del programa tienen criterios opuestos. Los gestores o administradores toman la iniciativa de pedir un estudio evaluativo que permita tomar decisiones con objetividad evitando apasionamientos. Sin embargo, en muchos casos los administradores ya tienen tomada la decisión antes de llamar a los evaluadores, pero desean revestirla del “ropaje legitimador” de la investigación-evaluación.
- c) Relaciones públicas. A veces la evaluación es considerada como un modo de darse “auto-bombo”. El administrador cree que cuenta con un programa exitoso y busca la manera de hacer visibles sus “glorias”. Un buen estudio “les vendrá de perlas”. Copias del informe serán elevadas a todas las instancias superiores, consiguiendo así efectos “autoglorificantes” o económicos, si se persigue la financiación del programa. Suchman (1970) habla de dos finalidades emparentadas: “echar humo en los ojos” y “adornarse”. En el primer caso se intenta justificar un programa débil, seleccionando para la evaluación únicamente aquellos rasgos que, superficialmente, tienen buen aspecto. La evaluación para “adornarse” trata de encubrir el fracaso de un programa evitando toda estimación objetiva. Por supuesto, los motivos del administrador no tienen por qué ser torcidos o egoístas. Por lo general, es necesario conseguir apoyos para un concepto o un proyecto en cuyas virtudes se cree. La evaluación puede ofrecer gran parte de esos apoyos.
- d) Cumplir los requisitos de una financiación externa o social. Muchas decisiones políticas de financiación de un programa exigen como requisitos la evaluación del mismo. Desde el punto de vista de los

que conceden ayuda económica, corriendo el riesgo de financiar un proyecto no probado, es razonable exigir que se aporten datos del grado en que el proyecto funciona bien. En este sentido, la evaluación corre el riesgo de ser “un ritual necesario” para convencer o “aplacar” al financiador, pero carente de utilidad real.

Todo lo anterior nos hace pensar, que la evaluación puede ser una empresa racional, que a menudo se efectuará por razones que no son racionales. Ello no debe hacernos desconfiar de ella, pero debe prevenimos de cualquier tipo de justificación. No todas son legítimas y es siempre necesario ponderar las razones que hacen que el evaluador acometa su tarea. Para empezar, será necesario preguntarse: ¿quiénes esperan obtener respuestas? Y ¿qué esperan obtener?

En resumen, la investigación evaluativa, en opinión de quienes la practican, es una manera de aumentar la racionalidad de las decisiones. Téngase en cuenta que, al contar con información objetiva acerca de los resultados de los programas, es posible tomar decisiones atinadas en materia de asignación de partidas presupuestarias y planificación de los programas. Los programas que rindan buenos resultados se ampliarán; los que no los rindan serán desestimados o modificados drásticamente. Por otra parte, sabemos que, si no se evalúan sistemática y objetivamente los programas de acción que se pongan en marcha, es fácil que los mismos programas sean desechados o ampliados por criterios “políticos” y maniobras en beneficio de grupos o personas particulares. Los datos, si se utiliza la investigación evaluativa, sustituirán a los favores y a otras negociaciones políticas, de manera que se tomarán decisiones más racionales.

No obstante lo anterior, es verdad que la breve historia de la investigación evaluativa no ha sido demasiado fructífera; es más, algunos autores la consideran decepcionante. Pocos ejemplos de aportaciones importantes a las política y/o los programas pueden citarse. Ello se debe, en parte, a las notables resistencias de las organizaciones a ser evaluadas, a que sus programas sean evaluados, a ofrecer la información que se precisa en la forma en que se necesita y al cambio no deseado. Otra parte de la culpa estriba en la manera en que a veces está estructurada, dotada de personal y ejecutada la misma evaluación. En ocasiones hay incoherencia entre los fines propuestos en la evaluación y el tipo de estudios o programas evaluados.

“Gran parte de la decepción que venimos comentando se debe al carácter poco realista de las expectativas que se ponen en las aportaciones que ofrecerá la investigación evaluativa. Por lo general, un estudio de evaluación no desemboca en la obtención de datos decisivos e inequívocos acerca

del valor de un programa. Sus resultados muestran a menudo pequeños cambios, o ambiguos, efectos “de poca monta”, resultados en los que han influido sucesos específicos de lugar y momento. Tal vez se necesitará un estudio continuado a lo largo del tiempo y comparando proyectos, para hablar con seguridad de éxito y de fracaso. Además, para quienes toman decisiones, la evaluación de la evidencia obtenida es sólo una “entrada” entre muchas otras. Deben considerar otros muchos factores: receptividad del público, reacción de los participantes, costo, cantidad de recursos y de personal de que se disponga, alternativas existentes, etc. Por otra parte, la puesta en marcha de un programa eficaz puede no aceptarse en una determinada comunidad o colectivo por considerarse que viola supuestos y valores aceptados por ese grupo”.

El reto está ahí.

11. Referencias Bibliográficas

- Ary, D. *et al.* (1984). *Introducción a la investigación pedagógica*. México: Intera-mericana.
- Béjar, I. I. (1983). Introduction to Item response models and their assumptions. En Hambleton (Ed.), *Applications of item response theory*. Vancouver: Educ. Research Institute of British Columbia.
- Berk, R. A. y otros (1975). *Educational evaluation methodology: the state of the art*. Baltimore: John Hopkins. Univ. Press.
- Brown, H. (1983). *La nueva filosofía de la ciencia*. Madrid: Tecnos.
- Bunge, M. (1984). Qué es y para qué sirve la ciencia de las Ciencias y la Tecnología. *El País*, 14-11-84 (p. 28).
- Buyse, R. (1949). Origen y desarrollo de la pedagogía experimental. *Revista Española de Pedagogía*, n.º 28 (pp. 5-9).
- Campbell, D. T. y Stanley, J. C. (1978). *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires: Amorrortu.
- Chalmers, A. F. (1982). *¿Qué es una cosa llamada ciencia?* Madrid: Siglo XXI.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49 (12) (pp. 977-1003).
- Cook, T. D. y Campbell, D. T. (1979): *Cuasi-experimentation*. Chicago: Rand Mac-Nally.
- Cook, T. D. y Reichardt, Ch. (1986). *Métodos cualitativos y cuantitativos en investigación*. Madrid: Morata.

- Cronbach, L. (1974). Beyond the two discipline of scientific psychology. Paper American Psychological Assoc. Los Angeles, California.
- Dendaluce, I. (1988). Una reflexión metodológica sobre la investigación educativa». En Dendaluce, I. et al. (1988), *Aspectos metodológicos de la investigación educativa*. Madrid: Narcea - II Congreso Mundial Vasco.
- Dewey, J. (1953/1989). *Cómo pensamos*. Buenos Aires: Paidós.
- Fernández Cano, A. (1995). *Métodos para evaluar la investigación en psicopedagogía*. Madrid: Síntesis.
- Gage, N. L. (Ed.) (1963). *Handbook of research on teaching*. Chicago: Rand McNally.
- (1989). The paradigm wars and their aftermath: «a historical sketch» of research on teaching since 1989. *Educational Researcher*, 18 (7) (pp. 4-10).
- García Ramos, J. M. (1986). Validación del constructo en el ámbito pedagógico. *Revista Española de Pedagogía* n.º 174 (pp. 535-554).
- (1989a). *Bases pedagógicas de la evaluación*. Madrid: Síntesis.
- (1989b): Evaluación de Programas, Evaluación de Centros Educativos. En Pérez Juste, R. y García Ramos, J. M. (1989), *Diagnóstico, evaluación y toma de decisiones*. Madrid: Rialp.
- (1991). Recursos metodológicos en la evaluación de programas. *Bordón*, vol. 43, 4. (pp. 461-476).
- Garrison, J. (1986). Some principles of postpositivistic philosophy of science. *Educational Researcher* (pp. 9-15).
- Guba, E. G. y Lincoln, I. S. (1989): *Fourth generation evaluation*. Newbury Park: Sage.
- Gullicksen, H. (1953). *Theory of mental tests*. New York: Wiley.
- Hernández Pina, F. (1994). Métodos de investigación en Educación. Memoria de Cátedra. Murcia.
- Howe, K. (1985). Two dogmas of educational research. *Educational Researcher*, 14 (8).
- (1988). Against the quantitative-qualitative incompatibility thesis. *Educational Researcher*, 17 (4).
- Howe, K. y Eisenhart, M. (1993). Criterios de investigación cualitativa (y cuantitativa). Prolegómenos. *Revista de Educación*, 300 (pp. 173-189).
- Husen, T. (1988). Paradigmas de la investigación en educación. Un informe del estado de la cuestión. En Dendaluce, I. (Ed.), *Aspectos metodológicos de la investigación educativa* (pp. 46-60). Madrid: Narcea.
- Kaplan, A. (1985). Research methodology: behavioral sciences. *The International Encyclopedia of Education*, 7 (pp. 4293-4300). Oxford: Pergamon Press.

- Kerlinger, F. N. (1989), 3.^a ed.: *Investigación del comportamiento. Técnicas y metodología*. México: Interamericana.
- Kunh, Th. S. (1984). *Segundos pensamientos sobre paradigmas*. Madrid: Tecnos.
- Lecompte, M. D.; Millroy, W. L. y Preissle, J. (Eds.) (1992), *The handbook of qualitative research in education*. San Diego: Academic Press.
- Linn, R. (1988). Medición educativa. Algunos problemas y tendencias actuales. En Dendaluze, I. (Ed.), *Aspectos metodológicos de la investigación educativa*. Madrid: Narcea.
- Lord, F. M. y Novick (Eds.) (1968). *Statistical theories of mental test scores*. Reading, Mass.: Addison-Wesley.
- Martínez Arias, R. (1995). *Psicometría: Teoría de los tests psicológicos y evaluativos*. Madrid: Síntesis.
- Morales, P. (1988). *Medición de actitudes en psicología y educación. Construcción de escalas y problemas metodológicos*. San Sebastián: Txartalo.
- Muñiz, J. (1992). *Teoría clásica de los tests*. Madrid: Pirámide.
- (1993). *Teoría de respuesta al ítem*. Madrid: Pirámide.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Orden, A. de la (1984). Investigación pedagógico-experimental y praxis educativa. En VV.AA., *Cincuentenario de los estudios de Pedagogía*. Madrid: Facultad de Filosofía y CC. de la Educación. Universidad Complutense de Madrid.
- (1985) (Ed.). *Investigación Educativa*. Diccionario de CC. de la Educación. Madrid: Anaya.
- (1989). Investigación cuantitativa y medida en educación. *Bordón*, vol. 41, 2 (pp. 217-236).
- (1995). Innovación e Investigación en el ámbito educativo. *Bordón*, 47 (2). (pp. 135-141).
- Pedhazur, E. (1982). *Multiple regresion in behavioral research* (2.^a ed.). New York: Holt.
- Pérez Juste, R. y García Ramos, J. M. (1989). *Diagnóstico, evaluación y toma de decisiones*. Madrid: Rialp.
- Savage, C. W. y Enrlich, P. (Eds.) (1992). *Philosophical and foundational issues in measurement theory*. Hillsdale, New Hersey: Earlbaum.
- Sechresth, L. y Figueredo, A. J. (1993). Program evaluation. *Annual Review of Psychology*, 44 (64) (pp. 645-674).
- Shadish, W. R. (1989). Critical multiplism: a research strategy and its attendant tactics. En Sechrest, L. *et al.* (Eds.), *Health services research: a fous on AIDS*. Agency health care policy res. Rockville, M.D. (pp. 5-28).
- Shulman, L. (1988). Paradigms and research programs in the study of teaching. A contemporary perspective. En Wittrock, M. C. (Ed.), *Handbook of research on*

- teaching*. New York: MacMillan. Traducido al castellano. Paidós/MEC 1989. (pp. 9-84).
- Smith, J. K. y Heshusius, L. (1986). Closing down the conversation: the end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher*, 15 (1) (pp. 4-12).
- Soltis, J. (1984). On the nature of educational research. *Educational Researcher*, 10, 13 (pp. 5-10).
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103 (pp. 677-680).
- (1951): *Mathematical measurement and psychophysics. A Handbook of experimental Psychology*. New York: Wiley.
- Stufflebean, D. et al. (1971). *Educational evaluation and decision making. Phi Delta Kappa*. Itaca, Illinois: Peacock.
- Tejedor, J. y otros (1994). Perspectivas metodológicas actuales de la evaluación de programas en el ámbito educativo. *Revista de Investigación Educativa*, 23 (pp. 93-127).
- Tiedeman, D. V. (1960). Experimental methods. *Encyclopedia of Educational Research*, 3.ª ed. AERA. New York: MacMillan.
- Tuckman, B. W. (1979). *Analyzing and designing educational research*. New York: Harcourt B.J. Pb.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, Mass.: Addison-Wesley.
- Van Dalen, P. B. y Meyer, W. J. (1966). *Manual de técnicas de investigación educativa*. Buenos Aires: Paidós.
- Weis, C. H. (1975). *Investigación evaluativa*. México: Trillas.
- Wittrock, M. C. (1986) (Ed.). *Handbook of research on teaching*. New York: MacMillan (3.ª ed.) (Traducción en Paidós/MEC).
- Wolf, R. L. (1990). *Evaluation in education*. Praeger Pb., 3.ª ed.