

La metaevaluación educativa

GUSTAVO ROSALES ESTRADA*

Educational metha-evaluation

Abstract. *The purpose of the present paper is to establish metha-evaluation (the evaluation of evaluation) as a system which allows us to make an objective evaluation of the educational process and place responsibility properly with the purveyors of educational service, and at the same time analyze its structural development and establish clear and objective criteria for its normative use.*

Introducción

La metaevaluación, como práctica de evaluar la evaluación (Scriven, 1974; Stufflebeam, 1974, 1978; Cook and Grunder, 1978), asume que los procesos correctos de evaluación requieren de contrastar sesgos, errores técnicos, costos excesivos y utilización de resultados, entre otros.

Estos contrastes permiten perfeccionar la evaluación y justificar la inversión de recursos de todo orden en estas empresas.

Sólo una pequeña proporción de estudios de evaluación son, a su vez, evaluados; de los pocos procesos de metaevaluación que se efectúan, una parte sustancial son de tipo interno, es decir, ejecutados por el mismo evaluador.

La resistencia a ser evaluado (Scriven, 1984) es una condición inherente al ser humano, la cual también afecta a los evaluadores.

Aunque la escasa frecuencia de metaevaluaciones es, en cierto grado comprensible, resulta injustificada la no detección de prácticas evaluativas poco correctas, cuya persistencia pone en tela de juicio la fiabilidad de la profesión docente.

I. Génesis de la metaevaluación

Scriven (1969) introdujo formalmente el término metaevaluación al finalizar la década de los años sesentas. La metaevaluación siempre apareció asociada a los programas evaluativos, pero fue hasta esa década cuando los evaluadores empezaron a discutir formalmente sus procedimientos y criterios (Stake, 1967, 1970; Stufflebeam, 1968).

Lessinger (1970) y Provus (1971) analizan el concepto bajo la denominación propia de “auditoría de programas”, y proponen procedimientos para realizar el seguimiento de estudios de evaluación, con la finalidad de asegurar su calidad y destacar la validez de sus conclusiones y recomendaciones.

La metaevaluación ha sido considerada tradicionalmente como un instrumento para evaluar proyectos específicos: la evaluación de los sistemas de evaluación y la investigación evaluativa aparecen como formas generales de metaevaluación; desde esta óptica, la metaevaluación no se limita a formular cuestiones únicamente sobre la metodología y efectos de un proyecto evaluativo en particular, sino que amplía su objeto para incluir la estructura de la evaluación en su conjunto y su impacto en la generación de información relevante.

II. Concepto

La definición de metaevaluación, a la luz del proceso educativo, debe ser coherente con la de evaluación, ya que se trata de un caso particular de la misma. Si la evaluación es un juicio de valor sobre algo, es decir, la determinación de su valor y mérito, la metaevaluación es un juicio de valor sobre la evaluación.

Stufflebeam (1981) la define como el proceso de diseñar, obtener y utilizar información descriptiva y de enjuiciamiento acerca de la utilidad, del valor práctico y de la adecuación ética y técnica de una evaluación, para guiarla y hacer público el informe de sus puntos débiles y fuertes.

El *Join Committee on Standars for Educational Evaluation* (1981: 166) define la metaevaluación como “la evaluación de la evaluación”, pero también habla del examen o auditoría de una evaluación como “examen y verificación independientes de la calidad de un plan de evaluación, de la suficiencia con que fue puesto en práctica, de la precisión de los resultados y de la validez de sus conclusiones”, lo que

* Centro de Innovación, Desarrollo e Investigación Educativa, UAEM.



supone una descripción más detallada de las implicaciones de la evaluación.

Stufflebeam (1987) da a la metaevaluación un carácter profesional, lo cual implica enjuiciar el valor y el mérito de lo que representa y hace la profesión del evaluador: servicios de evaluación, utilización de las evaluaciones, investigación evaluativa y desarrollo organizativo.

La metaevaluación utiliza las normas aceptadas por los evaluadores y trata de asegurar que se cumplan, sin descuidar su desarrollo y enfrentándose a las prácticas. El objetivo de la metaevaluación es ayudar a que la evaluación desarrolle todo su potencial; efectuándola correctamente producirá mejoras directas e incontestables en el sistema, los programas y las prácticas educativas. Si se le considera como una herramienta para analizar la calidad, la metaevaluación se puede aplicar a toda clase de estudios y metodologías, así como a distintos paradigmas de investigación. En un sentido más amplio De la Orden (1981) dice que "el poder de la metaevaluación permite optimar el objeto evaluado y, por consiguiente, la propia evaluación".

Estas consideraciones nos llevan a distinguir (Stufflebeam, 1978) entre la metaevaluación formativa y la sumativa. La formativa, aplicada antes de que la evaluación se efectúe, pretende ser una guía para ayudar a los evaluadores a llevar a cabo su proyecto de evaluación; la sumativa sería el estudio del mérito de una evaluación terminada. Ambas determinan el grado en que la evaluación es técnicamente adecuada, útil para guiar decisiones, ética al incidir sobre personas y organizaciones, y práctica en el uso de los recursos.

Stufflebeam (1974, 1981) distingue dos roles en el metaevaluador: el formativo, que guía la evaluación, examina su proceso y ayuda a los evaluadores a tomar decisiones acerca de la planificación, desarrollo e interpretación de los resultados, y el rol sumativo, que atestigua la fortaleza o debilidad de una evaluación, y realiza un examen dirigido a concluir con un informe sobre su mérito.

En el papel formativo, el metaevaluador actúa como un consultor; en el sumativo, es un agente externo que comprueba la calidad del proceso y del producto obtenido, con referencia a determinados criterios y normas.

1. ¿Cómo valorar el mérito de una evaluación?

No existe un consenso sobre un conjunto determinado de principios que definan la calidad de una evaluación.

Chelinsky (1987) afirma que intentar definir y medir la calidad de una evaluación no es tarea fácil, dado que la calidad depende de las condiciones impuestas a la evaluación examinada. Aunque pocos diseños son unánimemente aceptados, al aplicar adecuadamente criterios que definan una evaluación satisfactoria, se pueden elaborar diseños consistentes (Brinkerhoff, 1985). Sin embargo, la tarea de especificar criterios y defender esa especificación no es fácil. Los criterios no son evidentes y su selección implica siempre algún sistema de valores.

III. Criterios de calidad

Básicamente existen tres criterios principales de calidad:

1. Rigor: métodos precisos, correctamente aplicados.
2. Valor: mérito, utilidad e importancia de los resultados de la evaluación.
3. Eficiencia: efectividad del diseño y proceso de evaluación (Schwandt y Halpern, 1988).

1. Criterios de rigor

De las tres dimensiones citadas, destaca el rigor. En el plano teórico, todas las series de normas señalan que una evaluación rigurosa requiere, entre otras cosas, el uso de medidas válidas y fiables, así como métodos apropiados de análisis de datos. Pero en la práctica el consenso es más difícil, pues determinar procesos evaluativos aceptados por todos para documentar el rigor constituye una tarea problemática. Tres factores explican esta situación:

- a) Contexto del estudio: el tiempo y el costo afectan a la elección de métodos para establecer el rigor.
- b) Diseño del estudio: aunque los conceptos pueden ser análogos, los diferentes tipos de diseño requieren prácticas diferentes.
- c) Paradigma de investigación en el que está inscrito el proyecto: requiere de una lógica particular de justificación, la cual permite considerar que unas prácticas son más aceptables que otras para definir aspectos de rigor.

2. Criterios de utilidad

Stufflebeam (1981) basa la selección de criterios en la necesidad de obtener información de calidad de acuerdo a una doble perspectiva:

- a) Que la información sea un reflejo exacto de la situación, es decir, que sea consistente, donde los procesos de obtención y análisis deben conformarse con los criterios científicos actuales.

b) Referente al nivel del que la recibe y que podríamos llamar la capacidad de estar informado.

Estos criterios, que se denominan prácticos o de utilidad, son tan vitales que hacen que los criterios científicos señalados constituyan una condición necesaria, pero no suficiente, de calidad evaluativa.

3. Criterios éticos

Posavac y Carey (1989) interpretan las normas éticas para evaluadores e incluyen principios de conducta ética en investigación, así como descripciones de prácticas correctas de evaluación de programas.

La razón de la combinación de los dos aspectos es la creencia de que la ética en evaluación significa algo más que ser honesto con el dinero, los datos y los sujetos investigados. Aquí se encuentra implícita la idea de que la ética debe plasmarse en informes claros, útiles y precisos.

Especial énfasis adquiere la naturaleza política y aplicativa de las evaluaciones de programas, la cual se enfrenta más que a la investigación básica, a dilemas éticos. Por ello los evaluadores deben observar considerables precauciones al desarrollar sus estudios.

Objetividad, honestidad y justicia son algunos de los criterios que se deben usar para examinar evaluaciones de programas educativos, pues en estos elementos subyace todo un sistema de valores éticos, morales y profesionales inherentes a las condiciones de desarrollo cultural de quienes evalúan.

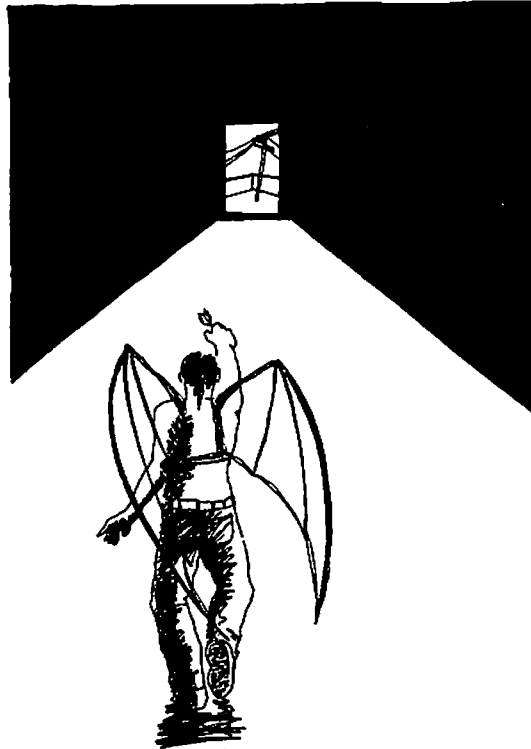
Estudios de la práctica evaluativa –informes, meta-evaluaciones, metaanálisis y comentarios críticos– proporcionan gran número de ideas y datos empíricos para generar y revisar teorías sobre cómo debe realizarse la evaluación, donde la participación dinámica y continua de evaluadores y evaluados es fundamental para retroalimentar al sistema que se analiza.

IV. Necesidad de establecer elementos normativos de evaluación

Una de las características de la práctica profesional es el respeto de normas referentes al servicio que garantizan. Sanders y Nafzinger (1974), dan tres razones para justificar la necesidad de establecer normas de evaluación:

1. El hábito en evaluación es un proceso que ha aumentado considerablemente, siendo difícil de abarcar en forma individual.

2. Se ha incrementado notablemente la evaluación en el ámbito educacional; en consecuencia, se hace necesario transmitir sugerencias y advertencias para la práctica de la evaluación educativa.



3. Ha habido un aumento del control sobre el modo de evaluar las prácticas evaluativas.

La evaluación de los programas puede influir en las decisiones y pensamientos de quienes se encuentran en situaciones de responsabilidad; la confianza en la evaluación frecuentemente la tienen audiencias que no pueden evaluar sus méritos; de lo anterior se desprende la necesidad de un estamento con autoridad sobre lo que constituye una buena práctica evaluativa, para que demandantes y evaluadores puedan tener un punto referencial común.

Se deben considerar cuatro dimensiones básicas de análisis en cualquier proceso evaluativo o metaevaluativo:

a) Utilidad: comprende las normas de evaluación necesarias para que ésta sea informativa, realista y transmisible. Requiere que los evaluadores se familiaricen con sus audiencias, se percaten de las necesidades de información de éstas, elaboren evaluaciones que respondan a tales necesidades y remitan de forma clara la información adecuada cuando sea necesaria. Todo esto debe basarse en normas de utilidad, las cuales se ocupan de determinar si una evaluación satisface las necesidades prácticas de información de los implicados en un programa. La utilidad se relaciona con el servicio que presta una determinada evaluación a una determinada audiencia.

b) Viabilidad: incluye las normas donde se reconoce que un proyecto de evaluación educativo debe ser realizable. Una preocupación es que el diseño de la evaluación sea operativo en el ambiente real en el que se ha de aplicar, y otra es que la evaluación sólo

consuma el material y tiempos necesarios para el logro de sus propósitos. En suma, la viabilidad requiere que las evaluaciones sean realistas, prudentes y económicas.

c) Legitimidad o corrección: establece cómo las evaluaciones educativas afectan al ser humano en formas distintas. Pretende que se protejan los derechos de las personas afectadas por una evaluación.

d) Precisión o exactitud: incluye aquellos elementos que determinan si una evaluación ha producido la información técnicamente adecuada y que sus conclusiones se ajusten de forma lógica a los datos obtenidos. Trata de asegurar que una evaluación revele y proporcione la información adecuada acerca de las características del objeto en estudio, así como determinar su mérito o valor.

Conclusiones

La pretensión de establecer los criterios que debe cumplir una evaluación o metaevaluación, permite que los investigadores educativos asignen diversas cualidades a la actividad evaluativa: que el juicio de

valor sea claro y riguroso, que permita llegar a conclusiones fundamentadas, que incluya una recolección sistemática de información y que se realice con rigor científico, entre otros requisitos. Pero aun incluyendo estos aspectos esas cualidades son insuficientes para resolver el problema de cómo definir la calidad de una metaevaluación.

Más allá de la satisfacción de uno o de varios criterios de los considerados por los expertos como ejemplos de buena práctica evaluadora, debe haber alguna característica más general que relacione todos ellos y que sea la esencia de la calidad evaluativa.

Esta calidad tiene que verse en el referente de la calidad de la educación, porque la evaluación es el motor de aquélla.

Sólo puede ser de calidad una evaluación cuando mejora la calidad educativa. Eficacia, eficiencia y funcionalidad, son las dimensiones fundamentales donde radica la construcción de la calidad de una metaevaluación, expresión de un sistema de coherencias múltiples entre los componentes del proceso de la metaevaluación. ♦

BIBLIOGRAFÍA

- Brinkerhoff, R. (1985). *Program evaluation: a practitioners guide for trainers and educators*. Boston.
- Cook, T.
- ____ y Campbell, D. (1979). *Quasi-experimentation. Design and analysis issues for field settings*. Chicago.
- ____ y Gruder, C. (1978). "Metaevaluation research", en *Evaluation Quarterly*. No. 2.
- Chelinsky, E. (1987). "What have we learned about the politics of program evaluation?", en *Educational evaluation and policy analysis*. 9:3.
- De la Orden, A. (1981). "Evaluación del aprendizaje y calidad de la educación", en *La calidad de la educación*. CSIC. Madrid.
- Join Committee on Standards for Educational Evaluation (1981). *Standards for evaluations of educational programs. Projects and materials*. McGraw-Hill. New York.
- Lessinger, L. (1970). *Engineering accountability for results in public education*. Phi Delta Kappa.
- Posavac, E. y Carey, R. (1989). *Program evaluation. Methods and case studies*. Prentice Hall. Englewood Cliffs, New Jersey.
- Provous, N. (1971). *Discrepancy evaluation for educational program. Improvement and Assessment*. Berkeley, CA. McCutchan.
- Sanders, J. y Nafzinger, D. (1974). *A basis for determining the adequacy for evaluation designs*. Kalamazoo. Western Michigan University.
- Schwandt, T. y Halpern, E. (1988). *Linking auditing and metaevaluation. Enhancing quality in applied research*. Sage. Newbury Park.
- Scriven, M.
- ____ (1969). "An introduction to meta-evaluation", en *Educational product report*.
- ____ (1974). "Standards for the evaluation of educational programs and products", en Borich, G. (Ed.). *Evaluating educational programs and products*. Educational Technology Publications. Englewood Cliffs.
- ____ (1984). "Evaluation ideologies", en Comer et al. (1984). (Eds.). *Evaluation studies review annual*. Vol. 9. Beverly-Hills Sage.
- Stake, R.
- ____ (1967). "The countenance of educational evaluation", en *Teacher's college record*.
- ____ (1970). "Objectives, priorities and other judgement data", en *Review of educational research*. 40 (2).
- Stufflebeam, D.
- ____ (1968). *Evaluation as enlightenment for decision making*. Columbus. Ohio State University. Evaluation Center.
- ____ (1974). *Metaevaluation: the evaluation center*. Kalamazoo. Western Michigan University Press.
- ____ (1978). *Metaevaluation: an overview*. Evaluation and the Health Professions. 1 (2).
- ____ (1981). "Metaevaluation: concepts, standards and uses", en Berk, R. (Ed.) (1982). *Educational evaluation methodology: The state of the art*. Baltimore, John's Hopkins. University Press.
- ____ (1987). "Professional standards for assuring the quality of educational program and personal evaluations", en *International journal of educational research*. Vol. 11, 1.