

LA FUNCIÓN OPTIMIZANTE DE LA EVALUACIÓN DE PROGRAMAS EVALUATIVOS*

Arturo de la Orden Hoz
Universidad Complutense de Madrid

Es evidente que en el mundo actual y, especialmente, en las sociedades más desarrolladas, la evaluación ocupa un lugar tan amplio y destacado que podríamos definir esta etapa histórica como «la era de la evaluación». Para explicar esta situación satisfactoriamente, habríamos de recurrir, sin duda, a diferentes vías y considerar muy diversos factores. Sin embargo, una constatación de las condiciones generales en que este hecho se produce nos permitirá un primer acercamiento a su explicación. En efecto, la evaluación es un fenómeno que se desarrolla y adquiere relevancia en momentos históricos caracterizados por el cambio social. En realidad, se trata de un mecanismo permanente de ajuste de las decisiones en las actividades sociales organizadas, cuya función es optimizar las alternativas. Si no hubiera cambio, no habría alternativas y, en consecuencia, la evaluación perdería su significado.

La evaluación como característica inherente a toda actividad humana intencional, se manifiesta tanto en la vida cotidiana, escasamente planificada, cuanto en la acción sistemática de las empresas individuales o colectivas del hombre. En el devenir de la rutina diaria, el hombre evalúa, valora constantemente personas, objetos, situaciones y acontecimientos, juzgándolos, previa contrastación con instancias o modelos, más o menos explícitamente representados.

En la acción individual o colectiva sistematizada (la política, la administración, la actividad económica, la medicina, etc.) la evaluación aparece como un componente esencial. La valoración de las consecuencias, el control de la calidad, el diagnóstico de la salud, que se utilizan como indicadores del proceso o de la adecuación de los resultados, son diferentes manifestaciones del fenómeno evaluativo.

* Este artículo está apoyado en otros trabajos del autor (De la Orden, 1982 y 1997).

Esta multiplicidad de situaciones evaluativas justifica la visión de este campo de actividad como un conjunto disciplinar relativamente disperso, cultivado como ocupación parcial por profesionales de distintas disciplinas para satisfacer necesidades, a su vez, dispersas, pero que, como afirma Morell (1990), comportan un núcleo de principios, doctrinas o creencias precisas, pero mal definidas. En la evaluación podemos rastrear elementos procedentes de las disciplinas de origen de los evaluadores —pedagogía, psicología, sociología, economía, política, etc.— cuya formación básica está vinculada a su área de procedencia, pero, en general, sienten una necesidad común de definir su afiliación profesional, al menos, parcialmente, en términos no estrictamente específicos del campo disciplinario en que se mueven.

No obstante, en los últimos 25 años, la evaluación ha sufrido un proceso de transformación profunda de sus bases estructurales y conceptuales. Como señala House (1992), estructuralmente se ha ido integrando cada vez más en el funcionamiento organizativo de las instituciones y contextualmente se ha flexibilizado, aceptando un alto grado de pluralismo teórico y metodológico.

Hoy, en un ambiente de cambio educativo generalizado, efecto tanto de las grandes reformas en marcha como de la innovación permanente, a nivel institucional y de aula, la evaluación educativa cobra también su máximo protagonismo y presenta características similares a las apuntadas para el campo en su conjunto, como demuestra el énfasis puesto en la evaluación del aprendizaje de los alumnos, la evaluación de los sistemas educativos en la perspectiva regional, nacional e internacional, los avances en la evaluación de centros escolares y profesores y especialmente la evaluación de programas educativos específicos y de formación continua de profesionales.

HACIA UNA PRECISIÓN CONCEPTUAL

Como premisa para la caracterización de la función optimizante de la evaluación de programas educativos, objetivo de este artículo, parece pertinente delimitar con la máxima precisión el propio concepto de evaluación como referente de los argumentos y análisis conducentes a la demostración de la tesis planteada.

La evaluación constituye esencialmente un juicio de valor sobre una realidad y, como todo juicio, se apoya en una comparación. Comparación y juicio son, pues, los componentes esenciales de la evaluación, su núcleo conceptual. Como la comparación exige dos términos, para evaluar precisamos la determinación de la instancia o modelo con la cual se ha de comparar la realidad objeto de evaluación, es decir, el patrón o criterio. La comparación implica, en consecuencia, de una parte el conocimiento, lo más completo posible, de la realidad (programas en nuestro caso) a evaluar, que necesariamente se apoya en un proceso previo de información (recogida de datos, medidas, etc.), y de otra, la descripción ajustada del término de comparación, del patrón, del criterio.

Por otra parte, la evaluación tiene un carácter instrumental. En educación, como en otros campos, evaluamos para algo que trasciende a la propia evaluación. Se evalúa para promocionar a los alumnos, o a los profesores, para adoptar tal o cual objetivo, programa o procedimiento didáctico, para mejorar, en suma la enseñanza y el aprendi-

zaje. La evaluación, pues, cumple unas funciones definidas en el contexto educacional; sus resultados son susceptibles de un uso determinado, de acuerdo con las funciones previstas. En general, podemos afirmar que se evalúa para tomar decisiones respecto al proceso o producto evaluados. Al juicio de valor en el que la evaluación consiste sigue una decisión.

En consecuencia, podríamos definir la evaluación de programas como el proceso sistemático de recogida, análisis e interpretación de información relevante y fiable para describir un programa educativo, o una faceta significativa del mismo, y formular un juicio sobre su adecuación a un criterio o patrón, que represente un valor aceptado, como base para la toma de decisiones sobre tal programa o faceta programática.

LAS FUNCIONES DE LA EVALUACIÓN

Dado el carácter instrumental de la evaluación, ya señalado, la determinación de las funciones y usos constituye una de las exigencias básicas, tanto en la perspectiva teórica —su conceptualización— como en la perspectiva de la práctica evaluativa. Cual sea la concepción del sentido último, del para qué de la evaluación y cual sea su concreción en los objetivos específicos de cada proyecto evaluativo, son factores altamente condicionantes, quizá determinantes, del efecto de la evaluación sobre el objeto evaluado y su contexto, es decir, del éxito o fracaso del instrumento evaluativo.

Convencionalmente se habla, por una parte, de la función de diagnóstico como base de la intervención educativa y de pronóstico y prevención; y, por otra, de la función de control y rendimiento de cuentas que supone informar a usuarios y consumidores de programas educativos, a las familias, a los gestores y decisores y, en último extremo, a la sociedad en general, acerca del valor o mérito del objeto evaluado —programas educativos— y certificar conocimientos y competencias generales y profesionales adquiridas en tales programas por los participantes, con garantía de objetividad.

Tradicionalmente estas dos categorías de usos de la evaluación han sido denominadas función formativa y función sumativa respectivamente. El primero en usar esta terminología fue Scriven (1967) quien caracterizó a la evaluación formativa como evaluación de procesos y a la evaluación sumativa como evaluación final, asumiendo, claro está, que la evaluación final de un programa se identifica con la evaluación de su producto. Poco después, Bloom (1969 y 1971) circunscribió la función formativa a la evaluación diagnóstica que, refiriéndose a los alumnos, trataría de determinar el estado de su aprendizaje en cualquier momento del proceso discente, mientras que la función sumativa sería la propia de una evaluación centrada en situar a cada alumno, respecto a su aprendizaje, en el punto que le corresponde en un continuo de competencia (del nivel más bajo —incompetencia— al nivel más alto —máxima competencia—). Finalmente, Stoffelebeam (1974) matiza estos conceptos asignando a la evaluación formativa el carácter de pro-activa y cuya función podría identificarse con la toma permanente de decisiones sobre los procesos del programa; y considerando a la evaluación sumativa como retro-activa con la función de exigir responsabilidades o rendición de cuentas, naturalmente como consecuencia de los resultados finales del programa.

Las diferencias entre ambos tipos de función responden básicamente a una doble consideración: el momento de su aplicación en el continuo del proceso educativo, por un lado, y la modalidad de las decisiones a que da lugar, es decir, el uso que se hace de la evaluación en cada caso, por otro lado. Ambas perspectivas combinadas determinan, a su vez, la faceta del programa en que con más probabilidad se centrará la evaluación (contexto, plan, proceso o producto).

En consecuencia, la evaluación de carácter formativo se centra en las fases previas al desarrollo del programa (la determinación de las necesidades a satisfacer, en el desarrollo mismo, en su plan y diseño y en los distintos estadios de su puesta en operación, lo que supone decisiones diagnósticas y de retroinformación a los planificadores y aplicadores del programa para el perfeccionamiento del mismo).

Más precisamente, la evaluación formativa proporciona la información adecuada para la mejora de los procesos educativos y a través de ellos, del producto.

Por su parte, la evaluación sumativa tiene generalmente carácter final y por tanto se centra básicamente en los resultados inmediatos, productos mediatos e impacto de los programas educativos.

Las decisiones que definen esta función evaluativa apuntan a la selección, clasificación y promoción de alumnos y a la certificación de conocimientos y competencias. Así mismo constituye la forma típica de evaluación para decidir la continuidad o terminación de un programa, modelo de organización, centro, etc. En resumen, controlar el valor de un programa (funcionalidad o pertinencia, eficacia y/o eficiencia).

Recientemente, y casi siempre en conexión con las discusiones en torno a los modelos de evaluación de instituciones educativas, se suele aceptar, en general de forma mecánica y acrítica, la existencia de dos modalidades de evaluación, la orientada al control, identificada en gran medida con la evaluación sumativa, y la orientada a la mejora, identificada con la evaluación formativa. La frecuente referencia al equiparar la evaluación formativa a la mejora y perfeccionamiento del proceso, y las escasas alusiones a decisiones de este tipo al caracterizar a la evaluación sumativa, puede ser una explicación plausible. No parece necesario documentar, por obvio, que la modalidad evaluativa orientada a la mejora es la que goza de una aceptación casi universal.

Sin embargo, la objeción más suave a esta posición es la de su simplicidad. La ausencia de matices a su vez convierte esta clasificación en un error. Una consideración atenta del concepto de evaluación y sus implicaciones pone claramente de manifiesto que la evaluación, toda evaluación, se justifica en su condición instrumental por su función optimizante del objeto evaluado y su contexto. Si esto es así, la evaluación orientada al control solo se justificará como evaluación educativa si mejora la educación. En esta perspectiva, la decisión de suprimir un programa, una disciplina, un plan de estudios e incluso una institución escolar como consecuencia de una evaluación sumativa o de control, se supone que mejora la educación. La diferencia con la evaluación formativa (de mejora) es que el perfeccionamiento, el proceso de optimización, es diferente, está pautado por decisiones distintas.

En otro trabajo (De la Orden, 1981) formulaba así lo que podría ser la base conceptual de la función general de la evaluación educativa propuesta, es decir, la optimización de la educación: Sea cual fuere la modalidad o tipo de evaluación adoptada en

cualquier contexto educativo, su influencia sobre el proceso y el producto de la educación es decisiva y afecta directa o indirectamente a su calidad, siempre claro está que la evaluación tenga consecuencias, es decir, siempre que se tomen las decisiones en función de las cuales se decidió la evaluación. Una evaluación sin consecuencias es inútil y puede ser perjudicial para el objeto evaluado y su contexto.

INFLUENCIA DE LA EVALUACIÓN EN LA CALIDAD DE LA EDUCACIÓN

La explicación de la influencia constatada en la evaluación en el proceso y el producto de la educación se apoya en el hecho de que las necesarias consecuencias de la evaluación modifican en algún sentido el objeto evaluado, en nuestro caso, el programa educativo, y también en alguna forma, la situación de las audiencias implicadas en el mismo. Estos efectos al ser percibidos por las audiencias como positivos o negativos —deseables o no deseables— condicionan el alto grado la conducta de estas audiencias en el contexto del programa educativo. Por otra parte, el elemento determinante del sentido de la modificación del objeto evaluado es, sin duda, la instancia de referencia. En efecto, los criterios indican lo que se exige del programa, del profesor, del alumno, para que el juicio de valor sobre el mismo sea favorable. La percepción por las audiencias críticas de esta situación guiará su conducta en la medida en que están interesadas realmente en el objeto evaluado y en las consecuencias de su evaluación. Parece evidente que los alumnos intentan satisfacer las exigencias (criterios) de los exámenes respecto a su aprendizaje; los profesores, en la medida en que la evaluación tenga consecuencias para su estatus y su profesión, tratarán de ajustar su docencia a las exigencias —criterios— del modelo evaluativo; los gestores de cualquier programa educativo tratarán de modelar los procesos y productos del mismo a las exigencias —criterios— de la evaluación. De forma general, por tanto, podemos afirmar que el modelo de comportamiento implícito en los criterios de evaluación, al constituir una expresión concreta de lo que se espera del programa realmente, puesto que es lo que exige de hecho, define de una manera efectiva los objetivos reales y operantes de la educación independientemente de lo estipulado formalmente y de que existan o no formulaciones expresas de los mismos.

En consecuencia, si consideramos el principio de la determinación teleológica de los medios (con toda la relatividad que se quiera, los fines y objetivos condicionan de alguna manera los medios para alcanzarlos), debemos concluir que la evaluación, al prescribir realmente los objetivos de la educación, determina, en gran medida, las características de la intervención y el aprendizaje resultante, lo que los alumnos aprenden y cómo lo aprenden, lo que los profesores enseñan y cómo lo enseñan, los contenidos y los métodos, en otras palabras, el producto y el proceso del programa. En efecto, si la calidad de la educación está vinculada a un sistema de coherencias en el proceso total de la educación, la virtualidad de la evaluación como estímulo o freno de la calidad educativa dependerá enteramente de las características del esquema evaluador: Si los criterios y los modos de evaluación, que determinan a su vez los objetivos reales, y en última instancia el producto del programa, son coherentes con los objetivos formalmente establecidos —o sobre los que existe un consenso implícito—, y, a través de

ellos, con el sistema de valores del que derivan, la evaluación actuará como el más poderoso factor de promoción de la calidad educativa al garantizar la congruencia y eficacia del sistema en su conjunto. Si, por el contrario, los criterios y modos de evaluación movilizan los procesos educativos hacia los objetivos no coherentes con las metas formalmente establecidas o implícitamente aceptadas del programa ni, en consecuencia, con el sistema axiológico en que se sustentan, la evaluación se constituye en el mayor obstáculo a la calidad.

CONDICIONES GENERALES DE LA EVALUACIÓN COMO INSTRUMENTO DE OPTIMIZACIÓN EDUCATIVA

Dado que la evaluación del aprendizaje puede influir positiva o negativamente en la calidad de la educación, resulta necesario determinar qué características debe poseer un modelo evaluativo para maximizar sus efectos positivos y minimizar o anular los negativos.

Evidentemente, la condición fundamental de un sistema de evaluación es que sea educativamente válido, lo que implica no solamente que los instrumentos de medida utilizados (test, pruebas, escalas de observación, procedimientos no formales, etc.) sean válidos en el sentido tradicional, sino también que exista un máximo de congruencia entre los criterios y modos de evaluación, la especificación de los objetivos del programa y el proceso de intervención. El elemento común entre criterios de evaluación, objetivos y proceso de intervención es el comportamiento cognitivo, afectivo y psicomotor de los alumnos claramente especificado. En efecto, los objetivos no son otra cosa que la especificación de este comportamiento o de la capacidad para el mismo en forma de resultados esperados del proceso educativo; los criterios de evaluación, lo que exige, constituyen, en conjunto, una muestra representativa de tal comportamiento; y el proceso didáctico se apoya en este comportamiento como base para su adquisición.

En consecuencia, la validez educativa de la evaluación del aprendizaje, como base de la calidad de la educación, significa:

- 1) Que los objetivos del programa sean claramente identificados y formulados sin ambigüedad. Para que la evaluación no se desvíe de los objetivos del programa, es necesario conocer con la máxima precisión tales objetivos.
- 2) Que los criterios de evaluación (lo que se exige del programa y cómo se exige) deben constituir una adecuada muestra representativa de los contenidos y conductas especificadas en los objetivos. Esta es la condición básica de la evaluación como palanca de la calidad de la educación. Todo profesor ha comprobado que los estudiantes aprenden mucho más y de manera más efectiva todo aquello que es objeto de evaluación y control. La evidencia experimental ha confirmado claramente esta hipótesis. Baste citar el ya antiguo estudio comparativo de Walker y Schaffarzick (1974) sobre evaluaciones del currículum en el que se puso de manifiesto que cuando las pruebas de rendimiento están en línea con las metas del programa innovador éste resulta más eficaz que el programa tra-

dicional con el que se compara. Por el contrario, cuando las pruebas no cumplen esta condición, resultan superiores los programas tradicionales. Por otra parte, el intento de alinear los criterios de evaluación con los objetivos contribuye decisivamente a clarificar a éstos y a resaltar su importancia en el pensamiento de profesores y alumnos. Las metas educacionales se hacen más definidas y significativas, incrementando la coherencia entre criterios de evaluación, objetivos formales y aspectos del aprendizaje considerados por alumnos y profesores como resultados educativos valiosos e importantes.

- 3) Que la forma, procedimientos e instrumentos de evaluación exijan los comportamientos especificados en los objetivos del modo más directo. Es decir, la técnica evaluativa empleada (observación de procesos, análisis de productos, informe, examen escrito tradicional, cuestiones de recuerdo o reconocimiento, pruebas con materiales o sin materiales, exámenes orales, etc.) debe estar directa y estrechamente relacionada con las características del aprendizaje o realización deseadas. Si un objetivo de un programa de Química es que los estudiantes manejen adecuadamente determinadas piezas del equipo del laboratorio, exigir la descripción oral de cómo manejaría tales piezas, o la elaboración de un esquema gráfico del proceso, son formas evaluativas menos válidas que requerir el desarrollo real del proceso en el propio laboratorio.
- 4) Que la evaluación sea fiable y objetiva en el sentido de que el azar o los errores instrumentales tengan un efecto mínimo en los resultados.

Una evaluación así concebida supone, en principio, el requisito esencial para asegurar el sistema de coherencias en que consiste la calidad de la educación. La relación directa y la correspondencia entre, objetivos formales, objetivos reales, procesos instructivos y resultados de un programa, de tal manera que resulten isomórficos, constituye la única garantía de un sistema con un alto grado de coherencia interna y externa con las expectativas y los resultados.

OPTIMIZACIÓN EDUCATIVA Y EVALUACIÓN DE LAS FASES DEL PROGRAMA

Una forma de incrementar la potencia optimizante de los programas educativos consiste en asociar la evaluación a su planificación, desarrollo y aplicación. Una de las características básicas de la evaluación es su carácter instrumental. La evaluación, como señalamos, más allá de un juicio de valor, se justifica porque permite el ajuste de las decisiones sobre el objeto o faceta del objeto evaluado, y este ajuste decisional supone en todas las circunstancias una mejora del objeto (programa, currículum, innovación de todo orden). La función evaluativa, como hemos señalado, cualquiera que sea el objetivo específico en cada caso, será siempre la optimización de las alternativas a la situación evaluada.

Cualquiera que sea el origen, objetivos, estructura y alcance del programa educativo, el paso inicial en evaluación consiste en determinar su necesidad. El análisis del contexto llevará a la selección de problemas, al diagnóstico situacional, que se plasmará en un conjunto de necesidades educativas priorizadas que habrán de ser satisfechas

a través del programa. La evaluación de necesidades se convierte en la fase previa de la planificación ya que su resultado será la identificación de las metas de la acción educativa. De esta forma, la evaluación de necesidades asegura la funcionalidad del programa —dimensión básica de la calidad— al facilitar su coherencia con las aspiraciones, expectativas y necesidades de individuos y grupos.

El proceso de planificación y desarrollo del programa debe ser también objeto de evaluación. En este sentido habrá de ser valorado el modelo de programa adoptado y el propio modelo de planificación y desarrollo que se sigue, así como la política de estímulo y de planificación de programas innovadores.

El producto de la planificación y desarrollo del programa es el plan o propuesta de acción que, a su vez, habrá de ser sometido a evaluación, antes de su puesta en operación. Esta fase de la evaluación debe asegurar la adecuación del diseño del programa y su coherencia interna, es decir, la congruencia inter e intra componentes del plan: objetivos, contenidos, estrategias, experiencias y actividades de aprendizaje y ambientes de intervención, materiales de apoyo y previsiones de evaluación. Un plan coherente es una condición necesaria, aunque no suficiente, de una intervención coherente, exigencia fundamental de la calidad educativa.

La evaluación de la puesta en operación del plan constituye una fase esencial para asegurar la calidad del programa y por tanto, la optimización educativa. Se trata de la evaluación de los procesos educativos generados por la aplicación del programa, a través de la observación, de pruebas piloto y seguimiento. La evaluación del proceso permitirá constatar en que grado es innovadora la conducta docente y discente y permitirá tomar decisiones adecuadas en orden a evitar que la inercia desdibuje el perfil de la innovación, como ha sido comprobado en diversos estudios. (Alexander, Murphy y Woods, 1996; Ekholm y Kull, 1996) Así mismo, la evaluación de los procesos educativos reales durante el periodo de ejecución del plan constituye un requisito imprescindible para la evaluación del producto del programa. Es absolutamente necesario saber si los resultados constatados se deben a los procesos planificados o a las formas de conducta, que sólo aparentemente se presentan como las exigidas por el programa.

La evaluación del producto o resultados educativos del programa apuntan directamente a determinar, por un lado, la eficacia, y por otro, la eficiencia del mismo y, por tanto, la optimización educativa. La eficacia del programa se constata bien en términos absolutos —grado en que el producto y el impacto de la acción educativa coinciden con los objetivos y productos previstos—, bien en términos relativos —grado en que los resultados de los procesos programados se acercan más o menos a los productos o resultados de los programas convencionales de educación o de otros programas diferentes—.

La evaluación, en suma, permite detectar los problemas educativos, su extensión, complejidad y relevancia a medida que se van produciendo. Esta constatación puede dar lugar a la planificación de cambios e innovaciones en programas concebidos como alternativas de solución a tales problemas. Una vez puesto en marcha el plan de trabajo, como hipótesis de solución, la evaluación formativa proporciona, desde el primer momento, información que permite juzgar en qué grado el programa o el cambio pla-

nificado produce los efectos previstos en el aprendizaje de los alumnos; y sugiere las eventuales modificaciones necesarias para el logro de los objetivos previstos.

En resumen, la evaluación de programas educativos, en la medida en que los criterios sean congruentes con los objetivos previstos y el diseño de evaluación sea el adecuado constituye la palanca más importante para canalizar los procesos y los productos educativos hacia objetivos valiosos y deseables social e individualmente y, por tanto, promueve la funcionalidad, la eficacia y la eficiencia de la educación y, en consecuencia, su calidad.

BIBLIOGRAFÍA

- Alexander, P.A., Murphy, P.K. y Woods, B.S. (1996). *Of the squalls and fathoms: Navigating the seas of educational innovation*. Educational Researcher. 25 (3) 31-39.
- Bloom, B.S. (1969). *Some theoretical issues relating to educational evaluation*. En R.W. TYLER (Ed.) «Educational Evaluation: new roles, new means». 68th Yearbook of the N.S.S.E. Chicago. University of Chicago Press.
- Bloom, B.S. y otros (1971). *Handbook on formative and sumative of student learning*. N. York, McGraw-Hill.
- De la Orden, A. (1982). *La Educación Educativa*. Buenos Aires, Docencia.
- De la Orden, A. (1997). *Evaluación y optimización educativa*. En H. Salmeron y otros (Eds.). Evaluación Educativa. Granada, Grupo Editorial Universitario.
- De la Orden, A. (1981). *Evaluación del aprendizaje y calidad de la educación*. En varios «La calidad de la educación». Madrid, C.S.I.C. Instituto de Pedagogía.
- Eckholm, M. y Kull, M. (1996). *School Climate and Educational Change*. EERA Bulletin 2 (2) 3-11.
- House, E.R. (1992). *Tendencias en evaluación*. *Revista de Educación*, n^o 299, p. 43.
- Scriven, M. (1967). The methodology of evaluation. AERA Monograph Series on Curriculum Evaluation, n^o1, pp. 39-83.
- Stufflebeam, D.L. (1974). *Metaevaluation*, The Evaluation Center, Kalamazoo, Western Michigan University Press.
- Walter, D.F. y Schaffarzick, J. (1974). *Comparing Curricula*. Review of Educational Research, n^o 44, pp. 83-111.