



## La delimitación del concepto

Aun cuando el término *metaevaluación* data ya de los años 60, en que fue acuñado por Michael Scriven, lo cierto es que el mismo ha despertado hasta ahora un escaso interés entre los especialistas en el campo de la evaluación. Pese a esta circunstancia -o quizá precisamente por ello- la coincidencia entre los diferentes autores, en lo que se refiere a la definición del concepto que ahora nos ocupa es, en los aspectos básicos, prácticamente total.

Así, por analogía con el término *evaluación*, la *metaevaluación* ha sido y es comúnmente considerada como aquella investigación sistemática cuyo objetivo es emitir un juicio acerca de la calidad o los méritos de una evaluación. La evaluación se convierte, de esta forma, en el objeto de estudio del metaevaluador (Stufflebeam, 1981, 1987; Scriven, 1967, 1991).

No obstante, el propio concepto de *evaluación* en nuestro idioma plantea considerables dificultades al analista. Y ello no sólo porque sea posible distinguir entre una acepción amplia -en la que evaluar es sinónimo de valorar- y una más precisa o científica -para la que la evaluación es aquel análisis sistemático que persigue la valoración de un determinado fenómeno- sino porque, con esta última, podríamos significar tantos tipos de evaluación como realidades u objetos evaluables.

Con no poca frecuencia, hallamos en las obras dedicadas al tema de la evaluación cierta vaguedad en el uso de la categoría. Parece en ellas que lo que se predique de una evaluación -por ejemplo, la del aprendizaje del alumno- es perfectamente aplicable a cualquier otra -por ejemplo, la de una política pública-. Pero esta imprecisión no es meramente lingüística: su trascendencia y sus repercusiones metodológicas resultan más que evidentes. Reflexiones tan genéricas son, por tanto, poco apropiadas para un análisis riguroso.

La especificación de la naturaleza del objeto, en otras palabras, del tipo de evaluación que pretenda analizarse, en la medida en que condiciona cuál deba ser el planteamiento analítico y los métodos apropiados para su estudio, se convierte así en un primer paso inexcusable para el estudioso que haya de acometer la realización de una metaevaluación.

En nuestro caso, la argumentación se centrará en el análisis de las evaluaciones de políticas públicas.

## El problema de los criterios

El principal problema con el que nos enfrentamos a la hora de abordar una metaevaluación es el de los criterios conforme a los cuales se evalúen las evaluaciones. De hecho, éste ha sido un tema recurrente tanto para los teóricos de la evaluación como para quienes se han dedicado a la metaevaluación. Aún son escasas las propuestas teóricas formuladas de manera específica para la metaevaluación. Buena muestra de ello es que algunas de las metaevaluaciones más significativas que se han llevado a cabo hasta la fecha no son sino aplicaciones de modelos teóricos elaborados para la realización de evaluaciones.

Las respuestas que ofrecen las distintas corrientes metodológicas a la cuestión de los criterios son bien distintas.

En líneas generales, los enfoques cuantitativistas tienden a hacer hincapié en los aspectos metodológicos y, en concreto, a hacer de las técnicas cuantitativas el paradigma de la "calidad" en evaluación. Por el contrario, las corrientes cualitativistas -a partir de la constatación de la naturaleza eminentemente política de la evaluación- suelen poner el énfasis en los aspectos éticos del proceso de investigación, en ocasiones incluso en detrimento de las cuestiones metodológicas.

El trabajo de metaevaluación de Bernstein y Freeman (1975) constituye un buen ejemplo de lo que en el llamado paradigma científico o hipotético-deductivo en evaluación se considera como "calidad" en la investigación evaluativa. A juicio de estos autores, la valía de una evaluación sólo quedaría garantizada con el empleo de procedimientos cuantitativos para la recogida de datos en el marco de un diseño de naturaleza experimental. Así pues, la metaevaluación tendría como objetivo calibrar la medida en que el estudio de evaluación se aproxime a este "modelo". En la línea de esta argumentación se inscribirían también autores como Rossi (1989), Campbell y Stanley (1966).

Pero han sido sin duda los autores *naturalistas* quienes mayores esfuerzos han dedicado al desarrollo teórico y a la práctica de la metaevaluación. Las corrientes naturalistas vinieron a poner de manifiesto las deficiencias de las argumentaciones científicas, esto es, la excesiva simplicidad de las formulaciones que equiparaban "calidad" con utilización de técnicas cuantitativas. La constatación de la naturaleza y las implicaciones políticas de la evaluación llevó a autores como House, Stake o Scriven, entre otros, a destacar la importancia de los aspectos éticos en el proceso de investigación evaluativa. Dichos aspectos pasaron a convertirse en el estándar con el que medir la valía de una evaluación: puesto que ésta ha de realizarse conforme a requerimientos de carácter ético (de respeto a la pluralidad de opiniones existentes, a la confidencialidad de los informantes, etc...), la metaevaluación habrá de centrarse en valorar el grado en que la investigación evaluativa cumple con esos requisitos.

En esta línea, Ernest House (1981, 1993), considera que la evaluación ha de ser no sólo veraz y creíble sino también "justa", tal y como se deduce tanto de sus estudios teóricos como de su crítica de la evaluación del programa "Follow Through" (1977). Según House la relevancia que ha ido adquiriendo en los últimos tiempos la práctica evaluativa hace necesario que ésta se asiente sobre principios tales como la igualdad de las distintas audiencias a la hora de tomar parte en el proceso de evaluación, el hecho de que las opiniones puedan ser expresadas libremente, con total autonomía, la imparcialidad en el juicio y la reciprocidad. El concepto de validez que House emplea no es el utilizado tradicionalmente en investigación: para él aquélla reside más bien en la medida en que la investigación contribuya a la justicia de la sociedad en general.

De igual forma, Robert Stake (1986), toma como referencia, en su estudio de la evaluación del programa "Cities-in-Schools", el modelo que él mismo ha dado en llamar "responsive evaluation". Aun cuando la traducción exacta del término "responsive" al castellano es la de "sensible" habitualmente se emplea la expresión de "evaluación responsiva" para referirse a este modelo. Como se aprecia en su obra *Quieting Reform*, los elementos que conforman su propuesta en evaluación son luego empleados como criterios de juicio en este trabajo de metaevaluación, en el que llama la atención sobre la relación que puede establecerse entre supuestos científicos simplistas e intereses políticos no orientados a la justicia y la solidaridad. El autor apuesta porque en la evaluación se integren las opiniones de todos aquellos que tienen un interés en el programa, los llamados "stakeholders", en definitiva, porque la evaluación sea "sensible" a los deseos de los distintos destinatarios. Se decanta por los estudios evaluativos de carácter naturalista, en los que el objetivo fundamental es la comprensión de la unicidad o singularidad del caso, no la búsqueda de conocimiento generalizable ni la elaboración de teoría.

Como señala muy acertadamente Patton (1986), el debate entre cuantitativistas y cualitativistas en evaluación no ha sido estéril. Ha servido precisamente para constatar la dificultad de hallar unos criterios o estándares universalmente válidos que guíen al metaevaluador en su análisis crítico de la evaluación. Los criterios habrán de ser establecidos en cada caso, por lo que dependen del

consenso al que sobre ellos lleguen los actores afectados a través de un proceso de negociación, concepto éste central en el enfoque naturalista. No obstante, considera que cualquier evaluación puede ser juzgada según unos criterios de referencia de carácter general, tales como su conveniencia, su utilidad, el grado en que sea factible, su credibilidad y su relevancia.

Sin lugar a dudas, M. Scriven (1991), ha realizado una de las más significativas aportaciones al campo de la metaevaluación. Como señala este autor, el metaevaluador puede servirse tanto de listas de estándares específicamente elaboradas para la metaevaluación como de otros catálogos de criterios formulados para la realización de evaluaciones. La "Meta-Evaluation Checklist" (MEC) incluye una serie de *items* -que el autor obtiene de diversas fuentes- y que apuntan a la conveniencia de que el estudio evaluativo sea comprensivo y comprensible, creíble, ético, que en él se expliciten los estándares de valor utilizados, así como su justificación, que sea factible de ser llevado a la práctica, suficiente pero no excesivamente preciso (la evaluación ha de ser rigurosa desde un punto de vista técnico, pero un exceso de rigor puede hacer también que se olviden cuestiones como la relevancia de la información proporcionada), sensible desde el punto de vista político, que atienda y responda a las necesidades de las audiencias, a las que será proporcionado, oportuno, seguro, válido y "rentable" en términos de coste-eficacia.

La "Key Evaluation Checklist" (KEC) fue en principio elaborada como referencia para la realización de evaluaciones pero también puede ser aplicada para la metaevaluación. Según Scriven, la KEC no debe ser considerada como una lista de *desiderata* sino más bien como un catálogo de *items* que deben ser analizados. Entre ellos se incluyen aspectos tales como el diseño del estudio, del personal de la evaluación, así como del calendario de la misma, la procedencia de los criterios empleados incluso para la elección del metaevaluador, los antecedentes y el contexto de la evaluación, la identificación de las audiencias, la estimación de los valores, costes y recursos, la valoración del proceso de investigación, los resultados del mismo, las alternativas existentes a la evaluación, la posibilidad de generalización de ésta y, por último, la significación global de la investigación evaluativa. En el informe, el metaevaluador puede adjuntar también recomendaciones sobre el uso o el destino de la evaluación.

Pero quizá lo más destacable de la obra de Scriven es su propuesta de que la metaevaluación se convierta en un imperativo profesional de los evaluadores, de que la evaluación, para decirlo en sus mismos términos, "empiece en casa". Como afirma Stufflebeam, "El objetivo de la metaevaluación es asegurar la calidad de los servicios evaluativos, evitar o enfrentarse a las prácticas ilegales o a los servicios que no son de interés público, señalar el camino para el perfeccionamiento de la profesión y promover una mayor comprensión de la empresa evaluativa" (1987: 52).

### **Una revisión crítica**

La propia naturaleza de la evaluación condiciona, obviamente, su análisis. En la medida en que la evaluación es, ante todo, investigación (algo obvio pero que con frecuencia parece olvidarse en ciertas propuestas), la tarea que habrá de centrar la atención del metaevaluador y que éste habrá de valorar será la *capacidad explicativa* de la investigación evaluativa que pretenda analizar. La determinación de los criterios que sirvan de referencia para la metaevaluación se realizará en función de los objetivos de la investigación y de la perspectiva de la que parta el analista. Como sugieren Schwandt y Halpern (1988), no resulta posible establecer unos estándares universalmente válidos para la realización de metaevaluaciones: dichos estándares dependerán de lo que el investigador estime que es la calidad en evaluación.

La metaevaluación, al igual que la evaluación es, en esencia, una actividad pluridisciplinar (Alvira, 1991). Ello quiere decir que ninguna disciplina científica puede arrogarse la pretensión de convertirse en la "ciencia evaluadora" por excelencia. Sin embargo, la primacía de materias como la Pedagogía -cuya orientación hacia los problemas de orden moral es clara- o la Psicología en el

análisis y la evaluación de políticas educativas, uno de los ámbitos en los que el desarrollo de la evaluación es más notorio, ha determinado las perspectivas y los criterios dominantes tanto en las propuestas teóricas sobre evaluación como en la metaevaluación misma. La evaluación de políticas educativas desde el enfoque de políticas públicas sigue siendo aún una excepción.

Como hemos visto, algunas reflexiones acerca de los criterios que han de emplearse en la metaevaluación han girado en torno a cuestiones de carácter metodológico -una evaluación será tanto más válida cuanto mayor sea la utilización de técnicas cuantitativas-; otras, en cambio, se han ocupado principalmente de la ética de la evaluación -ésta ha de ser justa, ha de promover la participación de todos los sectores afectados por la política pública-; otras, por último, se han centrado en las normas que deben regir la práctica profesional -la metaevaluación constituye así la garantía de calidad de la actividad evaluadora- o, han tratado diversos aspectos en unas propuestas que cabría calificar de eclécticas -como es el caso de Patton o, más singularmente, de Scriven.

En Ciencias Sociales, tan simplista sería decir que sólo es científica, sistemática y válida aquella evaluación que se realice según un diseño experimental o cuasi-experimental y que emplee métodos cuantitativos para la recogida y el tratamiento de los datos como afirmar que el hecho de que se obstaculice la participación de ciertos sectores en la misma plantea, ante todo y principalmente, un problema de carácter ético cuando, en realidad, tiene más que ver con los aspectos metodológicos de la investigación o, si cabe, con la regulación de la práctica profesional del evaluador.

Desde las últimas tres décadas, las sociedades democráticas han venido asistiendo al paulatino desarrollo de la evaluación, motivado por la creciente necesidad sentida en el Estado capitalista avanzado de valorar los efectos de las políticas públicas. Efectivamente, la evaluación ha pasado de ser una actividad periférica a convertirse en un tipo de investigación que despierta cada día un mayor interés entre especialistas de muy distintas disciplinas, que es respaldada e impulsada por los gobiernos y, de esta forma, va alcanzando progresivamente mayores cotas de profesionalización e institucionalización. El primero de estos aspectos se ha concretado en la existencia de una comunidad de expertos en la materia que posee sus propios foros de debate y discusión, sus propios estándares de crítica y revisión del conocimiento. La institucionalización se ha manifestado en el surgimiento de un gran número de organizaciones públicas o privadas cuyo principal cometido es, precisamente, la evaluación (House, 1993).

La importancia que la evaluación ha ido adquiriendo tanto en medios académicos como institucionales no reside sólo en su carácter de investigación científica -como en principio cabría pensar- sino muy especialmente, en sus notables implicaciones políticas. De hecho, buena parte de las más recientes elaboraciones teóricas en el campo de la evaluación giran entorno a esta circunstancia: a la dimensión política de la evaluación. Su reconocimiento es, sin duda, una de las más singulares aportaciones de las corrientes naturalistas. Son varias las razones por las que se puede subrayar la relevancia política de la investigación evaluativa:

i) Las evaluaciones tienen como objeto de estudio unos elementos -las políticas públicas- que son propuestos, elaborados, debatidos, financiados y aprobados *en el proceso político*. (Weiss, 1987; Palumbo, 1987).

ii) La evaluación reproduce la dinámica, los rasgos del contexto político en el que surge, y los procedimientos, las estructuras y los actores que conforman éste inciden notablemente en su realización.

iii) El impulso hacia su consolidación procede, fundamentalmente, de los poderes públicos, en forma de institucionalización y/o de financiación. Los gobiernos ven en ella un elemento de legitimación (el respaldo que suponen los informes de expertos)), al tiempo que una oportunidad

para el control. Como afirma House, el gobierno legitima la evaluación y la evaluación legitima la actuación gubernamental (1993).

iv) La evaluación constituye además una actividad intrínsecamente política, desde el momento en el que en ella se ven involucrados intereses contrapuestos, políticamente relevantes, que luchan por obtener una mayor cuota de poder e influencia en la negociación.

v) Los criterios conforme a los cuales se juzga la política, el diseño mismo de la evaluación no dejan de tener naturaleza política por cuanto implican adoptar un posicionamiento u otro frente a cuestiones tan controvertidas como el acceso de determinados grupos a la evaluación, la propiedad de la información producida, etc... (Weiss, 1987).

vi) Las evaluaciones se realizan con el propósito -tácito o explícito- de informar e influir, en un sentido o en otro, en el proceso de toma de decisiones políticas. No quiere ello decir que se orienten siempre a la mejora de la política o el programa en cuestión. Con frecuencia, se emplean como medio de rentabilizar el "esfuerzo político" o, incluso, se convierten en un arma arrojada entre opciones de distinto signo. La evaluación produce información y la información proporciona poder. Son pues especialmente proclives a ser instrumentalizadas, utilizadas para fines políticos o partidistas.

Este último punto ha de ser matizado. El hecho de que parezca, a la luz de lo anteriormente señalado, que las evaluaciones tienden o pueden ser instrumentalizadas políticamente, no debe llevar al estudioso a partir de una aproximación *patológica* en su análisis. El que una evaluación no sea utilizada tal y como fue previsto o el que su diseño técnico no resulte ser el más apropiado para los objetivos propuestos no quiere decir necesariamente que detrás de todo ello haya una más o menos evidente intencionalidad política.

Varios han sido los autores que han analizado el desarrollo de la evaluación en las sociedades avanzadas y, más concretamente, las pautas y tendencias de la política de las autoridades gubernamentales por lo que respecta a la evaluación (Weiss, 1991; Norris, 1990). Con frecuencia, estos estudios han servido, al tiempo, para denunciar la situación en la que se encuentra la investigación evaluativa financiada con fondos públicos, su falta de credibilidad ante la injerencia cada vez mayor del gobierno en la elaboración de la agenda y el contrato de evaluación (esencialmente estableciendo, cuando no imponiendo, restricciones al uso y la difusión de la información).

Creemos que las características reseñadas hacen de la evaluación un objeto de estudio sumamente interesante para el científico social. En virtud de anteriores consideraciones, parece lógico que la metaevaluación no pueda quedar reducida a un mero examen de la calidad técnica de la evaluación. La influencia de lo político ha de ser considerada en cualquier análisis metaevaluativo que pretenda comprender la evaluación en toda su complejidad. La relevancia de la perspectiva politológica en cualquier análisis metaevaluativo se hace incuestionable. La metaevaluación no sólo permite aprehender las claves de la política evaluativa gubernamental, sino que nos remite a una cuestión trascendental: el control que puedan ejercer los poderes públicos sobre los resultados de estas investigaciones.

---

## BIBLIOGRAFIA

Alvira, F., (1991), *Metodología de la evaluación de programas*, Madrid: Centro de Investigaciones Sociológicas.

Bernstein, I. y Freeman, H.E., (1975), *Academic and Entrepreneurial Research: Consequences of Diversity in Federal Evaluation Studies*, Nueva York: Russell Sage.

Campbell, D.T. y Stanley, J., (1966), *Experimental and Quasi-Experimental Design for Research*, Chicago: Rand McNally.

House, E.R., et. al., (1977), *No Simple Answer: Critique of the "Follow Through" Evaluation*, Urbana: CIRCE/University of Illinois.

House, E.R., (1981), *Evaluating with Validity*, Newbury Park: Sage.

- (1993), *Professional Evaluation*, Newbury Park: Sage.

Norris, N., (1990), *Understanding Educational Evaluation*, Londres: Kogan Page.

Palumbo, D.J., (1987), *The Politics of Program Evaluation*, Beverly Hills: Sage.

Patton, M.Q., (1986), *Utilization-Focused Evaluation*, Newbury Park: Sage.

Rossi, P.H., Freeman, H.E., (1979), *Evaluation: A Systematic Approach*, Beverly Hills: Sage.

Scriven, M., (1967), "The Methodology of Evaluation", *Perspectives on Curriculum Evaluation*, AERA Monograph Series on Curriculum Evaluation, 1, Chicago: Rand McNally.

- (1991), *Evaluation Thesaurus*, Newbury Park: Sage.

Stake, R.E., (1986), *Quieting Reform*, Urbana/Chicago: University of Illinois Press.

Stufflebeam, D.L. y Shinkfield, A.J., (1981), "Meta-Evaluation: Concepts, Standards and Uses" en R. Berk., ed., *Educational Evaluation Methodology: The State of the Art*, Johns Hopkins.

- (1987), *Evaluación sistemática. Guía teórica y práctica*, Madrid: Paidós/MEC.

Schwandt, T.A. y Halpern, E.S., (1988), *Linking Auditing and Metaevaluation*, Newbury Park: Sage.

Weiss, C.H., (1991), "Evaluation Research in the Political Context" en M.W. McLoughlin y D.C. Phillips, eds., *Evaluation in Education: A Quarter Century*, Chicago: University of Chicago Press.