

LA EVALUACIÓN DE PROGRAMAS EDUCATIVOS: CONCEPTOS BÁSICOS, PLANTEAMIENTOS GENERALES Y PROBLEMÁTICA

Ramón Pérez Juste

UNED

RESUMEN

La necesaria relación entre la Pedagogía y su objeto de conocimiento, la Educación, es objeto de análisis a la hora de plantear las reflexiones pertinentes sobre la evaluación educativa en general y la evaluación de programas educativos en particular.

La evaluación de programas es analizada en la doble perspectiva de la mejora de la acción educativa de profesores y educadores, por un lado, y de la investigación evaluativa por otro.

El autor hace objeto de análisis tres conceptos básicos: programa, evaluación y evaluación de programas, planteando una propuesta metodológica especialmente pensada para la evaluación de programas educativos.

Por último, se ofrece una reflexión sobre algunos temas especialmente críticos, tales como el papel de los expertos, la factibilidad de la evaluación, los aspectos éticos de la evaluación o la comparabilidad de los resultados.

SUMMARY

The necessary connection between Pedagogy and education, as its knowledge object, is the theme of analysis when we put forward accommodated studies on educational evaluation, in general, and on educational programmes evaluation, in particular.

The evaluation of programmes is analysed, on the one hand, with the double approach of the improvement of educational activity implemented by teachers and educators and, from the other one, of evaluative research.

The author offers as object of analysis three basic concepts: programme, evaluation and programme evaluation, by suggesting a methodological proposal particularly directed towards evaluation of educational programmes.

Finally, a consideration is offered on some themes particularly critical, such as the expert's role, the evaluation feasibility, the ethical implications of evaluation or the results comparability.

I. LA EDUCACIÓN COMO OBJETO DE CONOCIMIENTO E INVESTIGACIÓN

La educación es una actividad radicalmente humana, sistemática, orientada al perfeccionamiento, a la mejora de las personas, de *cada una* de las personas, por medio de acciones intencionadas de los educadores, generalmente concretadas en *planes* o *programas*. Como en cualquier otro ámbito de la realidad natural o social, el ser humano se ha propuesto su conocimiento, bien sea entendido como una *comprensión profunda* de su naturaleza, planteamientos, procesos y resultados, bien como una *explicación* de los mismos mediante leyes más o menos generales, a ser posible de naturaleza causal por ser éste el medio más adecuado para su control y predicción, bien como instrumento al servicio de la *transformación* liberadora de la humanidad¹.

No es mi intención afrontar, con una mínima profundidad, los ya viejos debates sobre la naturaleza de la educación de cara a valorar la adecuación de los diversos paradigmas para producir conocimiento relevante a través de la reflexión y de la investigación. Sin embargo, no es menos cierto que tal planteamiento es básico y que, al menos de forma implícita, subyace en las decisiones que toman los profesionales de la educación o de la Pedagogía cuando abordan este tipo de tareas.

Dejemos clara, eso sí, la compleja y diversa realidad de la Educación como objeto de conocimiento e investigación, de cara a justificar nuestras propuestas metodológicas en un campo de tanto porvenir —si se aborda adecuadamente— como el de la investigación evaluativa o evaluación de programas.

I.1. La educación como arte

En el ya añejo discurso sobre estas cuestiones, una de las propuestas es la de considerar a la educación como un *arte*; y no es descabellada la idea sino todo lo contrario, cuando nuestro foco de atención es la relación privilegiada, por única, profunda, intensa y continuada, entre un educador y un educando o todo un grupo de niños o jóvenes en su etapa de formación.

¹ Entiéndase esta distinción a los efectos meramente expositivos. Difícilmente se puede mantener la incomunicabilidad de estos tres grandes planteamientos sobre la finalidad del saber.

Sin embargo, el hecho de que esa relación sea irrepetible y, en esa misma medida, se acerque a la existente entre el artista y la creación de su obra², no impide en modo alguno la posibilidad de que el *artista / educador* encuentre en el saber pedagógico principios, enunciados, propuestas... con los que abordar con mayores posibilidades de éxito una tarea tan apasionante como difícil. Y es que el educador puede aplicar a su situación concreta esos saberes pedagógicos que le indican las regularidades, los grados de dispersión, las tendencias, las secuencias, los ciclos..., al igual que en la pintura, *arte* por excelencia, el pintor aplica de forma creativa los conocimientos ¿científicos? sobre proporciones, perspectiva, contrastes, coloridos ... imprimiéndole después su huella personal.

1.2. La educación como ciencia

Precisamente, la educación puede ser objeto de un conocimiento al menos con pretensión de científico, en la medida en que se le aplique un método riguroso al servicio de metas relevantes.

No insistiremos más el manoseado tema de los paradigmas, pero sí dejaremos constancia de que no todos aceptan —en realidad son cada vez menos los que lo hacen— que la única ciencia posible sea la derivada de la aplicación del método hipotético-deductivo-experimental.

Por tanto, en la medida en que el objeto del saber existe, que disponemos de *métodos* rigurosos y que están a la espera múltiples *problemas* y *preguntas* por responder, cabe pensar en una Pedagogía como ciencia, si no con el estatus de las Ciencias Naturales sí con el propio de las de su ámbito, las Ciencias Humanas o Sociales.

En su momento³ señalábamos como una de las mayores objeciones para aceptar la existencia de una ciencia de la Educación, junto a la relativa al control en los diseños, la posibilidad de establecer leyes generales. Sin embargo, como allí indicábamos, hay autores que defienden tal posibilidad en la medida en que tal concepto no es de significación unívoca. Acudíamos, y lo hacemos de nuevo, a la afirmación de M. Sherwood (*The logic of Explanation in Psychoanalysis*. Academic Press, 1969), que recoge varios tipos de generalizaciones aceptables, reseñando los siguientes:

2 Evidentemente, el símil es imperfecto: mientras el artista trabaja con un objeto dócil, estable, que no reacciona a sus intervenciones (o que, cuando lo hace, se comporta de conformidad con las leyes inmutables de la Naturaleza), en educación el objeto / sujeto es capaz de atender o no, de aceptar o no, de poner su esfuerzo o no, y hasta de rechazar las propuestas de sus educadores... El objeto humano es sujeto, con capacidad de tomar decisiones y de comprometerse con ellas con mayor o menor intensidad y esfuerzo. Usamos, por tanto, el símil únicamente en lo que vale: la dimensión creativa de esa relación, donde las reglas, leyes o principios únicamente sirven como referencia y deben ser interpretadas y aplicadas en cada momento en función de las circunstancias.

3 En la Memoria para la oposición de la plaza de profesor agregado en la Universidad de Santiago de Compostela, en 1980.

- «1 *AB se verifica (siempre)*
 2 *En las ocasiones $T_1 \dots T_n$, AB se verifica.*
 3 *En las ocasiones $T_1 \dots T_n$, en las condiciones $C_1 \dots C_k$, AB se verifica.*
 4 *En las ocasiones $T_1 \dots T_n$, en las condiciones $C_1 \dots C_k$, para los individuos $P_1 \dots P_l$, AB se verifica.*
 5 *AB se verifica en p % de las condiciones (lo que sería la afirmación 1 con la adición de una condición de probabilidad).*
 6 *En las ocasiones $T_1 \dots T_n$, en las condiciones $C_1 \dots C_k$, para los individuos $P_1 \dots P_l$, AB se verifica en el p % de casos.*
 Siendo AB la relación entre las variables A y B, resulta que la situación 6 es la forma más débil de generalización».

Sin embargo, aun esta «forma débil» de generalización no deja de ser un conocimiento útil para los educadores y, desde luego, adaptado a las características del campo objeto de estudio, donde la diversidad de los seres humanos, y hasta su irrepetibilidad, deben siempre ser tomadas en consideración.

1.3. La educación como tecnología

Tampoco es una opción descartable la consideración del estudio de la Educación como una *tecnología*. No olvidemos que según Bunge⁴, la *tecnología* nace de aplicar la investigación a mejorar las condiciones de vida del ser humano:

«La ciencia como actividad —como investigación— pertenece a la vida social; en cuanto se aplica al mejoramiento de nuestro medio natural y artificial, a la invención y manufactura de bienes materiales y culturales, la ciencia se convierte en tecnología. Sin embargo, la ciencia se nos aparece como la más deslumbrante y asombrosa de las estrellas de la cultura cuando la consideramos como un bien en sí mismo, esto es, como un sistema de ideas establecidas provisionalmente (conocimiento científico) y como una actividad productora de ideas (investigación científica)».

Parece evidente que en el marco de la educación hay mucho *objeto* especialmente apto y adecuado para un saber *tecnológico*: todo aquello que tiene que ver con la planificación, la programación, el diseño de cursos, de recursos, de materiales, de programas —se destinen al logro de saberes, a la capacitación en destrezas y estrategias o a la creación, desarrollo y mantenimiento de valores—, la organización escolar, los sistemas de evaluación... entran de lleno en es apartado.

4 Bunge, M. (1981). *La ciencia: su método y su filosofía*. Buenos Aires: Ediciones Siglo Veinte, p. 9.

La investigación, y la evaluación, se convierten aquí en instrumentos de enorme importancia para su mejora tanto en la fase de diseño como de implementación, desarrollo y resultados.

1.4. La educación: un campo heterogéneo y multidisciplinar

Las anteriores consideraciones, sobre cuya «corrección» no es fácil decidir dada su indemostrabilidad, si nos llevan, al menos a juicio de quien esto escribe, a una importante conclusión: la educación, una realidad compleja y diversa donde las haya, puede ser considerada desde perspectivas muy diferentes, probablemente todas ellas válidas y necesarias. En cada caso, eso sí, el estudioso deberá saber en qué plano se sitúa y actuar de modo coherente.

Pero hay una segunda e importante cuestión: ¿es la Educación un objeto disciplinar homogéneo? Probablemente muchas de las disputas metodológicas podrían desaparecer o, al menos, atenuarse sensiblemente, si se llegara a una respuesta negativa, esto es, a distinguir entre el estudio científico de la Educación, centrado en la Pedagogía, y las diversas Ciencias de la Educación, encargadas de abordar objetos relacionados con y necesarios para una educación de calidad. Es más, que dentro de la propia Pedagogía como ciencia de la Educación cabe hablar de «objetos» susceptibles de ser abordados multi o pluridisciplinariamente, lo que exigiría que, para alcanzar un acercamiento mínimo y suficiente a tan compleja realidad, se acudiera a unos planteamientos metodológicos acordes con la complejidad y diversidad de su objeto.

Nuestra posición en este punto se concreta en una respuesta negativa a la pregunta y, en consecuencia, en la aceptación de la necesidad de evitar cualquier manifestación reductiva en lo concerniente a la metodología cuando de elaborar conocimiento se trate.

Las dos cuestiones anteriores nos conducen a una misma respuesta en términos de metodología: cualquier propuesta metodológica que se autolimita a una única metodología, además de estar abocada al fracaso por inadecuada a la realidad objeto de estudio, no ofrecerá los frutos apetecibles al dejar fuera de su foco de atención aspectos relevantes de esa compleja realidad. Lo mejor es aplicar el *principio de complementariedad metodológica*, que va incluso más allá de la proposición de Cook⁵ conocida como *multiplismo*⁶. El principio de complementariedad reconoce que ante realidades complejas, el mejor acercamiento no reside en la elección de una entre varias opciones, ni siquiera en elegir la mejor, sino en abordarlas desde enfoques metodológicos complementarios⁷, siempre que lo permita la formación de los estudiosos, el tiempo y los recursos disponibles.

5 Cook, T.D. (1985). Postpositivist Critical Multiplism. En L. Shotland y M.M. Marck (Edits). *Social Science and Social Policy*. Beverly Hills: Sage Pub.

6 En efecto, mientras el *multiplismo* se inscribe en la búsqueda de salidas a las limitaciones del positivismo, el principio de complementariedad metodológica es capaz de dar cabida, en simultáneo o sucesivo, a los métodos cualitativos y cuantitativos, además de a los históricos, filosóficos y comparativos.

7 Podemos encontrar una concreción del *principio de complementariedad metodológica* en el discurso de ingreso del profesor García Hoz en la Real Academia de Ciencias Morales y Políticas. Allí reconoce la riqueza que representa conocer la evolución (historia) de un fenómeno, sus aspectos cuantitativos y cualitativos, su situación en relación con la de otros contextos (sean otros sistemas educativos, otros países...).

1.5. La evaluación de programas y la mejora de la actuación pedagógica

La evaluación de programas es una actividad metodológica que admite, al menos a los efectos expositivos, dos polos separados por un continuo de actuación: la acción reflexiva, ordinaria, de cada profesor o educador sobre su *programa*, entendido como *plan* al servicio del logro de sus metas educativas, y aquella otra llevada a cabo por expertos mediante la aplicación rigurosa de metodologías de diferente naturaleza y alcance, destinada a la evaluación de proyectos y programas de intervención social —en educación, formación, salud, ocio, empleo...— de gran amplitud, complejidad y duración.

En el campo pedagógico, caracterizado por un afán cientifista, los profesores universitarios nos hemos ocupado muy poco de la creación en el profesorado de una auténtica cultura evaluativa sobre sus programas, los que podríamos denominar *programas ordinarios*, a pesar de que es en la aulas, en cada una de ellas, donde se da en mayor grado y con superior intensidad la acción educativa, el hecho educativo.

Crear una auténtica *cultura evaluativa* tendría unos frutos espectaculares; tal cultura supone la existencia de aptitud y de actitud positiva, a lo que se uniría el conocimiento, la información relevante para la mejora que ofrece la evaluación. Al caer en tierra bien abonada por la actitud favorable, una información de tal naturaleza daría sus mejores frutos en el compromiso con las decisiones de mejora de los planes, proyectos, procesos y resultados.

Tal vez sería ésta la mayor aportación a la mejora de la Educación; sin embargo, y en el marco del hacer científico, un adecuado estudio de las actuaciones de los profesores podría ser organizado, sistematizado y tratado desde una perspectiva cercana a la replicación⁸, dando lugar al establecimiento de enunciados y conclusiones de cierto nivel de generalidad.

La evaluación de programas, pues, puede servir a esos dos grandes objetivos: la mejora de la calidad de la acción educativa dentro de cada aula y de cada centro, y la creación de teoría, aunque, por prudencia y rigor, demos aquí a tal expresión un nivel atenuado.

1.6. La evaluación de programas como investigación evaluativa

Pero la evaluación de programas es una actividad nacida en otro contexto y orientada hacia otro tipo de objetivos; fundamentalmente nos situamos ante una evaluación externa de programas educativos, realizada por una sola vez, desde el exterior, por expertos, en ocasiones de diversos ámbitos —disciplinarios, metodológicos, sociales— respondiendo a lo que conocemos como *investigación evaluativa*. A ella se han dedicado las mejores de nuestras fuerzas ya que se trata de una actividad lucida y recompensada en términos profesionales y, en ocasiones, económicos.

⁸ Vid. al respecto los planteamientos de M. Sidman sobre replicación directa y sistemática en *Táctica de la investigación científica*, Barcelona: Editorial Fontanella, 1978.

Sin embargo, son muchos los que dudan de su valor práctico, de su utilidad, aunque sea forzoso reconocer que los grandes programas institucionales, al igual que las innovaciones de cualquier tipo, suelen llevar unida como una realidad socialmente exigida, una *evaluación de programas*, destinada por lo general a rendir cuentas, a ganar apoyo público (Nevo: 1986) y, al menos teóricamente, a la mejora del programa a través de la toma y aplicación de las adecuadas decisiones de mejora⁹.

Lo que se debe pedir a estas evaluaciones es que se realicen con el rigor propio de toda investigación evaluativa, siguiendo hasta donde sea posible y la naturaleza de la situación lo aconseje y lo permita, los estándares propuestos por el Joint Committee¹⁰ y, de modo concreto, en el ámbito de la metodología, aplicando el *principio de complementariedad metodológica* con la mayor riqueza posible.

Vale la pena señalar, por último, que en este ámbito los pedagogos estamos afrontando evaluaciones que no siempre se relacionan en sentido pleno con la educación y que, en consecuencia, muchas de las reflexiones anteriores —y de las que siguen— sólo les serán de aplicación de forma limitada. La ampliación de nuestro campo tanto de intervención como de su evaluación puede responder a causas diversas, no siempre concordantes pero sí perfectamente explicables y muchas defendibles y asumibles. Entre éstas conviene destacar, y apoyar:

- a) Aquellas que derivan del reconocimiento profesional de personas que trabajan en el ámbito educativo en general y pedagógico en particular, que les hace ser llamados para que desempeñen, con el rigor y la profesionalidad acreditados, la actividad evaluativa de programas no estrictamente educativos ni pedagógicos.
- b) La ampliación del campo de actuación de los pedagogos desde el estrictamente educativo al más amplio de la formación, de gran porvenir e interés en el marco de la *educación permanente*.

Entre las demás, es preciso tomar conciencia de que, en ocasiones, los profesionales de la educación abordamos ámbitos más o menos relacionados con el nuestro como consecuencia, tal vez en simultáneo pero a veces como fenómenos independientes, de, por un lado, nuestra falta de capacidad —o de interés— para desarrollar lo que debe ser nuestro campo de atención preferente y, por otro, del empuje de profesionales de ámbitos cercanos que se hacen cargo de parcelas teóricamente propias de los pedagogos, ya sea por nuestro abandono, ya porque nuestras aportaciones no alcanzan la calidad y los resultados que se esperaban, ya por la indefinición de los

9 Forzoso es reconocer que, con frecuencia, y dados los evidentes condicionamientos, estas evaluaciones pueden caer en lo que Stufflebeam y Shinkfield denominan *cuasievaluaciones*, a las que tratan con una enorme dureza al afirmar de quienes las practican que «utilizan la evaluación de una forma inadecuada para lograr sus objetivos», señalando a continuación lo siguiente:

«Puede que recopilen la información rigurosamente, pero no revelan las verdaderas conclusiones, o lo hacen de un modo selectivo, o incluso las falsifican, engañando tanto a la gente como a sí mismos» (págs. 67 y 68). Vid. Stufflebeam, D.L. y Shinkfield, A.J. (1987). *Evaluación sistemática. Guía teórica y práctica*, Barcelona: Paidós-MEC.

10 JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (1994). *The Program Evaluation Standards*, Sage.

respectivos ámbitos de actuación, por no señalar sino algunas de las posibles explicaciones.

Dejemos constancia aquí de que las reflexiones y propuestas que siguen quieren centrarse en ámbitos estrictamente educativos, no por desdén o menosprecio de los demás sino por considerarlos como el campo prioritario de los profesionales de la Pedagogía.

2. CONCEPTOS BÁSICOS

Señalábamos más arriba que el acto educativo tiene mucho de *relación artística* entre el educando y el educador, que trata de aplicar, con la debida prudencia y adecuación¹¹, los conocimientos científicos, las reglas técnicas y la propia experiencia al aquí y ahora de tal relación.

Sin embargo, el que cada acto educativo tenga o pueda tener ese carácter cuasi artístico no es óbice para que sea conveniente, e incluso necesaria, una acción reflexiva y meditada en la que el educador planifica su acción educativa para el alumno o el grupo, fijando objetivos, seleccionando medios, diseñando las condiciones adecuadas y estableciendo criterios y pautas de valoración¹². A tal actividad suele dársele el nombre de programación y el de programa a un concepto de muy amplio espectro, pero que, en el campo educativo, no es sino un *plan de acción*, por tanto una actuación *planificada, organizada y sistemática*, al servicio de metas educativas valiosas.

A los efectos de la evaluación de programas conviene clarificar los conceptos fundamentales, destacando entre ellos los de *programa, evaluación y evaluación de programas*, a los que nos referiremos a continuación.

2.1. Programa

Como hemos señalado, en el campo pedagógico la palabra *programa* se utiliza para referirse a un *plan sistemático* diseñado por el educador como medio al servicio de las metas educativas. Tanto a los efectos de su elaboración como de su posterior evaluación —dos procesos que deberían guardar armonía y coherencia— deben tomarse en consideración algunas importantes cuestiones:

11 No olvidemos que la *educación personalizada* tiene uno de sus principios básicos en la adecuación de la intervención educadora a la diversidad de los educandos, a sus características específicas, a su individualidad —concepto preferido por Allport— tanto en sus aptitudes como en las diferentes dimensiones de su personalidad —desde sus intereses y motivaciones a su modo de ser— y en su capacidad para abordar los diferentes tipos de aprendizaje.

12 La tarea específica del educador consiste en facilitar el éxito de sus alumnos ayudándoles a alcanzar los objetivos planificados. De todos es conocido el planteamiento de Bloom cuando diferencia entre estructura de la disciplina y la estructura del proceso de aprendizaje, «*que debe permitir al estudiante pasar con éxito de una fase de aquél a otra*», concluyendo que «*la estructura de la enseñanza y el aprendizaje deben basarse en consideraciones pedagógicas y no tienen por qué reflejar el punto de vista del especialista sobre su campo de estudio*» (Bloom: 1981, p. 30).

- a) Todo programa debe contar con unas *metas y objetivos* que, obviamente, han de ser *educativos*. Esta afirmación parece una obviedad, pero no son pocas las ocasiones en que la evaluación de programas se «olvida» de plantearse esta tan importante como básica cuestión.
- b) Metas y objetivos deben estar acomodados a las características de los destinatarios en su contexto de referencia y ser asumidos como propios por los agentes del programa.
- c) Tanto a la hora de su implantación como de su evaluación, en este caso en cuanto *variable independiente*, el programa ha de estar claramente especificado y detallado en todos sus elementos fundamentales: destinatarios, agentes, actividades, decisiones, estrategias, procesos, funciones y responsabilidades del personal, tiempos, manifestaciones esperables, niveles de logro considerados *a priori* como satisfactorios ... De no ser así adolecerá de limitaciones en uno de los criterios para su posterior evaluación: la *evaluabilidad*.
- d) Ha de incorporar un conjunto de *medios y recursos* que, además de ser *educativos*, deben ser considerados como *suficientes, adecuados y eficaces* para el logro de metas y objetivos.
- e) Necesita de un sistema capaz de apreciar tanto si metas y objetivos se logran como de poner de relieve, en caso de que así no sea, o de que lo sea a niveles insuficientes o insatisfactorios, dónde se producen las disfunciones y carencias y, a ser posible, las causas de las mismas (objetivos muy elevados, aplicación inadecuada de los medios, ambiente incoherente, relaciones humanas insatisfactorias, tiempos insuficientes, rigidez en las actuaciones, momentos inadecuados, efectos no planeados...).

Obviamente, tal sistema no es otro que el de la *evaluación*, que debe formar parte de los medios —evaluación *en* el programa— pero que ha de permitir su propia mejora —evaluación *del* programa, como un componente fundamental de su *mejora continua*— además de constituirse, por su propio potencial, en una actividad de mejora profesional¹³.

2.2. Evaluación

Aunque es difícil ofrecer una definición que cuente con suficiente consenso y respaldo, a los efectos de este documento el autor la entiende como

«...la valoración, a partir de criterios y referencias preespecificados, de la información técnicamente diseñada y sistemáticamente recogida y organizada, sobre cuantos factores relevantes integran los procesos educativos para facilitar la toma de decisiones de mejora»¹⁴

¹³ Véase al respecto la aportación de Hopkins (Hopkins, D. (1989): *Evaluation for School Development*. Milton Keynes, Open University Press) cuando distingue entre la evaluación **para** la mejora, la evaluación **de** la mejora y la evaluación **como** mejora.

¹⁴ Vid. al respecto Pérez Juste, R. (1995): Evaluación de programas educativos, en A. Medina Rivilla y L.M. Villar Angulo (Edits). *Evaluación de Programas Educativos, Centros y Profesores*. Madrid: Universitat, 73-106.

Cuatro son, por tanto, los componentes fundamentales de la evaluación:

- a) *Los contenidos a evaluar*: «cuantos factores relevantes integran los procesos educativos». Nuestra posición se concreta en una evaluación integral capaz de armonizar la evaluación de la enseñanza y la del aprendizaje, del alumnado y del profesorado, del programa y de la institución o centro educativo¹⁵.
- b) *La información a recoger*: dado su carácter técnico debe ser «técnicamente diseñada y sistemáticamente recogida y organizada». La información es la base de la evaluación, por lo que ésta nunca será mejor que aquélla. En consecuencia, se deberá acudir a cuantas fuentes, técnicas e instrumentos sea preciso para captar de modo suficiente y adecuado cuantos objetos evaluativos se haya decidido evaluar, evitando toda tentación de sesgo o de insuficiencia.

Este punto debe cuidarse con esmero si no se quiere caer en un riesgo especialmente peligroso: el de que la información recogida, por sí misma o por las técnicas e instrumentos utilizados, pueda llegar a desvirtuar los objetivos educativos, que, sin duda, deben ser los elementos rectores tanto del proceso educativo como del propiamente evaluador.

Por último, la información deberá ser organizada de forma que sirva eficazmente a los propósitos de la evaluación, sin llegar a convertirse, como a veces ocurre, en una fase cuasi independiente de la misma, bien por el exceso en su tratamiento bien por su planteamiento al margen de los objetivos a los que sirve.

- c) *La valoración de la información*, donde distinguimos los *criterios*, aplicados a las diversas unidades de información, y las *referencias*, utilizadas para emitir juicios globales de evaluación, en nuestro caso sobre el programa. *Criterios* pueden ser los de *exactitud, precisión, claridad, profundidad, variedad...* En cuanto a las *referencias*, las tres fundamentales son las conocidas como *normativa* (baremo), *criterial* y *personalizada o idiosincrásica*.

Vale la pena insistir de nuevo: los *criterios* deben derivar, de modo coherente y armónico, de los *objetivos* educativos a los que sirve el programa; los riesgos derivados de no cumplirse tal exigencia no son otros que los ya apuntados en el apartado anterior, esto es, la desvirtuación de sus objetivos. Algo similar debe afirmarse de las *referencias*, que deben ser seleccionadas en función de los objetivos y de las finalidades a las que sirve la evaluación¹⁶.

15 No estaría de más poder afirmar lo mismo de la evaluación del *sistema educativo*, pero es preciso reconocer que aquí están implicadas dimensiones no estrictamente pedagógicas o educativas, como son las de tipo político, social o económico, por no reseñar sino las más importantes.

16 No parece difícil mantener la superioridad de una evaluación de referencia *personalizada* cuando de lo que se trata es de apreciar el *progreso* educativo de los educandos en pos de su formación personal; o de la *criterial* cuando, además de lo anterior, entran en juego las expectativas y demandas sociales junto a las exigencias de la disciplina objeto de enseñanza y aprendizaje, algo especialmente importante en niveles educativos no obligatorios; o la *normativa*, cuando sea preciso situar los resultados de un alumno en relación con los de compañeros del propio o de otros centros educativos, o los de un centro —incluso de un sistema educativo— en relación con los de otros, sea para situarse en un contexto en que es preciso competir —becas, selección de personal...— sea para orientar decisiones de diversa índole, como ocurre en el caso de los indicadores y estándares nacionales o internacionales...

- d) *La finalidad*: «facilitar la toma de decisiones de mejora». Caben otras finalidades o funciones¹⁷, pero, desde una perspectiva estrictamente educativa, la más coherente es la *formativa* o para la mejora, sea del alumno y su aprendizaje, sea del profesor y su actividad educativa, sea del centro educativo como unidad.

Debemos hacer notar que el carácter de herramienta al servicio de los objetivos educativos que atribuimos a la evaluación exige de ésta que sea sea continua y formativa, razón por la que, en un primer ciclo o bucle, deberá sobrepasar la clásica etapa final para alcanzar a los procesos e incluso el diseño y la calidad misma del programa; pero, en realidad, nuestra concepción nos conduce a su institucionalización, con lo que, en ciclos o bucles evaluativos posteriores, se da un enlace y entrelazamiento entre la evaluación final de un ciclo y la inicial del siguiente, lo que permite extraer el máximo potencial de mejora a esta función educativa.

Por último, pero tal vez como lo más importante en tanto que representa una cuestión de principio, deseo resaltar tres características, derivadas de todo lo anterior a la par que su fundamento y soporte, que dan a la evaluación un especial alcance: su carácter *integral*, *integrado* e *integrador*.

La dimensión o nota de *integral* nos hace ver la necesidad de que la evaluación, además de abarcar la realidad educativa toda, lo haga de modo armónico, al servicio de una gran meta común: la mejora de la persona, bien por su contribución directa a tal objetivo bien por la incidencia que la mejora de los centros, de los programas o del propio profesorado tengan sobre aquella; la nota de *integrada* es una consecuencia de la anterior: la evaluación es un elemento más, junto al resto y, por ello, en necesaria armonía con todos ellos, al servicio de esa finalidad, no debiendo quedar relegada a una actividad añadida, extrínseca y meramente yuxtapuesta; por último, el carácter *integrador* da a la evaluación una dimensión activa y dinámica, y no meramente pasiva, en esa armonización de los medios al servicio de la finalidad de la educación, por tanto en clara contraposición a planteamientos que puedan subvertir el orden, haciendo del instrumento —la evaluación— la finalidad a servir¹⁸, sea por una reordenación de los medios sea por el tiempo —excesivo— que requiera¹⁹.

Obviamente, caben, y así son la mayoría, evaluaciones parciales —sólo los aprendizajes, sólo los programas, sólo los centros, sólo los materiales de aprendizaje...— pero, en tanto que sean evaluaciones pedagógicas, deberían responder a un sistema integralmente concebido y organizado.

17 Nevo habla de la *sumativa*, la *sociopolítica* y la *administrativa*.

18 Piénsese, por ejemplo, lo que puede ocurrir cuando de los resultados de una evaluación no centrada en los objetivos educativos puedan seguirse consecuencias no deseables para los educandos, para los profesores o para los propios centros educativos (selectividad, promoción del profesorado, evaluación de centros y sus consecuencias de carácter económico —financiación— o social —certificación y acreditación de la calidad—).

19 La evaluación no puede consumir un tiempo que deba dedicarse a la acción educativa.

2.3. Evaluación de programas

En este marco, entendemos la evaluación de programas como

«un proceso sistemático, diseñado intencional y técnicamente, de recogida de información rigurosa —valiosa, válida y fiable— orientado a valorar la calidad y los logros de un programa, como base para la posterior toma de decisiones de mejora, tanto del programa como del personal implicado y, de modo indirecto, del cuerpo social en que se encuentra inmerso»²⁰

Al igual que en el caso anterior, debemos clarificar los conceptos incluidos en la definición, en particular aquellos aspectos que representen alguna peculiaridad respecto de lo ya dicho en el punto 2.2., dedicado a la evaluación en general.

En primer lugar, el *contenido* u objeto de evaluación: sabemos que nos estamos refiriendo a *programas*, pero de todos es conocida la existencia de *modelos* de muy diversa naturaleza en función de su objeto —algunos centrados sólo en los resultados para la constatación del grado de *eficacia*—, de su alcance —evaluaciones por una sola vez— o de sus funciones, fundamentalmente sumativas. Nuestra posición es la de abarcar el programa tanto en su *calidad* —la de sus propios objetivos, de su diseño, de sus medios o del propio sistema de constatación de sus resultados— como en sus *logros* o resultados.

La *finalidad* se centra decididamente en la *mejora*, en perfecta sintonía con la esencia de los actos educativos. Precisamente por esa orientación a la mejora es necesario hacer objeto de evaluación, también, los propios procesos educativos al igual que el proceso de implantación e implementación del programa. Los modelos clásicos centrados en los productos únicamente resultan satisfactorios cuando, tras su aplicación, se constata su eficacia. Pero, en caso de ineficacia, o de niveles insatisfactorios, se revelan incapaces de aportar a los educadores la información necesaria para las decisiones de mejora. En resumidas cuentas, planteamos una evaluación de programas en cuatro grandes etapas —a las que nos referiremos más adelante—, las tres primeras imprescindibles en una primera evaluación, y la cuarta para hacer realidad la integración de la evaluación como una realidad asimilada por el profesorado o la organización.

En cuanto a la *información* ya hemos señalado con anterioridad que es la base de la evaluación, que no puede ser mejor que la información que le sirve de base. Su calidad de *rigurosa* hace referencia a una característica propia de todo procedimiento de acceso científico al saber: no sirve cualquier tipo de información sino aquella que ha sido recogida mediante instrumentos o técnicas debidamente contruidos y adecuadamente contrastados en cuanto a sus características básicas de fiabilidad y validez. En este caso, añadimos la nota de *valiosa* para dar a entender la necesidad de que tal información, además, sea *útil* para facilitar la *toma de decisiones de mejora*. Con frecuencia, la naturaleza de la información recogida es lo suficientemente genérica e inespecífica

20 Vid. al respecto Pérez Juste, R. (1995): obra citada, p. 73.

como para que las decisiones o no sean adecuadas o, en realidad, no puedan tener como base la información recogida, lo cual haría de la información un componente superfluo²¹.

En cuanto a la valoración de la información, daremos por dicho lo señalado al respecto en el punto anterior, señalando, eso sí, la importancia que cobra, dentro del *contexto* en el que el programa se aplica, el *Proyecto educativo* y el *centro educativo*, cuando de programas ordinarios se trate, que, obviamente, deben ser *coherentes* y encontrarse en *sintonía*, en *armonía*, con ambos. La definición y la delimitación de tales contextos, por otro lado, resultan básicas a la hora de establecer tanto los niveles de generalización o, si se prefiere, de aplicación de los resultados a situaciones, ámbitos o contextos diferentes de aquellos en que se aplicó el programa, como de valorar su propia eficacia²². Por idénticas razones, en el caso de programas nuevos, diferentes de los ordinarios, será necesario a tales efectos, definir con claridad el contexto de que se trate, sea el familiar, el amical, el social o el colegial.

3. PLANTEAMIENTOS GENERALES: COMPONENTES BÁSICOS DE UNA PROPUESTA PARA LA EVALUACIÓN DE PROGRAMAS EDUCATIVOS

Dada la entidad del concepto *modelo*, tal vez sea preferible acudir a uno menos fuerte, como puede ser el de *propuesta*. Para elaborar un procedimiento completo de evaluación de programas es preciso tomar en consideración, además de lo anterior, algunos elementos con repercusiones metodológicas de diversa naturaleza y alcance, y que, dado el espacio disponible, nos limitaremos a enunciar:

- **Lo educativo** como referencia constante. Tanto en el diseño como en la implementación, desarrollo y evaluación de programas, se deben tener como referencia: que los objetivos, medios y recursos deben ser educativos —no manipuladores— y adecuados; que el programa coexiste con otros programas en el aula y en el centro, con sus correspondientes posibles interacciones; que el *factor humano*, —tanto sus agentes como sus destinatarios, agentes también— es fundamental: sin su decidida implicación y colaboración no hay programa eficaz; o que el contexto, por lo general orientado por «valores» contrarios o, cuando menos, indiferentes a los que subyacen en el programa implantado y evaluado, juega un papel condicionante de primer orden.

21 Un ejemplo: en evaluaciones sobre el sistema *a distancia* ha sido habitual preguntar a los alumnos por su preferencia por textos ordinarios o específicos para esta modalidad de aprendizaje, o solicitar su valoración de los textos utilizados en el curso anterior. Piénsese, sin embargo, si en el primer caso no hubiera sido conveniente matizar las respuestas en función de la asignatura de que se trate, de la carrera o del ciclo o curso estudiados; y en el segundo, parece imprescindible acudir a una valoración de *cada* texto ya que una valoración en términos de Media o Mediana, apenas sirve para otra cosa que para quedarse tranquilo si es muy elevada o preocupado en caso contrario. A esta cuestión nos referimos al hablar de información *valiosa*.

22 En ocasiones, un buen programa puede no funcionar adecuadamente por falta de adecuación al contexto de referencia.

- El **tamaño o amplitud del programa**: desde los reducidos —un fin de semana— a los que duran uno o varios cursos, con diferente amplitud y ambiciones, lo que puede alterar sensiblemente la metodología a utilizar²³.
- La **complejidad del programa**: con objetivos de muy diversa naturaleza²⁴, con las correspondientes implicaciones para su formulación, actuaciones pedagógicas y evaluación.
- El **estatus del evaluador**: como responsable, simultáneamente, del programa y su evaluación, lo que orienta especialmente hacia enfoques autoevaluativos y continuos, centrados en la mejora, o como sólo evaluador, más propio de evaluaciones externas, de expertos, por una sola vez, de carácter sumativo, que concluyen en un informe entregado —previa discusión o no— a los responsables del programa para la toma de las decisiones que correspondan.
- Las **unidades de análisis**: los programas educativos tienen su destinatario natural en las personas de los educandos. Sin embargo, en ocasiones es necesario, bien sea por motivos intrínsecos bien por interés de los responsables del programa o de su evaluación, tomar en consideración otro tipo de unidades, como el aula, el centro, u conjunto de centros (públicos, privados, de educación infantil, de una comunidad autónoma...).
- La relación entre el **programa y su evaluación**: la evaluación debería ser un componente necesario de cualquier programa entendido en forma integral, al menos en su función formativa, pero, con frecuencia, aparece como algo yuxtapuesto o sobrevenido. El hecho de que la evaluación esté o no integrada en el programa condiciona notablemente la metodología, reduciendo su alcance y limitando sus posibilidades en caso negativo.
- La **información a recoger**: dada la complejidad de los fenómenos educativos, una de las tareas más delicadas consiste en la construcción de lo que se considera dato a los efectos de evaluación del programa.
 Vale la pena *exigir* una plena coherencia entre los objetivos del programa, los de su evaluación, y los datos o la información recogida para decidir sobre él; esto, obviamente, no implica desprecio hacia el otro tipo de datos, pero nunca éstos pueden sustituir a aquéllos, sino complementarlos. Y es que, a pesar de la importancia que concedemos a los datos *duros*, dado que la educación se concreta en resultados académicos y, sobre todo, en determinados comportamientos, no se debe olvidar que las *percepciones* de agentes y destinatarios son importantes en la medida en que pueden condicionar la realidad misma del programa. De ahí el valor de formas de triangulación en el análisis de datos.
 Con todo, no son pocas las ocasiones en que los efectos fundamentales del programa son diferidos, por lo que caben evidencias intermedias, aunque sean reco-

23 Esta doble dimensión de *tamaño y duración*, no sólo es una cuestión de cantidad sino que, incluso, da lugar a ciertas exigencias cualitativas a la hora de evaluarlo, con repercusiones en temas como la necesidad o no de *muestreo*, tanto de contenidos como de informantes, de estar abierto a posibles efectos *no planeados* o de cuidar determinadas interacciones.

24 Piénsese, en el caso de nuestro actual sistema educativo, la diferencia que representa para la evaluación el que los objetivos sean de tipo *conceptual, procedimental o actitudinal*, o de todos ellos a la vez.

gidas una vez finalizada la aplicación del programa, que pueden tomarse como predictoras de efectos posteriores esperables. Valdría la pena, no obstante, dejar clara constancia de ello y no abandonar su constatación posterior, con frecuencia olvidada porque ya no se rentabilizaría política o socialmente el esfuerzo.

3.1. Componentes de la propuesta

Abordaremos este aspecto desde dos enfoques diferentes, pero complementarios: los correspondientes al *programa* y a su *evaluación*.

A) En relación con el *programa*:

Los elementos a tomar en consideración son:

- El *programa* en cuanto tal y, por tanto, con todos sus componentes: sus objetivos y metas con sus correspondientes contenidos; sus medios y recursos y su evaluación.
- El *programa* en su proceso de implantación y desarrollo.
- El *programa* como realidad llevada a cabo, se piense o no en una nueva aplicación.
- El *programa* como una realidad dinámica: la evaluación debe dar lugar a *consecuencias*, que pueden ser teóricas, prácticas, metodológicas y, sobre todo, educativas, tanto sobre los contenidos educativos como sobre el propio programa.

B) En relación con la *evaluación del programa*:

Desde la perspectiva de la evaluación del programa, sus componentes son:

- La *función* o *funciones* a las que sirve, en cuanto condicionantes del resto de decisiones.
- La *metodología* a utilizar, concepto amplio en el que caben diversos aspectos. Destacamos entre ellos los referentes a la calidad técnica del propio diseño de evaluación del programa; al *diseño* o *plan* general de la evaluación: preguntas a las que responder, objetivos a lograr, niveles a alcanzar (significación estadística, significación pedagógica); *métodos* a utilizar, donde, dadas las consideraciones anteriores, optamos por el *principio de complementariedad metodológica*, bien en forma simultánea bien de modo sucesivo; *información* a recoger en torno al contexto, al programa, a su implantación, a su aplicación, a sus resultados y a sus consecuencias; *momentos* en que habrá de ser recogida la información; *técnicas* e *instrumentos* que habrán de utilizarse; sistema de *registro* y los procedimientos de *análisis*; *criterios* para valorar cada una de las dimensiones del programa, y *referencias* que permiten una apreciación global del mismo; tipo de *decisiones* previsibles y los procedimientos para su toma, concreción, seguimiento y evaluación.
- El *informe*: su formato, elaboración, debate, difusión.

3.2. Líneas generales de la propuesta evaluativa

La concreción básica de todo lo anterior puede tomar la forma de la siguiente propuesta, estructurada en torno a los cuatro momentos o etapas arriba citadas, siendo las tres primeras fundamentales e imprescindibles en una primera evaluación de un programa educativo. La cuarta representa una situación deseable, según la cual la evaluación se institucionaliza y se integra en sucesivas aplicaciones del programa.

PRIMER MOMENTO: evaluación del programa en cuanto tal

Se trata de la actividad evaluativa más importante tanto por ser la primera y base de todas las demás como por su contenido —abarca el programa en su globalidad— y sobre todo, por sus grandes aportaciones a la mejora y optimización del programa incluso antes de que sea puesto en marcha (función preventiva de la evaluación).

Finalidad:

Establecer la calidad técnica del programa, su viabilidad práctica y su evaluabilidad. Poner en marcha el programa en condiciones óptimas.

Función:

Formativa: tomar por anticipado las decisiones de mejora que puedan elevar las potencialidades del programa.

En ocasiones puede ser sumativa, sea sobre el programa sea sobre su evaluación.

Metodología:

Análisis de contenido de documentos, estudios prospectivos, técnica Delphi
Juicio de expertos multidisciplinares: metodológicos, científicos, técnicos, pedagógicos.
Registros, pruebas diversas de evaluación inicial (prerrequisitos)

Información a recoger:

Sobre el programa: su fundamentación, su formulación y su relación con las necesidades, carencias, demandas y expectativas de los destinatarios.

Criterios:

*Calidad, pertinencia*²⁵ y *realismo* de objetivos y metas y *adecuación* a los destinatarios y al contexto.

Suficiencia de los apoyos, medios y recursos para conseguir los objetivos, en particular sobre la *formación, implicación y compromiso* de los agentes y demás implicados.

Calidad técnica de los planteamientos de evaluación.

²⁵ La *pertinencia social* es un criterio de calidad recogido en la Conferencia Mundial sobre Educación Superior (París, octubre de 1998) en el documento *La Educación Superior en el Siglo XXI. Visión y Acción* (pág. 25). Allí se puede leer: «La calidad es inseparable de la pertinencia social», desarrollándose esta idea como «una búsqueda de soluciones a las necesidades y problemas de la sociedad y más especialmente a los relacionados con la construcción de una cultura de paz y un desarrollo sostenible».

Calidad técnica del programa: *coherencia* con las bases teóricas y con las necesidades a las que trata de responder; *congruencia* interna entre sus componentes.

Viabilidad del programa.

Evaluabilidad: relevancia, suficiencia, claridad y accesibilidad de la información disponible sobre el programa y de la necesaria en los diversos momentos o etapas de su aplicación y evaluación. Análisis de las dificultades detectadas y previsibles para evaluar el programa

Decisiones:

Generalmente formativas (de mejora previa).

En casos graves, sumativas (retirada del programa, no realización de la evaluación encargada o asumida).

SEGUNDO MOMENTO: evaluación del proceso de implantación del programa.

Finalidad:

Facilitar la toma a tiempo de decisiones de mejora.

Acumular información para introducir mejoras en futuras ediciones del programa.

Función:

Formativa; en casos extremos, sumativa.

Metodología:

Recogida de información: observación, diálogos, entrevistas, análisis de tareas, pruebas formativas intermedias...

Análisis de la información: sesiones de grupo, debates...

Información a recoger:

Sobre el desarrollo del programa, sobre resultados intermedios, sobre efectos no planeados.

Criterios:

Cumplimiento: desfases, desajustes, flexibilidad; *coherencia institucional;* *eficacia parcial;* *satisfacción de los implicados:* agentes, colaboradores, destinatarios

Decisiones:

Ajustes parciales. En casos extremos, suspensión de la aplicación del programa.

TERCER MOMENTO: evaluación de los resultados de la aplicación del programa.

Finalidad:

Comprobar la eficacia del programa

Función:

Fundamentalmente sumativa. Cuando sea posible, sumativa formativizada.

Información a recoger:

Resultados en relación con los objetivos. Efectos —positivos o negativos— no planeados.

Criterios:

Eficacia: grado de logro de los objetivos propuestos. *Eficiencia*: resultados en relación con los medios disponibles y las circunstancias en que el programa se aplica.

Efectividad: efectos beneficiosos no previstos. *Satisfacción* de destinatarios, agentes y otro personal interesado / afectado. *Impacto* del programa en el contexto en que se aplica.

Referencias:

Situación de partida: el programa frente a si mismo (**progreso**).

Niveles preespecificados (evaluación criterial) (**logro del dominio**).

Niveles de otros programas (**superioridad**).

Decisiones:

Sumativas: mantener o suprimir el programa.

Formativizadas: mejorar el programa para una nueva edición.

CUARTO MOMENTO: institucionalización de la evaluación del programa.

Ciclos sucesivos de evaluación / mejora / evaluación / mejora.

En este momento se logra la plena integración del programa y de su evaluación, lo que viene a redundar en la mejora tanto del programa como de la propia metodología de la evaluación, así como en el perfeccionamiento profesional de los educadores.

4. PROBLEMÁTICA

Como se ha podido constatar, la evaluación es una actividad compleja, en la que se cruzan elementos de muy diversa naturaleza, lo que conduce a soluciones muy diferentes y, en ocasiones, muy alejadas unas de otras.

Con todo, la aludida *complejidad* da lugar a una serie de problemas sin resolver, sobre los que habrá que tomar decisiones en cada momento. Reseñamos algunos de los fundamentales:

4.1. La existencia de expertos

Los modelos o propuestas evaluativas externas tienen como base la consideración de la evaluación como una actividad técnica, necesitada de profesionales expertos que la conduzcan a buen puerto.

Con frecuencia, el diseño mismo de la evaluación lo realizan expertos, y personas expertas son las encargadas de recoger, organizar y elaborar la información como base para la configuración de un juicio técnico y hasta para la propuesta de medidas de mejora.

El problema es grave en ciertos tipos de evaluaciones, como lo pone de relieve la situación de la evaluación de la docencia o de la investigación en el Plan de Evaluación de la Calidad de las Universidades, donde no se cuenta ni de lejos con el número de *expertos* necesarios para llevar a cabo tal labor. Pero las consecuencias de una evaluación externa, propia de expertos, llevada a cabo por personas con una formación apresurada, concentrada en una o dos sesiones de fin de semana, son fáciles de imaginar. De por sí, la evaluación levanta recelos en las personas e instituciones evaluadas; si a ello se une que quienes han de llevar a cabo la tarea no son reconocidos por los evaluados como dotados de la competencia y de la experiencia necesarias, lo más probable es que la evaluación pierda una gran parte de su efectividad y hasta pueda llegar a ser la gran ocasión perdida: en efecto, cuando los resultados de una evaluación defraudan y son insatisfactorios, y ello tras un largo proceso destinado a crear las actitudes positivas tanto para aceptar la presencia de personas ajenas como para colaborar activamente en las tareas, resulta muy difícil intentarlo de nuevo.

4.2. La factibilidad

En el campo de la evaluación se corren muchos riesgos, y uno de ellos es el intento de perfeccionismo. Cuando un grupo de expertos es invitado a proponer modelos evaluativos, lo hace más desde la perspectiva de sus colegas que de la correspondiente a las personas e instituciones a las que se aplicará la evaluación.

Este hecho suele conducir a propuestas muy elaboradas, amplias, profundas y complejas, cuya aplicabilidad es muy costosa en tiempo, recursos materiales y dedicación del personal.

Nos encontramos ante un importante reto: el de conjugar una evaluación que merezca la pena por la aportación de información que pueda representar, y el tiempo, los recursos y el personal necesarios para lograrlo. La investigación debería orientarse a identificar los contenidos mínimos necesarios para una evaluación que permita ser valorada como *suficiente*, dejando fuera todos aquellos otros aspectos que podríamos denominar *superfluos*. Se trata, en definitiva, de elaborar propuestas centradas en lo esencial, suficientes y factibles.

Recuérdese que las normas del Joint Committee se estructuran en torno a cuatro grandes criterios, siendo el de *factibilidad o viabilidad* uno de ellos, al lado de los de *utilidad, probidad o eticidad y precisión o exactitud*. Los estándares de *factibilidad* se orientan a lograr una evaluación realista, prudente, diplomática y austera, y se concretan en las tres manifestaciones siguientes: *procedimientos prácticos, viabilidad política y coste-efectividad*.

La exigencia de factibilidad resulta especialmente necesaria si se quiere crear una cultura autoevaluativa, la que viene considerando el autor como la que representa mayores posibilidades de mejora de la acción educativa en aulas, centros y demás ámbitos educativos.

4.3. Juicios expertos o juicios democráticos?

En la propuesta que hemos ido configurando nos hemos decantado, en lo fundamental, por: a) una evaluación orientada a la mejora; b) desde enfoques internos; inicialmente global, pero, en ciclos sucesivos, parcial, centrando la atención en los puntos *fuertes* y *débiles* previamente detectados; c) realizada por la propia comunidad educativa, con el asesoramiento de expertos cuando no los haya en la propia comunidad; d) aplicando el principio de complementariedad metodológica; e) con la participación de los miembros de la comunidad.

Pues bien, tanto la *participación* como el enfoque de *complementariedad metodológica* conducen a un problema sin resolver: ¿quién o quiénes deben emitir los *juicios* valorativos «definitivos»? y, por tanto, ¿quién ha de tomar las decisiones de mejora?

Lo ideal sería que tanto unos como otros fueran el resultado de *sesiones de evaluación* en las que, con una adecuada dinámica de grupo, los responsables de la institución alcanzaran el consenso en los aspectos fundamentales o, al menos, la aceptación y asunción como propios de los acuerdos mayoritarios. Pero si no fuera así, resolver las cuestiones en forma de votación no parece ser la respuesta más correcta.

Si se acepta que la *participación* es una cuestión de *grado*²⁶ y que, en consecuencia, admite diversas manifestaciones, desde la simple recepción y aportación de información hasta la co-gestión de las instituciones, pasando por la delegación de competencias..., parece adecuado pensar que, en casos de falta de acuerdo final, sea en el diagnóstico de los puntos fuertes y débiles, sea en la selección de las medidas de mejora a aplicar, sea en la elaboración de los correspondientes planes tanto de mejora de la institución o del programa como de su misma evaluación, las posiciones últimas corresponden a los responsables de su gobierno. No obstante, nunca se insistirá lo suficiente en la importancia de conseguir, mediante el diálogo y las sesiones de grupo debidamente conducidas, un acuerdo básico tanto en el diagnóstico de los puntos débiles como en las medidas a tomar: el apoyo a los planes de mejora depende en gran medida de este importante hecho.

4.4. Los planteamientos éticos

Uno de los problemas más serios de las evaluaciones institucionales, sobre todo cuando es realizada por parte de expertos comisionados o contratados por los responsables del sistema educativo, de los centros educativos o de los programas educativos, es el relativo a la independencia profesional de los evaluadores. Junto a ello, es preciso referirnos al propio comportamiento ético de tales profesionales.

Recordemos al efecto las palabras de Stufflebeam y Shinkfield sobre las pseudoevaluaciones, llegando a denominar charlatanes y manipuladores a quienes no se comportan en este punto de modo adecuado:

26 El lector puede consultar Ciscar, C. y Uría, M.E. (1986). *Organización escolar y acción directiva*. Madrid: Popular; o Muñoz, A. y Román, M. (1989). *Modelos de organización escolar*. Madrid: Cincel.

«Puede que recopilen la información rigurosamente, pero no revelan las verdaderas conclusiones, o lo hacen de modo selectivo, o incluso las falsifican, engañando tanto a la gente como a sí mismos»²⁷.

El problema resultará de difícil solución mientras los evaluadores se encuentren en situación de desigualdad con respecto a quienes encargan las evaluaciones; la falta de asociaciones profesionales fuertes en el campo de la evaluación deja a los evaluadores en la situación de asalariados en relación con sus patronos. Con frecuencia, el evaluador debe optar entre asumir las condiciones de quienes le contratan o arriesgarse a que su trabajo no tenga la difusión necesaria como para influir en la mejora de las personas y de las organizaciones sobre las que ha trabajado profesionalmente.

Este importante problema ha sido recogido en las Normas del Joint Committee, dedicándole uno de sus cuatro grandes apartados, el de *honradez o probidad*, cuya meta se encuentra en asegurar que la evaluación sea realizada legal y éticamente, y con el debido respeto a todos cuantos intervienen y están interesados en sus conclusiones. El apartado de *honradez* está integrado por los ocho estándares de calidad siguientes: *orientación al servicio, convenios o acuerdos formales, respeto a los derechos humanos, respeto a la dignidad e integridad humanas, informes completos y claros, declaración de los resultados, conflicto de intereses y responsabilidad fiscal*.

4.5. La comparabilidad de los datos

Cada vez con mayor frecuencia se plantea el problema de la comparabilidad de los resultados, sobre todo cuando se trata de programas en gran escala o de evaluaciones institucionales amplias que afectan al sistema educativo, a las universidades públicas (o a las privadas), a un conjunto de colegios de una misma institución o titularidad... La comparabilidad puede ser una manifestación más de un fenómeno tan reciente como arrollador: el de la globalización.

En ocasiones, este hecho viene impulsado por la función evaluativa de rendición de cuentas o por el cumplimiento del principio de derecho a la información para facilitar las decisiones de los ciudadanos.

El problema de la comparabilidad es de elevada complejidad desde una perspectiva puramente técnica²⁸, pero se complica todavía más por algunas de las exigencias de nuestro tiempo en lo relativo al *currículum* de las instituciones o por los derechos consagrados tanto a la libertad de cátedra de los profesores como a la autonomía institucional de las universidades u otro tipo de instituciones.

²⁷ *Ibíd.*, págs. 67 s.

²⁸ De hecho, hasta en el marco de los *indicadores* resulta difícil de lograr. Así, en *Análisis del Panorama educativo. Los indicadores de la OCDE. 1995*, editado en Madrid por Mundi-Prensa, se lee lo siguiente: «*Comparabilidad: mejoras y problemas persistentes*: Como resultado de las mejoras en el instrumento e instrucciones de recogida de datos y el continuo diálogo entre la Secretaría de la OCDE y los suministradores nacionales de datos, estos son, de algún modo, comparables. (...) Sin embargo, sigue habiendo importantes problemas de comparabilidad...» (Pág. 11).

De un tiempo a esta parte, en efecto, se está tomando conciencia de la conveniencia de que los *curricula* reduzcan su grado de centralización en los diferentes países, dejando autonomía a las diversas comunidades y regiones en que se estructura cada nación y hasta a los propios centros para acomodar y completar las líneas básicas del currículo nacional. En España, junto al denominado *Diseño Curricular Base*, de aplicación en todo el país, las diferentes Comunidades Autónomas, en un primer nivel, y los centros educativos, en un segundo, cierran el currículo a cursar por los estudiantes, algo que podrá sufrir adaptación en el nivel del aula. ¿Cómo compaginar este hecho con la necesidad de contar con datos que sitúen los resultados de un país en relación con otros de su entorno? ¿Cómo hacer posible la comparación entre los resultados de diferentes colegios de una misma sociedad o empresa?²⁹

El caso de las Universidades puede ser todavía más agudo. Por principio constitucional, las universidades gozan de un elevado grado de autonomía en su organización, gestión y desarrollo de las actividades académicas. ¿Cómo comparar resultados de universidades con tamaño y complejidad muy diferentes, planes de estudio con notable diversidad, número de alumnos, profesores y personal distintos, *ratios* muy alejadas unas de otras...?

Un problema añadido lo es el relativo a las dificultades que este hecho representa para la movilidad estudiantil, tanto para las convalidaciones dentro del propio sistema cuanto para el reconocimiento de Títulos entre diferentes países. Dentro del propio país puede darse el caso de que haya dificultades para poder aplicar una misma prueba para el acceso a la Universidad³⁰; las dificultades se acrecientan para homologar los títulos universitarios de diferentes países, incluso dentro de la propia Comunidad Europea.

Una solución, al menos parcial, siempre que la comparación se oriente a la *mejora*, puede encontrarse en seleccionar como referencia no los estándares nacionales o internacionales sino los datos de la propia institución en una perspectiva evaluativa diacrónica. Estaríamos hablando en tal caso de *superación, avance, mejora, progreso* en lugar de *superioridad y competitividad*. Esta perspectiva podría hacerse compatible con una evaluación criterial: la referencia serían los niveles de eficacia, eficiencia, satisfacción... que la propia institución, tras el análisis de su situación, considerara razonablemente alcanzables. Las comparaciones entre los indicadores de los países quedarían limitadas a aquellos aspectos estructurales, funcionales y organizativos en que los datos brutos pueden ponerse en relación con datos contextuales. Por ejemplo: % de alumnos escolarizados en los diferentes niveles; *ratios* de éxito (superación de

29 Sobre esta cuestión reflexionó el profesor De la Orden en la conferencia de clausura del XII Congreso Nacional de Pedagogía y Iberoamericano, celebrado en Madrid en septiembre del año 2000.

30 En España, la prueba de selectividad aplicada tras el Bachillerato para acceder a la Universidad no puede tener como base unas mismas materias debido a que se dan diferencias entre las que se estudian en los diversos cursos. En efecto, las materias que en unas Comunidades Autónomas se estudian en segundo curso de Bachillerato (sobre el que se estructura la citada prueba) en otras se cursan en primero, dándose el caso de que, en otras, se reparten entre primero y segundo.

los diferentes cursos) por etapas educativas; % del producto interior bruto invertido en educación; % de alumnos del Centro educativo que superan las pruebas de selectividad y cuota de estos que culminan, después, los estudios universitarios en cada una de las carreras...

Sin embargo, no estaría de más un elevado esfuerzo de los responsables políticos por encontrar el equilibrio entre la adecuada respuesta a la diversidad social, institucional y personal y el establecimiento de un currículo básico centrado en la cultura propia de los grandes ámbitos territoriales: el Estado, la Comunidad Europea, Occidente.... Ambos polos deben complementarse y nunca darse la espalda.

BIBLIOGRAFÍA

- Abraira, C. (1994). *Efectos de la evaluación formativa en alumnos de Matemáticas de Escuela Universitaria de formación de Profesorado de Educación General Básica*. Tesis doctoral. Madrid: UNED.
- Aguado Odina, M^a T. (1993). *Análisis y evaluación de modelos y programas de educación infantil en el marco de la Reforma de la Enseñanza*. Madrid: UNED.
- Alvira, F. (1991). *Metodología de la evaluación de programas*. Madrid: CIS.
- Aparicio, F. y González Tirados, R.M^a (1994). *La calidad de la enseñanza superior y otros temas universitarios*. Madrid: ICE de la Universidad Politécnica.
- Bartolomé, M. (1994). La investigación cooperativa. En V. García Hoz (Dir.). *Problemas y métodos de Investigación en Educación Personalizada* (pp. 376-403). Madrid: Rialp.
- Bartolomé Pina, M. (1997). Panorámica general sobre Educación Intercultural en Europa. *Revista de Investigación Educativa*. Vol. 15, 1, 7-39.
- Brenes Chacón, O.E. (1998). *Influencia de las metodologías activas en la enseñanza de las Ciencias sobre los resultados académicos y socioafectivos*. Tesis doctoral. Madrid: UNED.
- Buendía Eisman, L. (1996). La investigación sobre evaluación educativa. *Revista de Investigación Educativa*. Vol. 14, 2, 5-24.
- Bunge, M. (1981). *La ciencia: su método y su filosofía*. Buenos Aires: Ediciones Siglo Veinte.
- Burgen, A. (Edit) (1999). *Metas y proyectos de Educación superior*. Madrid: Fundación Universidad-Empresa.
- Cabrera, F. y Espín, J.V. (1986). *Medición y evaluación educativa*. Barcelona: PPU.
- Castillo Arredondo, S. (1987). Evaluación de la eficacia del contenido del consejo orientador. Tesis doctoral. Madrid: UNED.
- CECE-Club Gestión de Calidad (1997). *Guía de autoevaluación. Centros Educativos o formativos no universitarios. Modelo Europeo de calidad total en la gestión*. Madrid: ITE-CECE.
- CIDE. (1995). *Calidad de la Educación y Eficacia de la Escuela*. Madrid: Autor.
- Colás, P. y Rebollo, M^a A. (1993). *Evaluación de Programas. Guía práctica*. Sevilla: Kronos.
- Consejo de Universidades (1991). *La evaluación de las instituciones universitarias*. Madrid: Secretaría General del Consejo de Universidades.
- Consejo de Universidades (s.f.). *Plan Nacional de Evaluación de la calidad de las universidades*. Madrid: Secretaría General del Consejo de Universidades.

- Cook, T.D. (1985). Postpositivist Critical Multiplism. En L. Shotland y M.M. Marck (Edits). *Social Science and Social Policy*. Beverly Hills: Sage Pub.
- Cristóbal Martín, M. (1991). *Análisis y evaluación de las prácticas en las Escuelas Universitarias de Formación del Profesorado de EGB de la comunidad de Madrid*. Tesis doctoral. Madrid: UNED.
- Dueñas Buey, M^a L. (1988). Análisis y valoración de un ensayo de integración escolar en Madrid-Capital. Tesis doctoral. Madrid: UNED.
- Dueñas Buey, M^a L. (1990). Evaluación de programas de apoyo a la integración escolar de deficientes mentales ligeros. *Revista de Investigación Educativa*. Vol. 8, 15, 403-408.
- Dueñas Buey, M^a L. (1993). Evaluación de un programa de apoyo a la integración escolar de deficientes mentales ligeros. En P. Colás y M.A. Rebollo. *Evaluación de Programas*. Sevilla: Kronos, 209-214.
- Etxeberria, J., García Jiménez, E., Gil Flores, J. y Rodríguez Gómez, G. (1995). *Análisis de datos y textos*. Madrid: Ra-ma.
- Fernández Ballesteros, R. (Coord.) (1996). *Evaluación de Programas*. Madrid: Síntesis.
- Gamazo Pérez, T. (1999). *Efectos de la aplicación de algunos instrumentos del Programa de Enriquecimiento Instrumental de R. Feuerstein en los educandos de primero y segundo ciclo de Educación Primaria*. Tesis doctoral. Madrid: UNED.
- García Hoz, V. (1991). El principio de complementariedad en la investigación pedagógica y en la Educación personalizada. *Anales de la Real Academia de Ciencias Morales y Políticas*. 68, 299-317.
- García Hoz, V. (Dir.). (varios años). *Tratado de Educación Personalizada*. 33 volúmenes.
- García Jiménez, E. (1994). Investigación etnográfica. En V. García Hoz (Dir.). *Problemas y métodos de Investigación en Educación Personalizada* (pp. 343-375). Madrid: Rialp.
- Gaviria, J.L. (1999). El enfoque multinivel en la evaluación de sistemas educativos. *Revista de Investigación Educativa*. 17, 2, 429-443.
- Gento Palacios, S. (1994). *La gestión de calidad en Educación*. Madrid: La Muralla.
- Gento Palacios, S. (1998). *Implantación de la calidad total en las instituciones educativas*. Madrid: UNED.
- Ginés Mora, J., I.V.I.E. y Universidad de Valencia (1998). *Indicadores de gestión y rendimiento para las universidades españolas*. Valladolid.
- Hernández Pina, F. (1993). Concepciones en el estudio del aprendizaje de los estudiantes universitarios. *Revista de Investigación Educativa*. 22, 117-154-24.
- House, E.R. (Edit). *New Directions in Educational Evaluation*. Lewes: Falmer Press.
- INCE (1998). *Diagnóstico del Sistema Educativo 1997*. Madrid: Ministerio de Educación y Cultura.
- INCE (2000). *Sistema Estatal de Indicadores de la Educación*. Madrid: MECD.
- INCE (2000). *Evaluación de la enseñanza y el aprendizaje de la lengua inglesa. Educación primaria 1999. Avance de resultados*. Madrid: Ministerio de Educación y Cultura.
- INCE (2000). *Evaluación de la educación primaria. Datos básicos 1999*. Madrid: Ministerio de Educación y Cultura.
- Jiménez Fernández, C. (2000). Cómo conocer a los más capaces: diagnóstico de la alta capacidad. En C. Jiménez Fernández: *Diagnóstico y educación de los más capaces*. Madrid: UNED.

- Joaristi Olariaga, L. y Lizasoain Hernández, L. (1998). *BMDP. New System para Windows*. Madrid: Paraninfo.
- Joint Committee on Standards for Educational Evaluation (1981). *Standards for Evaluations of Educational Programs, Projects and Materials*. New York: McGraw-Hill.
- Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standards*. Thousand Oaks, CA.: Sage.
- Jornet Meliá, J. y Suárez Rodríguez, J. (1994). Evaluación referida a criterio. Construcción de un test criterial de clase. En V. García Hoz (Dir.). *Problemas y métodos de Investigación en Educación Personalizada* (pp. 419-443). Madrid: Rialp.
- Hopkins, D. (1989). *Evaluation for School Development*. Milton Keynes: Open University Press.
- Kemmis, S. (1989). Seven Principles for Programs Evaluation in Curriculum Development and Innovation. En E. R. House (Ed.). *New Directions in Educational Evaluation* (pp. 117-140). Lewes: Falmer Press.
- Luján, J. y Puente, J. (1996). *Evaluación de centros docentes. El Plan EVA*. Madrid: MEC.
- Marchesi, A. y Martín, E. (1998). *Calidad de la enseñanza en tiempos de cambio*. Madrid: Alianza Editorial.
- Martín, M^a T. y Sarrate, M^a L. (1999). *Evaluación y ámbitos emergentes en animación socio-cultural*. Madrid: Sanz y Torres.
- Martínez Mediano, C. (1991). *Factores de eficacia de los Centros Asociados de la UNED en el curso de acceso para mayores de veinticinco años*. Tesis doctoral. Madrid: UNED.
- Martínez Mediano, C. (1996). *Evaluación de programas educativos*. Madrid: UNED.
- Mateo Andrés, J. (2000). La evaluación del profesorado y la gestión de la calidad de la Educación. Hacia un modelo comprensivo de evaluación sistemática de la docencia. *Revista de Investigación Educativa*. Vol. 18. 1, 7-36.
- Mateo, J., Escudero, T., Miguel, M. de, Ginés Mora, J. y Rodríguez Espinar, S. (1996). La evaluación del profesorado. Un tema a debate. *Revista de Investigación Educativa*. Vol. 14, 2, 73-94.
- Michavilla, F. (Edit). (1998). *Experiencias y consecuencias de la evaluación universitaria*. Madrid: Fundación Universidad - Empresa.
- Miguel Díaz, M. de (Coord) (1999). La evaluación de programas: entre el conocimiento y el compromiso. *Revista de Investigación Educativa*, Vol. 17, 2, 345-380.
- Miguel Díaz, M. de (1997). La evaluación de los centros educativos. Aproximación a un enfoque sistémico. *Revista de Investigación Educativa*. Vol. 15, 2, 145-178.
- Ministerio de Educación y Cultura (1997). *Modelo Europeo de Gestión de Calidad*. Madrid: Centro de Publicaciones. Secretaría General Técnica.
- Medina Heredia, M^a T. (1998). *Evaluación de programas educativos. Los materiales didácticos del Centro Nacional de Educación a Distancia*. Tesis doctoral. Madrid: UNED.
- Mozas Flores, J.C. (1997). *Diseño, aplicación y evaluación de un programa integrado en el currículo de Educación Primaria, sobre el fomento de actitudes contrarias al tabaco y al alcohol en un marco interactivo con los principios de una vida sana*. Tesis doctoral. Madrid: UNED.
- Nevo, D. (1986). The Conceptualization of Educational Evaluation. En E.R. House (Edit): *New Directions in Educational Evaluation*. (pp. 15-30). Lewes: Falmer Press.

- Nevo, D. (1990). Role of the Evaluator. En H. J. Walberg y G. D. Haertel (Edits.). *The International Encyclopedia of Educational Evaluation*, (pp. 89-90). Oxford: Pergamon Press.
- OCDE. (1991). *Escuelas y calidad de la enseñanza. Informe internacional*. Madrid: Paidós-M.E.C.
- OCDE (1995). *Análisis del panorama educativo*. Madrid: OCDE-Mundi-Prensa.
- Orden Hoz, A. de la (1993). Influencia de la evaluación del aprendizaje en la eficacia de la enseñanza. *Revista de Investigación Educativa*. 22, 7-42.
- Orden Hoz, A. de la (1999). La investigación educativa: naturaleza, funciones y ambigüedad de sus relaciones con la práctica y la política educativas. *Revista de Investigación Educativa*. Vol. 17, 1, 7-32
- Orden Hoz, A. de la (2000). La globalización y los estándares nacionales de rendimiento educativo. Conferencia de clausura del XII Congreso Nacional de Pedagogía y I Iberoamericano. Documento no publicado.
- Patton, M.Q. (1980). *Qualitative evaluation methods*. Beverly Hills: Sage.
- Pérez Juste, R. y García Ramos, J.M. (1989). *Diagnóstico, evaluación y toma de decisiones*. Madrid: Rialp.
- Pérez Juste, R. (Dir) (1989). La Universidad Nacional de Educación a Distancia: luces y sombras de una institución innovadora. *Bordón*. 41, 3. Monográfico.
- Pérez Juste, R. (1989). Evaluación del material didáctico. *Revista Iberoamericana de Educación Superior a Distancia*. Vol. I, 3, 93-112.
- Pérez Juste, R. (Dir) (1991). *La Universidad Nacional de Educación a Distancia: aproximación a la evaluación de un modelo innovador*. Madrid: CIDE.
- Pérez Juste, R. (Dir) (1992). *Estudios sobre el curso de acceso a la UNED para mayores de 25 años*. Madrid: UNED.
- Pérez Juste, R. (1993). Evaluación de las adaptaciones curriculares. *Revista de Ciencias de la Educación*. 153, 51-68.
- Pérez Juste, R. (1993). Evaluación de Programas de Orientación. *Actas del V Seminario Iberoamericano de Orientación*, págs. 46-65.
- Pérez Juste, R., Martínez Mediano, C. y Rodríguez Gómez, G. (1994). *Centros educativos de calidad*. Madrid: ITE de CECE.
- Pérez Juste, R. (1994). Investigación y mejora de programas. Programas comunes y programas personales. En V. García Hoz (Dir.). *Problemas y métodos de Investigación en Educación Personalizada* (pp. 510-536). Madrid: Rialp.
- Pérez Juste, R., García Llamas, J.L. y Martínez Mediano, C. (Coords.) (1995). *Evaluación de Programas y Centros Educativos*. Madrid: UNED.
- Pérez Juste, R., Cristóbal Martín, M^a. y Martínez Aznar, B. (Coords.) (1995). *Evaluación de Profesores y Reformas Educativos*. Madrid: UNED.
- Pérez Juste, R. (1997). La evaluación del material impreso como elemento sustancial en la evaluación de programas. Ponencia presentada a las *Jornadas Internacionales sobre el Material Impreso en la Enseñanza a Distancia*. Madrid: IUED de la UNED.
- Pérez Juste, R. (1998). La calidad de la Educación universitaria. *Revista Iberoamericana de Educación a Distancia* 1, 1, 13-38.
- Pérez Juste, R. (1999). Nuevas perspectivas de gestión, control y evaluación. *Bordón*, 51, 4, 449-478.

- Pérez Juste, R. (2000). La calidad de la Educación. En R. Pérez Juste y otros: *Hacia una Educación de Calidad*, (pp. 13-44). Madrid: Narcea.
- Pérez Juste, R. (2000). La evaluación como medio para la mejora de la calidad y de la eficacia del aprendizaje, de la educación y de las instituciones. Tercera ponencia de la Sección VI del XII Congreso Nacional de Pedagogía. Madrid.
- Repetto Talavera, E. (1994). Últimas aportaciones en la evaluación de programas de orientación metacognitiva de la comprensión lectora. *Revista de Investigación Educativa*. 23, 314 -323.
- Rodríguez Espinar, S. (1997). La evaluación institucional universitaria. *Revista de Investigación Educativa*. Vol. 15, 2, 179-216.
- Rodríguez Gómez, G. (1995). El centro escolar como objeto de investigación: la evaluación interna de centros. En V. García Hoz (Dir.). *Del fin a los objetivos de la Educación Personalizada* (pp. 308-332). Madrid: Rialp.
- Rodríguez Gómez, G., Gil Flores, J. y García Jiménez, E. (1996). *Metodología de la investigación cualitativa*. Málaga: Aljibe.
- Rossi, P.H. y Freeman, H.E. (1989). *Evaluación: Un enfoque sistemático para programas sociales*. México: Trillas.
- Sanz Moreno, A. (1992). *Diseño y evaluación de un programa para desarrollar las habilidades implicadas en la comprensión y síntesis de textos expositivos*. Tesis doctoral. Madrid: UNED.
- Sanz Oro, R. (1990). *Evaluación de programas en orientación educativa*. Madrid: Pirámide.
- Scriven, M. (1967). «The Methodology of Evaluation», en *Perspectives on Curriculum Evaluation*, AERA Monograph Series on Curriculum Evaluation, Chicago, Rand McNally.
- Sidman, M. (1978). *Táctica de la investigación científica*. Barcelona: Editorial Fontanella.
- Stenhouse, L. (1984). *Investigación y desarrollo del currículum*. Madrid: Morata.
- Stufflebeam, D.L. y Shinkfiel, A.J. (1987). *Evaluación sistemática. Guía teórica y práctica*. Barcelona: Paidós-MEC.
- Tejedor, F.J. (1990). Perspectiva metodológica del diagnóstico y evaluación de necesidades en el ámbito educativo. *Revista de Investigación Educativa*. Vol. 8, 16, 15-37.
- UNESCO (1998). *La Educación Superior en el Siglo XXI. Visión y Acción*. París: UNESCO.
- Varios (1991). Evaluación de Programas. *Bordón*, 43, 4, monográfico.
- Varios (1998). La evaluación de la Universidad Española. *Revista de Educación*, 315. Monográfico.
- Varios (1996). Evaluación de la calidad de la Educación. *Revista Iberoamericana de Educación*, 10, Monográfico.
- Villar Angulo, L.M. y Medina Rivilla, A. (Edits.) (1998). *Evaluación de Programas Educativos, Centros y Profesores*. Madrid: Universitat.
- Wilson, Jh. D. (1992). *Cómo valorar la calidad de la enseñanza*. Madrid: Paidós-M.E.C.
- Woods, P. (1987). *La escuela por dentro*. Madrid: Paidós-M.E.C.