



La evaluación de programas en el marco de la educación de calidad¹

Ramón Pérez Juste

UNED

El artículo presenta una concepción de la evaluación caracterizada como integral, integrada e integradora, orientada por dos principios básicos: un enfoque radicalmente pedagógico, educativo, y un planteamiento unitario, armónico y coherente, tanto entre la evaluación de los diferentes objetos como entre ésta y el conjunto de elementos que integran los procesos educativos de calidad.

This article shows a conception of the evaluation characterised as integral, integrated and integrator, guided for two basic principles: a radical, pedagogical and educational approach and a harmonic, coherent and unitary approach, of the evaluation of the different objects and the elements that take part in the quality educational processes.

1. Introducción

Es generalmente conocida la afirmación de Buyse en el sentido de que la mejor enseñanza consiste en un análisis entre dos síntesis. La primera, un tanto oscura, destinada a ofrecer una panorámica general del tema, mientras la segunda, a partir del análisis, se orienta a recomponer la realidad analizada integrando las partes de forma armónica. Entiéndase este primer punto como un intento de síntesis inicial del tema.

La exposición se va a estructurar en tres grandes apartados. En el primero haremos un análisis del concepto de evaluación; en el segundo centraremos nuestra atención en la función docente, marco de referencia para abordar lo que debe constituir la parte central de nuestra aportación, concretada, ya en la tercera parte, en una reflexión y propuesta en relación con la evaluación de programas.

Los planteamientos que siguen tienen dos hilos conductores de extraordinario relieve. Por un lado, el reconocimiento de la evaluación como una herramienta, potente eso sí, pero sólo herramienta, puesta al servicio de la realidad evaluada; por otro, la necesidad de articular una concepción evaluativa armónica, tanto *ad intra*, esto es, capaz de dar sentido unitario a la actividad evaluativa aplicada a los

diferentes objetos, como *ad extra*, o lo que es lo mismo, situando la evaluación en un marco de coherencia con el conjunto de los diferentes elementos integrantes de las propuestas educativas de calidad. El primero de tales «hilos» sitúa a la evaluación en su verdadero papel, el de instrumento, el de medio, cuya principal meta ha de consistir no ya en la constatación del grado de eficacia de las actuaciones educativas, aunque no sea algo desdeñable, sino en su contribución al logro de los objetivos planteados. El segundo no es sino una consecuencia del primero, pero tiene importantes repercusiones² para la práctica educativa en general y para la evaluativa en particular.

1.1. Evolución del concepto de evaluación

Sin embargo, las cosas no siempre han sido así, ni lo son todavía, de hecho, en muchas ocasiones y ámbitos. Y es que el de la evaluación es uno de los conceptos que ha sufrido mayores cambios, fruto, sobre todo, del gran enriquecimiento conceptual que ha venido experimentando.

Inicialmente, la evaluación fue una actividad desempeñada por quienes se encontraban en posición de poder, autoridad o superioridad sobre las personas evaluadas, bien en cuanto tales bien en sus actividades o realizaciones. De esta forma, la evaluación ha servido para la selección de personas, para la calificación de los aprendizajes, para la promoción dentro del sistema o para la certificación de titulaciones socialmente reconocidas.

Esta posición se ha concretado en una clara función de control, en su discontinuidad en relación con la continuidad de la realidad evaluada, en su formalización al final de determinados periodos de tiempo tras la realización de las actividades evaluadas, y en la asignación de juicios de valor a partir de criterios diversos, no siempre transparentes ni públicos y, en cualquier caso y por lo general, desconocidos en el momento en que se desarrollaban las actividades y los comportamientos que serían objeto de evaluación.

1.2. El salto cualitativo en la evolución del concepto de la evaluación

A juicio de quien escribe, el salto cualitativo se produce cuando a la evaluación se le asigna, o se le reconoce, la función de mejora, una función, por cierto, plenamente coherente con la esencia de los actos educativos (la educación es una actividad intencional y sistemática al servicio de la mejora o perfeccionamiento de las personas).

La función de mejora, propia de la evaluación conocida desde Scriven como formativa, implica grandes cambios, algunos, como el primero de los que siguen, de especial relevancia en nuestro contexto de reflexión:

- La ampliación de los «objetos» evaluados, pasando a incorporar no sólo la enseñanza al lado de los aprendizajes, sino todo el conjunto de elementos que contribuyen al aprendizaje: los medios, los contextos en que se lleva a cabo y hasta los educadores, el medio fundamental junto a las actividades programadas al servicio de los objetivos (García Hoz, Yinger).
- La ampliación de los «sujetos» o evaluadores; no son sólo las personas investidas de poder o autoridad, sino todas aquellas que puedan aportar información relevante y adecuada para promover las mejoras. Por tanto, no debe sorprender que incluso los alumnos puedan ser evaluadores, y no sólo de sus propias actividades y logros sino de la actividad de los educadores³.
- La ampliación de los momentos: la evaluación no se limita a los momentos en que algo ha concluido sino que se extiende a los procesos, en la medida en



que tales procesos pueden ser objeto de mejora y en que, de su mejora, se derivarán necesariamente mejoras en los productos o resultados. Es más, en ese marco de referencia cabe hablar de una evaluación previa, inicial, de carácter preventivo, que aumente la probabilidad de que los procesos educativos puedan llegar a ser eficaces.

- La continuidad frente a la discontinuidad: dado que los procesos son continuos, no parece adecuado relegar la actuación evaluadora al momento final de tales procesos, cuyo troceamiento con frecuencia es artificial (un trimestre, un cuatrimestre, un curso académico); sino que puede, y debe, realizarse tan en paralelo como sea posible a fin de extraer los máximos beneficios de la actividad.

1.3. Repercusiones para la evaluación educativa

Asumir estos planteamientos nos conduce a una reconsideración conceptual de la evaluación, que puede quedar sintetizada y definida por tres adjetivos: integral, integrada e integradora. Si el primero repercute sobre el objeto de la evaluación, el segundo tiene consecuencias técnicas y metodológicas de gran relieve, mientras el tercero es clave para uno de nuestros hilos conductores de reflexión: la armonía tanto ad intra como ad extra.

1.1.1. Evaluación integral

El carácter de herramienta o instrumento que atribuimos a la evaluación hace de ella un poderoso medio al servicio de los objetivos educativos. Por tanto, en la medida en que una educación de calidad se plantea la formación integral de las personas, la evaluación no puede resultar disfuncional, esto es, no puede sugerir a quienes en determinados momentos son objeto de evaluación que ciertos objetivos son los que, de verdad, importan mientras el resto sólo tienen una presencia testimonial. Una actuación tal altera el currículo al orientar el esfuerzo y el trabajo hacia aquellos objetivos que son objeto de evaluación, dejando al resto en un segundo plano. Es preciso potenciar, pues, la armonía y la coherencia entre los proyectos -educativos y curriculares- y los programas de los profesores con la evaluación practicada en centros y aulas.

Si la educación integral representa la necesidad de cultivar, promover, estimular y favorecer el logro de objetivos no sólo cognoscitivos y hasta cognitivos, sino afectivos, sociales, morales o estéticos⁴, la evaluación no puede centrarse en los dos primeros ni, mucho menos, como la investigación revela, en los meramente cognoscitivos. Es más, la posición reduccionista lleva en muchas ocasiones a limitar la atención, dentro de estos, a los de las categorías inferiores de las taxonomías, como pueden ser los aprendizajes simplemente memorísticos o prácticos, sin apenas comprensión ni asimilación por parte de los educandos.

Por último, y en la medida en que la educación se pone al servicio de la mejora de la persona, nada mejor que acudir a la «autoevaluación» como medio de promover la capacidad de metacognición, de toma de conciencia de las propias posibilidades y limitaciones, de aceptación de sí mismo y de compromiso autónomo y libre con determinadas metas libre y autónomamente seleccionadas y asumidas.

1.1.2. Evaluación integrada

El carácter formativo concedido a la evaluación hace de ella una actividad en perfecta sintonía y armonía con la actividad educativa; no cabe, por tanto, que, en cuanto tal, la evaluación sea una actuación añadida, yuxtapuesta, al servicio de finalidades o funciones no educativas.

La integración armónica de la evaluación con la actividad educativa la convierte en un medio más, junto al resto de los programados, diseñados y desarrollados

por el profesorado para alcanzar los objetivos de su programa (el programa en cuanto plan a su servicio). Incluso la evaluación sumativa puede formativizarse y extraer de ella las conclusiones pertinentes para la mejora del aprendizaje, de la enseñanza y hasta del propio profesor.

Así las cosas, no parece que deba extrañar el conceder a la evaluación un papel diagnóstico a la hora de asignar programas -evaluación inicial, preventiva- o de diseñar actuaciones correctivas o de mejora.

1.1.3. Evaluación integradora

En la medida en que la evaluación responde a las anteriores características, se convierte en una realidad que juega un papel activo y no meramente pasivo, como el que suele corresponderle cuando se limita a emitir juicios de valor sobre los resultados alcanzados.

La evaluación así entendida, dinamiza las actuaciones de profesores y alumnos, promueve la reflexión, el análisis y la autocrítica, fomenta la innovación y estimula el deseo de mejora continua, base de cualquier actuación pedagógica de calidad.

El carácter integrador de la evaluación ha sido puesto de relieve por aquellos expertos que llegan a decir que sólo cuando las innovaciones educativas alcanzan a la evaluación, su contenido se hace realidad.

Con frecuencia hemos asistido a situaciones en que los cambios en la evaluación han sido únicamente nominales: se cambió el concepto de examen por el de evaluación; el paso de los exámenes tradicionales a la denominada evaluación continua sólo se tradujo en la realización de más pruebas (alguien llegó a decir que se había pasado del Examen de Estado⁵ al estado de examen); de la evaluación continua se ha pasado a la evaluación formativa sin que eso se haya traducido en el ofrecimiento de la información y de la ayuda al alumnado a partir de los fallos, las carencias y los errores detectados en la evaluación.

Junto a ello, probablemente, el fallo más común ha sido el de plantear unos objetivos educativos más ambiciosos, ligados a lo que denominamos educación integral, mientras las técnicas de evaluación se autolimitan -carácter reductivo de las mismas- a los objetivos más bajos de la escala taxonómica⁶ dentro de los puramente cognoscitivos.

Pues bien, sólo los cambios que se reflejan en la evaluación se hacen realidad. Y es que la principal característica de una evaluación de calidad no es otra que la existencia de la máxima coherencia entre los objetivos planteados, los procesos y actividades educativas a su servicio, las técnicas utilizadas en la evaluación y el uso de sus resultados al servicio de los propios objetivos y de la mejora de los procesos.

1.2. Un paso más: la extensión de la evaluación a programas y centros educativos

Aunque cabe entender la evaluación de programas y centros como una actividad sistemática, organizada, rigurosa y científica con sentido en sí misma, o con finalidades propias diferentes de las aquí reseñadas⁷, no parece difícil asignarle objetivos coherentes y armónicos con los anteriores.

1.2.1. La evaluación de programas

Un programa es un documento técnico, elaborado por el profesorado, en el que se quedan reflejados tanto sus planteamientos docentes cuanto sus propias intenciones y metas educativas.



El programa es la integración de los objetivos educativos, a lograr a través de determinados contenidos, con un plan de acción a su servicio, plan en el que incluimos la evaluación de carácter formativo como un medio más, en sintonía y armonía con todos ellos.

No parece que tenga demasiado sentido desligar la evaluación del aprendizaje - tradicional objeto de evaluación- de la correspondiente a la enseñanza, dos caras de una misma moneda, si entendemos que ésta sólo cobra sentido en aquél, esto es, si admitimos que no se da verdadera enseñanza si no se plasma en el correspondiente aprendizaje.

En consecuencia, y al margen de la investigación evaluativa, con otras finalidades como objeto, cabe perfectamente integrar de modo armónico la evaluación de programas con la evaluación del aprendizaje de los alumnos. Con ella no sólo se mejorará el aprendizaje del alumno sino el programa del profesor y la propia actividad profesional de éste.

1.2.2. La evaluación de centros educativos

De nuevo, aunque los centros educativos pueden ser objeto de evaluación desde muy diversas perspectivas y al servicio de metas diversas, no cabe duda de que tales organizaciones pueden ser entendidas como un conjunto armónico de medios al servicio de un «gran programa», en concreto del Proyecto educativo.

El Proyecto educativo, en efecto, se concreta en los Proyectos curriculares, en los programas de los profesores, en los sistemas de disciplina y convivencia, en los planes de mejora... y, si las actuaciones están bien concebidas, todas ellas se ponen al servicio de una gran meta: la formación integral de las personas que son todos y cada uno de los educandos.

Así las cosas, se aprecia con toda claridad cómo la evaluación de centros o de instituciones educativas de diversa naturaleza puede ser considerada como un elemento integrador de la evaluación, que acoge en su campo de actuación a la evaluación de los programas y de los aprendizajes.

Concluyendo, la evaluación, teniendo únicamente el carácter de medio o instrumento, resulta ser de excepcional importancia para conseguir los objetivos educativos, siempre que responda a las tres grandes notas apuntadas: las de ser integral, integrada e integradora. Por tanto, si los objetivos recogidos en los Proyectos, programas e instituciones son de calidad, la evaluación es, ya, un destacado medio al servicio de la calidad, y de la eficacia, como ha puesto de relieve toda la investigación sobre escuelas eficaces.

A tales efectos se debe evitar todo reduccionismo en sus planteamientos, se ha de asegurar la máxima coherencia con los objetivos y los procesos educativos y ha de utilizarse al servicio de la mejora tanto de los educandos como de los propios educadores y, desde luego, de sus planteamientos técnicos, sean estos los programas diseñados al servicio de los objetivos o las propias organizaciones e instituciones en que desempeñan su tarea.

Como veremos, eso no impide que, cuando la sociedad demande a los educadores la asunción de funciones no estrictamente educativas, como son las propias de la evaluación sumativa (calificación, selección, acreditación), estos deban asumirlas con seriedad y las realicen con el debido rigor, traducido esencialmente en la recogida de información válida, valiosa y fiable al servicio de calificaciones justas.

2. Concepto de Evaluación y sus componentes

Como hemos tenido ocasión de indicar la evaluación es una actividad instrumental, por tanto al servicio de otra, sea ésta de carácter económico, médico, industrial, deportivo... o educativo. Pero, precisamente por ello, parece lógico pen-

sar que la naturaleza de la realidad evaluada represente alguna exigencia para la propia evaluación y, por tanto, para su concepto.

1.1. La función docente-formativa y la evaluación

La actividad docente y, en mayor medida, la educadora, podrían muy bien ser catalogadas como propias de las profesiones de ayuda. Nos referimos a profesiones destinadas a acompañar a otras personas a lograr sus objetivos. En nuestro caso, la educación, la ayuda se ofrece a los educandos, se trate de nuevas generaciones o de personas adultas, a progresar en su proceso de planificación y desarrollo como personas⁸, sirviéndose para ello de la orientación, la explicación, el asesoramiento, el apoyo o el estímulo.

En el caso de la docencia, su objeto fundamental se centra en el logro de las metas de aprendizaje y de formación intelectual por parte de los alumnos. A su servicio, el profesor diseñará toda una serie de actuaciones que, genéricamente, podríamos englobar bajo la denominación de plan destinado a lograr los objetivos previstos en currículos, proyectos y programas.

Bloom, centrando su atención en la actividad académica, expresa estas ideas con gran claridad cuando reflexiona sobre las diferencias existentes entre la estructura de una disciplina científica y la de su aprendizaje (Bloom y otros, 1981: 30). Veamos:

«La utilidad de una estructura para el aprendizaje tiene que ver con la habilidad de los estudiantes para comprenderla y utilizarla como factor organizador. No hay una relación clara entre la utilidad de una estructura para los estudiosos y su utilidad (y significación) para los estudiantes. [...] En contraste con la estructura de una disciplina se encuentra la estructura del proceso de aprendizaje, que deberá permitir al estudiante pasar con éxito de una fase de aquél a otra (...) la estructura de la enseñanza y del aprendizaje debe basarse en consideraciones pedagógicas y no tienen por qué reflejar el punto de vista del especialista sobre su campo de estudio».

Esta distinción entre el especialista en una materia y el profesor de la misma es de sumo interés y de notable trascendencia: si bien es cierto que nadie puede enseñar lo que no sabe, también lo es que la función docente no la ejercen con el mismo éxito todas las personas que conocen un campo del saber. Y ello es así porque tal función exige una serie de conocimientos, habilidades, destrezas y hasta talante humano, capaces de suscitar en los alumnos el interés por el estudio, de promover la motivación para el esfuerzo, de hacer asequible un determinado concepto, de captar las relaciones existentes entre diversas ideas, de descubrir el principio que explica un problema... Obviamente, las dificultades se multiplican, y las cualidades necesarias pueden ser mucho más complejas si de formación humana se trata.

Pues bien, responsabilidad del profesor es la de establecer las «consideraciones pedagógicas» a las que se refiere Bloom, destinadas a facilitar al alumnado el éxito en el aprendizaje, sea cual sea su naturaleza: conocimientos, destrezas, habilidades, actitudes, valores o comportamientos.

Las aludidas consideraciones se suelen concretar en el «programa», si bien en la medida en que lo entendamos no como un «temario» más o menos desarrollado de una materia o asignatura sino como un auténtico «plan de acción» al servicio de los objetivos que deben alcanzarse por medio de los contenidos seleccionados, adecuadamente organizados y estructurados.

El programa así entendido es una realidad compleja, como se desprende de la simple consideración del cuadro 1. Pues bien, como podemos apreciar en el cuadro, la evaluación puede, y hasta debe, formar parte del conjunto de medios integrantes del plan al servicio del logro de los objetivos educativos..., aunque, además, sirva a funciones diferentes, como las relativas a la comprobación de los niveles de logro alcanzados, la calificación académica y las consecuencias de la misma (promoción o no, selección...)



Complejidad del programa
Necesidad de compaginar los aspectos científico / técnicos propios de la materia de enseñanza con otros psico-pedagógicos (el alumno que aprende y la forma de facilitar el aprendizaje), y sociales (las demandas de la sociedad en relación con la materia).
Los contenidos, de diversa naturaleza. a. El enfoque de estudio, ligado a los objetivos. b. Los medios y recursos, facilitadores del aprendizaje: <ul style="list-style-type: none">- Información.- Metodología y recursos didácticos.- Motivación.- Enseñanza.- Actividades.- Materiales.- Clima de centro y aula.- Evaluación (formativa). c. La evaluación (sumativa).

Cuadro 1. El programa como plan de acción al servicio de los objetivos educativos

1.1. El doble papel de la evaluación

En el cuadro 1 apreciamos una importante realidad: el doble papel de la evaluación. Si tradicionalmente viene siendo practicada en su función sumativa, como una actividad yuxtapuesta, añadida, claramente diferente y diferenciada de la actividad de planificación, diseño y enseñanza, su inclusión en el esquema como medio quiere hacer notar que, por medio de ella, el profesor debe «enseñar», debe «educar»; dicho de otro modo, que la evaluación, al igual que la información, la explicación, la motivación, los propios materiales de aprendizaje, las actividades, y, en general, la metodología y recursos didácticos, debe ayudar a los alumnos a alcanzar las metas u objetivos previstos por el profesor en relación con los temas o contenidos de la disciplina que integran el programa.

La consideración de medio nos sitúa ante una importante cuestión: la relativa a las funciones a las que sirve la evaluación, que, como vamos a tener la ocasión de señalar, no sólo son las dos anteriores sino que incluye otras que, eso sí, son de menor importancia en el marco de las tareas educativas.

1.1.1. Las funciones de la evaluación y sus consecuencias

Tanto en el concepto de evaluación, al que en breve nos referiremos, como en el de las funciones a las que sirve, se da una gran diversidad de planteamientos⁹. A David Nevo, en una interesante aportación a la obra dirigida por House¹⁰, se debe la caracterización de la evaluación a través de las diez grandes cuestiones recogidas en el cuadro 2, reseñando en el segundo de los puntos, el relativo a las siguientes funciones: sumativa, formativa, sociopolítica y administrativa (Nao, 1986: 15-30).

- | |
|--|
| <ul style="list-style-type: none">- Cómo se define- Cuáles son sus funciones- Cuáles son los objetos a evaluar.- Qué tipo de información debe recogerse.<ul style="list-style-type: none">• Qué criterios deben utilizarse para juzgar el mérito o el valor de un objeto• A quién debe servir la evaluación.• Cuál es el proceso a seguir.• Qué métodos deben utilizarse.• Quién debe realizarla.• Qué estándares de calidad han de tomarse para valorar las evaluaciones. |
|--|

Cuadro 2. Caracterización de la evaluación según D. Nevo

a) Evaluación formativa o para la mejora

Si aceptamos la caracterización de la docencia como una profesión de ayuda, no debería extrañarnos la preferencia por la función formativa de la evaluación, una aportación de Scriven. La evaluación formativa tiene como meta la mejora, bien sea del aprendizaje individual de cada uno de los alumnos bien lo sea del programa como unidad.

La función formativa de la evaluación implica su consideración como medio; lo que representa, por un lado, una mayor continuidad en el proceso de recogida de información para potenciar el paralelismo entre los procesos de enseñanza y de aprendizaje; por otro, el traslado de la información al alumnado, haciéndole sabedor y consciente de sus carencias, errores y limitaciones, así como de sus puntos fuertes, dado el carácter motivador que suele representar este hecho; y, por último, su utilización para la toma inmediata y a tiempo de decisiones de mejora sobre cada alumno y sobre el programa en cada una de las unidades en que esté estructurado. Es más, la propia continuidad de la recogida de información puede verse superada, yendo más allá hasta aplicarse, con un sentido anticipador y preventivo, antes de la aplicación del programa docente, dando lugar a lo que se conoce como evaluación inicial.

Y es que, en efecto, la información recogida por el profesor en el proceso evaluador, en la medida en que esté adecuadamente diseñada y se plantee en forma sistemática y organizada, le permite corregir deficiencias, reorientar su actividad y la de sus alumnos, prestar apoyo, ofrecer información suplementaria... todo como un medio para lograr el éxito.

Dadas las limitaciones de tiempo, en esta tarea, correctiva en muchos casos y de desarrollo y optimización en otros (con los alumnos que van bien, con la finalidad de profundizar y afrontar nuevos objetivos de superior amplitud), resulta de especial utilidad acudir al sistema de enseñanza mutua o de monitores, mediante el cual los alumnos especialmente capacitados pueden ayudar -y lo hacen con éxito¹¹- a sus compañeros practicando, de paso, los valores de compañerismo y solidaridad entre colegas.

No parece difícil apreciar en esta forma de proceder la integración armónica de la evaluación en el conjunto de las tareas docentes, algo mucho más difícil de lograr cuando la evaluación es sumativa.

b) Evaluación sumativa

La sociedad ha depositado en el profesorado una importante tarea, alejada en sí misma de la función docente en cuanto tal, pero de gran trascendencia social. Se trata de la sumativa en terminología de Scriven, orientada a la certificación, rendición de cuentas, valoración global de un programa e, incluso, a la selección de las personas que han de ocupar determinados puestos.

Cada vez que un profesor se limita a calificar los resultados de una prueba o a emitir una valoración global del aprendizaje alcanzado durante una unidad de tiempo más o menos amplia, como un trimestre, semestre o curso académico, está ejerciendo esta función; cuando, en función de tales resultados, un estudiante promociona o no, se selecciona o no, alcanza un diploma o no..., nos encontramos ante la función sumativa de la evaluación. Del mismo modo, estaremos en una situación semejante cuando, tras la aplicación de un proyecto o programa, se decide si conviene mantenerlo como está, modificarlo o suprimirlo.

c) Otras funciones de la evaluación

Si las dos anteriores funciones son las principales en el marco de las tareas encomendadas al profesorado, en ocasiones la evaluación puede ser ejercida al servicio de funciones diferentes, siendo las principales las denominadas por Nevo como sociopolítica y como administrativa.

En el primer caso, se trata de evaluaciones destinadas a conseguir el apoyo, institucional y económico, para determinados planes, proyectos o programas.



Las demandas sociales son tantas que, siendo limitados los recursos disponibles, los organismos encargados de financiar proyectos suelen exigir argumentaciones sólidas sobre su valor y, a posteriori, evidencias de que el esfuerzo ha merecido la pena y hasta es conveniente seguir manteniéndolo.

Desde esta perspectiva, la evaluación cumple una importante función social en la medida en que exige responsabilidad en el uso de los recursos y evita el despilfarro de los mismos, siempre insuficientes por elevados que puedan ser en términos absolutos.

Por otro lado, la sociedad debe asegurar mediante los procedimientos pertinentes el cumplimiento de la normativa establecida. Si bien lo deseable, y lo esperable en la mayoría de los casos en las sociedades maduras y avanzadas, es que los ciudadanos cumplan las normas, no es menos cierto que la realidad nos pone día a día ante flagrantes incumplimientos de la legalidad.

No debe extrañar, por tanto, que puedan llevarse a cabo, y hasta que sean necesarias, evaluaciones centradas en el control del grado de cumplimiento de las normas, de las obligaciones docentes y discentes, de la legislación vigente hablando ya en términos generales. Algunas manifestaciones de este tipo de evaluaciones han dado lugar a los sistemas de auditoría, realizados desde fuera, por expertos, y en momentos concretos, generalmente finales, de los procesos objeto de atención. Nos encontramos ante la función administrativa o para el control.

No parece difícil mantener, a partir de lo hasta ahora señalado, que son las dos primeras funciones, y, en particular, la primera, las más propias de la actividad de los profesores. Hasta tal punto son así las cosas que, en el marco del nuevo sistema educativo español, en vigor desde 1990 con la progresiva implantación de la Ley de Ordenación General del Sistema Educativo (LOGSE), la función inspectora no se limita al control sino que se mueve entre este y el asesoramiento en temas de evaluación de centros, en lugar de quedar reducida su tarea a aquella, como podría desprenderse de su propia denominación (art. 61). Pero ello no obsta para que los profesores, a título individual o de equipos, se enfrenten con evaluaciones de esta otra naturaleza.

2. Nuestro concepto de evaluación

Ha llegado el momento de presentar, de forma ordenada y sistemática, el concepto de evaluación al que venimos refiriéndonos continua y reiteradamente sin haberlo definido ni, mucho menos, desarrollado.

Parece una obviedad que nuestra concepción deba estar en armonía con nuestra idea de educación. Así las cosas, la evaluación será integralmente pedagógica en la medida en que cumpla dos condiciones fundamentales:

- Centrarse en todos los factores relevantes que integran los procesos educativos.
- Orientarse predominantemente a la mejora de los educandos.

De modo subsidiario deberá proponerse la mejora, la optimización de otras personas -el personal de los centros-, de los programas, de los ambientes, de los materiales y recursos..., considerados, siempre, como medios o factores que contribuyen a la mejora de los educandos.

Pues bien, con tales planteamientos de base, nuestro concepto de evaluación, recogido a continuación, debe ser lo suficientemente amplio como para acoger todos los factores aludidos. Este es el concepto que defendemos:

«La evaluación pedagógica consiste en 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».

Este concepto puede apreciarse en el cuadro 3. Dejemos, por el momento, la debida constancia, tanto de los elementos clave -información, criterios de valora-

ción, toma de decisiones de mejora- como de los objetos de la evaluación, respondiendo así a la nota de integral a que nos referíamos en el punto 1: cuántos factores relevantes integran los procesos educativos, superando claramente la idea de una evaluación centrada en los alumnos y, más en concreto, en sus aprendizajes, por lo general limitados a los niveles más bajos (conocimiento, comprensión, aplicación) de las taxonomías educativas, como muy bien han puesto de relieve las diversas investigaciones al respecto.

Función esencial de la evaluación: la mejora	
Momentos	Elementos
Previo	<ul style="list-style-type: none"> • Los objetivos: <ul style="list-style-type: none"> - Del objeto evaluado - Del propio sistema evaluativo.
Núcleo de la evaluación	<ul style="list-style-type: none"> • Recogida de información • Valoración, por medio de criterios y de referencias preespecificados • Toma de decisiones orientadas a la mejora (planes)
Posterior	<ul style="list-style-type: none"> • Seguimiento de los planes de mejora.

Cuadro 3: Concepto de evaluación pedagógica

Tres son las ideas básicas que nos proponemos desarrollar: a) aunque la evaluación se concreta en lo que hemos denominado «núcleo» de la misma, hay dos momentos más de gran importancia, uno «previo» y otro «posterior»; b) la calidad de la evaluación nunca superará la de la información recogida; y c) no hay evaluación pedagógica sin toma de decisiones de mejora.

3.1. Los tres momentos de la evaluación

Con reiteración, tal vez hasta excesiva, venimos insistiendo en el carácter de medio que atribuimos a la evaluación. Este importante principio se traduce en el cuadro 3 en la importancia concedida a un momento previo a la evaluación y que, como se puede apreciar, se concreta en los objetivos no sólo de la evaluación (para qué se evalúa, algo ligado a la función formativa) sino, sobre todo, del proceso educativo.

Son los objetivos educativos los que se convierten así en directriz de la evaluación, cuya calidad dependerá de su adecuación, coherencia y suficiencia con aquellos. En efecto, como ocurre con todo tipo de instrumentos, medios o herramientas, aparte de sus propias características técnicas, a las que, en breve, nos vamos a referir, la principal de las exigencias es que resulten adecuados a la realidad a la que sirven, que sean coherentes con ella (que no desvirtuen la realidad evaluada¹²) y que su aportación resulte suficiente.

Cuando este hecho se respeta, queda patente no sólo el carácter integrado de la evaluación –la evaluación al servicio de los objetivos- sino su carácter integral –por coherencia con los objetivos del proyecto educativo- y hasta su potencial integrador, ya que, al darse tal coherencia, educandos y educadores, en pos del éxito en el logro de aquellos, logran la dinamización de todos los demás medios recogidos en el cuadro 1.

Dejando para el siguiente apartado el análisis del núcleo de la evaluación, haremos alguna consideración sobre el momento posterior, concretado en el seguimiento de los planes de mejora.

De nuevo se aprecia ese carácter subordinado de la evaluación a los objetivos educativos: no se evalúa para calificar, seleccionar, acreditar..., sino para mejorar. Y como las decisiones de mejora, parte fundamental del núcleo de la evaluación, no suelen ser algo que dé sus frutos de modo inmediato, es preciso convertirlas en planes de acción y, por lo tanto, con sus propios objetivos, medios y sistema de evaluación,... por si el diseño inicial no da el fruto apetecido.



2.2. El núcleo central de la evaluación

Evaluar es valorar¹³, pero esta actividad no es posible llevarla a cabo sin disponer de la pertinente información sobre la realidad a valorar. De ahí que en este punto clave incluyamos dos elementos esenciales: la información, los datos -disponibles o a recoger- y los criterios y las referencias con las que emitir los correspondientes juicios de valor.

2.2.1. La información

Como muy bien afirma Bunge refiriéndose a la investigación, los datos, a pesar de la etimología «dato» viene del latín *datum*, algo dado- «no son nada dado, sino que hay que producirlos, y muchas veces laboriosamente» (Bunge, 1979: 743).

Podríamos afirmar que la elaboración rigurosa de instrumentos para la recogida de datos se constituye en los cimientos de una evaluación de calidad, útil para la mejora del aprendizaje y de la enseñanza, e imprescindible para otorgar calificaciones válidas y justas.

Pues bien, en ese «construir los datos», en ese «recoger información» para su posterior valoración o evaluación, se aprecia, de nuevo, la necesidad de subordinación de los mismos a los objetivos (momento previo), de modo que la cualidad fundamental que tiene que adornar la información o los datos es la de su coherencia con aquellos, además, claro está, de las condiciones técnicas de fiabilidad, validez y valía¹⁴. Esta coherencia debe apreciarse en aspectos no siempre tomados en consideración, como es el referente, por un lado, a las técnicas elegidas para la evaluación y, por otro, a su propio contenido.

Si nos referimos a las pruebas, debemos ser conscientes de que no existe ninguna modalidad de estas que sea en sí misma ni buena ni mala, sino adecuada, coherente y suficiente para evidenciar el nivel de logro de los objetivos, las carencias y errores existentes y, en la medida de lo posible, el origen o la causa de unas y otros.

Si tenemos en cuenta que, en cualquier materia, cabe una amplia variedad de objetivos, desde la simple memorización de ciertos contenidos a la capacidad para analizar, sintetizar, relacionar, demostrar, argumentar, valorar... comprenderemos que un mismo tipo de pruebas -por ejemplo, las denominadas objetivas- pueden ser adecuadas para algunos de ellos y totalmente inadecuadas para otros; pero, junto a ello, que cualquier uso de un tipo de pruebas de forma claramente predominante o exclusivo, lleva aparejado un sesgo, una orientación del estudio hacia esos objetivos en detrimento del resto. De ahí nuestra recomendación de que, a la hora de recoger la información, acudamos al principio de complementariedad metodológica que, en este caso, se concreta en la posibilidad de utilizar técnicas variadas¹⁵, trabajos que obliguen al alumno a realizar aprendizajes según los objetivos¹⁶: el alumno debe saber que, según sea el tipo de objetivos que el profesor evalúe, deberá haber alcanzado el dominio de tipos de aprendizaje muy diversos y que ello se pondrá de manifiesto en técnicas de recogida de información muy diferentes unas de otros.

Cuando de las pruebas pasamos a su contenido, preciso es hacer referencia expresa a la necesidad de que responda a las exigencias técnicas de validez, valía y fiabilidad; de no ser así, la valoración posterior se verá afectada por esas mismas carencias y limitaciones. Sin embargo, la investigación pone de relieve que las cosas no suelen ser así, detectándose la clara tendencia a centrar la atención en los objetivos de menor nivel de las taxonomías, dejando de lado los procesos más complejos y, por tanto, más humanos.

La validez de la información en el caso de la evaluación del aprendizaje va ligada a una específica modalidad, conocida como validez de contenido, esto es, como el grado en que la prueba de que se trate refleja con fidelidad, con la debida representación y ponderación la realidad objeto de evaluación, sea una asignatura completa, una parte más o menos amplia o el contenido de una unidad, como puede ser un tema o unidad didáctica.

La importancia de este hecho es fundamental; véase al respecto lo manifestado por Hills (1982: 312). «Las preguntas de las pruebas surten un importante efecto sobre la índole del verdadero currículo de la escuela, quizás en oposición al currí-

culum nominal, debido a un principio psicológico fundamental del aprendizaje: los individuos aprenden lo que se premia. [...] quien controla las pruebas y evaluaciones en una escuela, en realidad controla el programa. [...] Si las pruebas de un maestro no concuerdan con el currículo fijado para la enseñanza, lo que se aprenda coincidirá con lo que se examine no con lo que dice el currículo».

Dos son las recomendaciones a seguir si se desea asegurar tales características, en especial cuando las pruebas se tienen consecuencias para los alumnos (calificación, selección... propias de la evaluación sumativa). En primer lugar elaborar un documento técnico, conocido como tabla de especificaciones, en el que quede debidamente reflejada la relación entre los contenidos de la materia y los objetivos a lograr a través del proceso educativo; la evaluación no hará sino asegurar una presencia suficiente, representativa y debidamente ponderada de unos y otros en los diversos instrumentos de evaluación¹⁷. La segunda, respetar determinadas características de la información a recoger, en especial las recogidas en el cuadro 5.

La tabla de especificaciones debe reflejar la relación entre una serie de contenidos, relativos a diversas unidades de aprendizaje y distintos objetivos, entendidos como destrezas, habilidades y funciones mentales. La tabla reflejará, además, si ciertas metas se consideran no obligatorias, por lo que, al construir las oportunas pruebas, y, sobre todo, al valorarlas y calificarlas, su dominio o su carencia tendrán un tratamiento diferente.

En determinadas tablas de especificaciones el profesor puede indicar en cada uno de los cruces el número de ítems a utilizar para tener la seguridad de que, por medio de ellos, ese contenido/objetivo resulta suficientemente evaluado. Una forma alternativa de proceder consiste en ponderar el resultado correspondiente, concediendo un peso superior al dominio o carencia de la realidad evaluada por el ítem.

La segunda cuestión, decíamos, tiene que ver con las características de la información a recoger, de las que forman parte, por cierto, las ya citadas de representatividad, suficiencia y ponderación. En el citado cuadro 4 aparecen las cualidades y su correspondiente descripción.

Cualidad	Descripción
Coherencia	Plena concordancia entre los objetivos planteado, las actividades docentes desarrolladas y el contenido de la evaluación.
Suficiencia	Su amplitud debe permitir al profesor contar con la necesaria seguridad para emitir un juicio sobre el grado de dominio del ámbito de aprendizaje objeto de evaluación.
Representatividad	La composición de la información ha de permitir apreciar el grado de dominio de las diferentes partes en que se estructura el campo evaluado -contenidos y objetivos- con una presencia equilibrada de todos ellos.
Ponderación	Se trata de que el peso de los diferentes elementos del contenido contribuyan en forma adecuada a la configuración de las diferentes calificaciones.
Relevancia social	El contenido debe seleccionarse de forma que el saber apreciado sea importante para la sociedad ante la que se acredita.
Relevancia científico/académica	El contenido debe seleccionarse sobre la base de lo que en cada disciplina es importante y representa el dominio de la misma.
Relevancia pedagógica	El contenido debe permitir la identificación de las carencias y errores del alumno, de las dificultades experimentadas por él en el proceso de aprendizaje y de sus posibles causas



2.2.2. La valoración

La disponibilidad de la información nos sitúa ante el momento crucial de la evaluación: el de la valoración, que hemos de realizar aplicando referencias y criterios que, para bien, deben ser públicos y estar preespecificados. Públicos, para orientar el aprendizaje del alumno y disciplinar al profesor; y preespecificados, para dirigir los procesos docentes y formativos en coherencia con los objetivos del programa.

a) Los criterios.

Entendemos por criterio una norma para juzgar una cosa. La fijación de criterios ayuda a discernir en relación con la realidad evaluada. El establecimiento de los criterios por parte del evaluador debe estar, obviamente, en sintonía, en armonía, en coherencia, con los objetivos tanto los más amplios, (los educativos), como los más limitados (los de la enseñanza).

Por lo general, determinados «productos» educativos pueden ser valorados aplicando diversos criterios. Por tanto, la selección de los que se van a utilizar en cada caso tendrá que ver con las metas a lograr en ese momento. Así, por ejemplo, una redacción puede ser juzgada atendiendo a su corrección tanto ortográfica como morfológica o sintáctica; a la precisión y a la riqueza del léxico, a la sistematización y ordenación de las ideas; al estilo personal; a la originalidad... Pues bien: lo probable, y parece que lo adecuado, será ir trabajando cada uno de ellos y someterlos a valoración una vez que se hayan realizado las actividades pertinentes. No tiene sentido valorar la riqueza expresiva si únicamente estamos trabajando la corrección morfológica y ortográfica.

En el caso de los problemas es importante decidir si la atención se debe localizar hacia la exactitud o la rapidez de los cálculos, la corrección del planteamiento o la interpretación de los resultados.

Merced a los criterios, discriminamos lo correcto de lo incorrecto, lo suficiente de lo insuficiente, lo adecuado de lo inadecuado... y en función de ello tomamos decisiones. Pero la clave está en tenerlos claros, decididos y publicados para su adecuada aplicación y para orientar la actividad de los alumnos.

b) Las referencias.

Una dimensión relacionada con los criterios es la relativa a las referencias. A la hora de tomar decisiones de carácter global sobre las realizaciones del alumnado se plantean problemas nada fáciles de resolver, como los siguientes:

- Si tuviéramos que seleccionar a un alumno por relación con otros, ¿dónde se situaría cada uno a lo largo de un continuo?
- ¿Es suficiente el nivel alcanzado por un alumno?
- ¿Ha alcanzado un alumno un nivel satisfactorio?

Preguntas similares podríamos formular en relación con un programa: ¿qué programa alcanza mejores resultados?; ¿se alcanzan las metas prefijadas?; dados los medios y recursos disponibles, ¿se puede dar por satisfactorio lo alcanzado?...

Básicamente nos hemos planteado las tres grandes referencias para calificar: la primera es conocida por lo general como normativa. Sitúa los logros de un alumno, o del conjunto de los que trabajan un mismo programa, en el marco de lo que consiguen los que forman su grupo, constituidos en baremo o regla para medir. Es una referencia «relativa» dado que, por definición, en un baremo siempre hay resultados por encima y por debajo de la media: se puede ser el último de un grupo -el superior- o el primero del grupo siguiente (el medio o el inferior). Pero, cuando hay que comparar a unos con otros no parece haber otra forma mejor de evaluar, aunque lo verdaderamente difícil sea, justamente, establecer el baremo¹⁸.

La segunda referencia es la criterial¹⁹. ¿Domina un alumno, o el grupo/clase, los niveles mínimos fijados por el profesor?. Esto es, ¿alcanza la suficiencia?

Con frecuencia lo que un profesor califica como suficiente –el alumno, o el grupo, alcanza el criterio o nivel general fijado de antemano– puede ser valorado por otro profesor como insuficiente, como bueno y hasta como excelente; es más, lo valorado hoy como suficiente por un profesor puede ser calificado como no suficiente o como bueno por el mismo profesor en una ocasión posterior²⁰.

Una forma clara de evitar estos hechos, que pueden tener repercusiones negativas para ciertos alumnos en procesos de selección por méritos, radica justamente en el establecimiento claro y preciso del criterio²¹ para decidir los diferentes niveles, en particular el más decisivo: el de la suficiencia o insuficiencia.

En la fijación de tal nivel los profesores deben acudir a criterios técnicos, relacionados con la importancia objetiva de determinados contenidos, su utilidad práctica, su valor como condicionante o prerrequisito de futuros aprendizajes... Dado que las materias de un curso pueden haber sido estudiadas en cursos anteriores y tener su continuación en posteriores, y dada la relación existente entre materias de un mismo curso, parece conveniente que la fijación de tales niveles se realice por equipos de profesores en lugar de establecerlos de modo independiente cada profesor.

Una valoración criterial debe facilitar el progreso de los alumnos a lo largo de cursos, ciclos y etapas en que se divide su discurrir a lo largo de su formación.

Una última referencia es la que se conoce como personalizada o idiosincrática. Con ella se pretende valorar como satisfactorio o insatisfactorio el nivel alcanzado por una persona o por un grupo concretos.

En el caso de personas concretas, si se tiene en cuenta que el rendimiento depende de sus capacidades y de su base o rendimiento previo, pero también de sus intereses o nivel de motivación, además de estar condicionado por las circunstancias sociales y familiares²², no debe extrañar que cuando se trata de valorar no sólo al aprendiz sino a la persona que cada aprendiz es, convenga tener en cuenta no sólo su dimensión intelectual sino la afectiva, la técnica y sus circunstancias familiares. De esta forma es posible satisfacer alguna de las necesidades básicas del ser humano, como la de ser reconocido y apreciado²³, algo que da fuerza y motivos para seguir esforzándose.

Obviamente, una referencia personalizada no debe mezclarse con otra criterial, aunque puedan yuxtaponerse. No tendría sentido decir que el rendimiento de un estudiante de cirugía es suficiente por el mero hecho de que se ha esforzado y, en función de sus circunstancias personales y familiares, el nivel de logro alcanzado merece más reconocimiento que el de otro alumno que, sin apenas esfuerzo, ha conseguido no ya el suficiente sino una calificación elevada. Cabe, de hecho, que un alumno con notable merezca, a la vez, una valoración de insatisfactorio. Sin embargo, en la medida en que las calificaciones tienen una función de acreditación social, el valor decisivo socialmente hablando es el de suficiencia²⁴, lo que no limita el profundo valor humano de la referencia personalizada en el proceso educativo de hacerse persona.

2.2.3. Evaluación y toma de decisiones

Una evaluación que termine en el acto de valoración no es, en sentido estricto y pleno, una evaluación pedagógica ya que, como tal, no contribuye a la mejora del aprendizaje de los alumnos. En efecto, una evaluación de tal naturaleza se sitúa fuera del proceso educativo: el profesor deja de actuar como tal para comportarse como un juez, que decide sobre el grado de cumplimiento del programa o del nivel alcanzado por cada alumno, pero en modo alguno le permite la mejora de aquel o la corrección de los fallos y carencias de éste.

Lo que en verdad sitúa la evaluación como un medio más es la utilización de los resultados de la evaluación para detectar lo que va bien y lo que va mal, los puntos



fuertes y débiles, los aspectos positivos y las áreas de mejora, y, si la información es correcta, suficiente y adecuada, la naturaleza de los aspectos positivos y negativos, su «causa» y, por tanto, los previsibles remedios. He ahí, por tanto, el punto de arranque de las decisiones propias del profesor: optimizar lo que va bien, corregir lo que va mal, enderezar lo que no va suficientemente bien, y ello lo mismo sobre alumnos concretos que sobre el grupo como unidad, en cuanto sujeto del programa.

Aquí se aprecia, de nuevo, la importancia de la información: pruebas mal construidas, poco válidas –por insuficientes, sesgadas, poco representativas...– dan lugar a una información que, si bien permite asignar un número (calificación) a los trabajos del alumnado, en modo alguno ayudan a identificar tales puntos fuertes y débiles y, menos aún, a tomar decisiones precisas para potenciar aquellos y mejorar estos²⁵. Nunca la evaluación será mejor que la información en que se basa.

En este momento concreto los profesores pueden colaborar activamente, sobre todo en lo que tenga que ver con los objetivos evaluados²⁶; puede apreciarse, en efecto, si los alumnos tienen dificultades para aplicar conocimientos, para comprender teoremas, para clasificar, organizar y elaborar de forma personal la información, para asimilar los contenidos, para demostrar, para razonar, para descubrir las ideas fundamentales, para reconocer la naturaleza de un problema, para operar con precisión y rapidez, para interpretar gráficos... Cuando las cosas son así, las decisiones pueden ser «transversales», con lo que la eficacia es muy superior, por el mutuo refuerzo y por la percepción del alumno de que el conjunto de sus profesores colabora a un proyecto común.

Las denominadas sesiones de evaluación, si se plantean en estos términos, resultan ser muy eficaces y contribuyen a una visión unitaria y armónica del proceso educativo: todos los profesores están interesados en el progreso de cada uno de sus alumnos, y, a su servicio, trabajan en auténtico equipo.

2.3. Seguimiento

Por lo general, las decisiones no producen un efecto inmediato, en particular cuando lo que se pretende es introducir cambios, sean perfectivos o correctivos, de largo alcance. Es por ello que tales decisiones son, de hecho, auténticos programas, que necesitan un seguimiento y hasta una evaluación para decidir sobre su eficacia y hasta para su modificación, si del seguimiento se desprendiera su inadecuado desarrollo o la falta de resultados.

3. Una propuesta de actuación

Una vez dejados claros los aspectos conceptuales, básicos para una práctica educativa con el debido soporte, nos proponemos en este apartado ofrecer una propuesta, armónica con todo lo anterior, y suficientemente concreta para iluminar la práctica educativa.

3.1. La evaluación en el marco educativo

El Proyecto Educativo, en cuanto elemento directriz de la vida de una institución, debe incluir los planteamientos básicos sobre el tipo de educación que se desea ofrecer, sobre los medios adecuados para alcanzarla y sobre el sistema para decidir sobre los niveles de logro y la propia eficacia de los medios planificados. Allí debe quedar de manifiesto el papel de la evaluación y su concepción global.

A mi juicio, dos deben ser las cuestiones allí reflejadas: a) la preeminencia de la función formativa de la evaluación sobre cualquier otra, y b) la concepción de la evaluación como un todo, armónico y coherente con el Proyecto, y que, por tanto,

se debe aplicar al centro, a los programas de los profesores y al aprendizaje de los alumnos, huyendo de la tendencia habitual a descuidar, cuando no olvidar, los dos primeros aspectos.

3.2. La evaluación institucional o del centro educativo²⁷

La dirección de los centros puede plantearse una evaluación entendida como control o, también, en términos o funciones sumativas; es más, puede verse afectada por iniciativas evaluadoras externas, como las de la Inspección.

En tales ocasiones, conviene actuar de forma que los resultados, e incluso los propios procesos, se formativen, esto es, permitan extraer consecuencias para la mejora del centro como unidad y de cada uno de sus elementos componentes fundamentales.

Sin embargo, lo verdaderamente importante es institucionalizar procesos de evaluación formativa de los centros, utilizando herramientas muy sencillas –para que no ocupen demasiado tiempo y no distraigan al profesorado de su función: ayudar al alumnado en sus procesos de formación– que tengan como consecuencia la instauración de la mejora continua. Estas evaluaciones internas, con el apoyo de expertos si fuera necesario, deben focalizar su atención en los elementos clave de la vida de los centros:

- Los Proyectos educativos.
- La organización y la gestión, al servicio de aquellos. La participación de la comunidad en el desarrollo del proyecto.
- Los programas de los profesores, debidamente organizados y coordinados al servicio de las metas comunes, incluidos los propios de los especialistas que colaboran con el profesorado (orientadores, logopedas...).
- El ambiente de los centros, como una realidad a construir activamente a través de actuaciones específicamente diseñadas.
- Los sistemas de mejora continua, con la evaluación formativa al frente.
- Los resultados realmente alcanzados.

3.3. La evaluación de los programas de los profesores²⁸

En el programa se encarna la acción educativa al servicio de los objetivos concretos de cada profesor que, obviamente, deben resultar armónicos y coherentes con los del Proyecto Educativo.

La actitud de mejora del profesorado se plasma en la asunción del compromiso de evaluación de su programa, una evaluación que puede concretarse en tres momentos diferentes: a) antes de iniciar el programa, para decidir sobre su calidad técnica, sobre su viabilidad y su evaluabilidad; b) durante el proceso de realización, para tomar a tiempo, las decisiones de mejora (evaluación formativa); y c) tras su aplicación, para apreciar su eficiencia, efectividad y eficacia. Este ciclo evaluativo –evaluación inicial, procesual o continua y final– deja de serlo cuando, como hemos indicado en lo relativo a los centros, se institucionaliza: en tal caso, se hace evidente que la evaluación «final» es, a la vez, la inicial del siguiente ciclo evaluativo.

3.4. La evaluación del aprendizaje de los alumnos²⁹

Con este marco de actuación como contexto de referencia, cobra pleno sentido la evaluación del alumnado, formativa y orientada a la potenciación y mejora de



sus logros, aun en el supuesto de que se hayan alcanzado los niveles prefijados (evaluación de referencia criterial) por parte del profesorado.

De hecho, la valoración del aprendizaje del alumnado³⁰ es un componente básico de la evaluación del programa del profesor, al igual que aquellos -los aprendizajes- y éstos -los programas- lo son para decidir sobre lo que va bien, regular o mal dentro de un centro educativo.

En lo relativo a los aprendizajes de los alumnos, preciso es recordar la importancia de asegurar que tanto las técnicas de evaluación como sus contenidos: a) sean adecuadas, suficientes y coherentes con los objetivos del programa y, en consecuencia, con los del Proyecto educativo; b) que los criterios y referencias a utilizar en la valoración sean públicos, estén preespecificados y se apliquen adecuadamente; c) que la evaluación no concluya en la valoración sino que dé lugar a las correspondientes decisiones, tanto de optimización de lo que va bien como de corrección de lo que va menos bien o, incluso mal; y d) que, en la medida en que tales decisiones impliquen procesos optimizadores o correctivos de alguna duración, cuenten con el adecuado seguimiento, lo que permitirá lograr la eficacia y mejorar los propios procedimientos de mejora para sucesivas aplicaciones.

4. La Evaluación de Programas

Después de presentar nuestro marco de referencia, dedicaremos una atención específica a la «evaluación de programas», campo éste susceptible de enfoques, desarrollos y actuación muy diversos. De entrada conviene dejar constancia de un posible doble enfoque: la evaluación de programas ordinarios de los profesores, realizada desde dentro, por ellos mismos o por el grupo de compañeros de Seminario o Departamento, y la evaluación de programas realizada por expertos, desde fuera, con mentalidad propia de investigación, en concreto de la denominada «investigación evaluativa».

Si, desde el primero de los enfoques, las posibilidades de mejora de la práctica educativa³¹ se nos antojan enormes, en el segundo de los sentidos, la evaluación de programas educativos tal vez sea, o deba serlo, una de las modalidades de investigación pedagógica más utilizada y, desde luego, con mayores posibilidades y aplicaciones.

No debe extrañarnos que así sea si tenemos en cuenta que la educación, por su propia naturaleza, es una actuación organizada y sistemática al servicio de la mejora o perfeccionamiento del ser humano, concepción totalmente compatible con la de programa, que no es sino un plan sistemático de actuación al servicio de metas pedagógicamente valiosas. Podríamos afirmar que los educadores encarnan sus planteamientos educativos en programas de mayor o menor amplitud, complejidad y duración. Otra cosa será si los programas de los educadores responden a las exigencias técnicas que les son exigibles.

La complejidad de la educación en general, y de los programas concretos en particular, aconseja abordar su evaluación con diversos fines, entre los que destacan los de conocer sus resultados, potenciar sus efectos, rentabilizar los recursos y corregir los fallos, disfunciones y actuaciones inadecuadas, insuficientes, insatisfactorias o ineficaces.

En esa línea van los movimientos de calidad, con sus diversos modelos más o menos elaborados e implantados, al reconocer en la evaluación la principal estrategia o herramienta al servicio de uno de sus principios fundamentales: el de la mejora continua.

Ahora bien, presentar una propuesta sobre evaluación de programas es una ardua tarea dada la amplitud de los aspectos que pueden ser objeto de consideración. Lo sensato, por tanto, será seleccionar algunos aspectos y concentrar en ellos nuestra atención. En concreto, estudiaremos los siguientes:

- El hecho educativo como objeto de la evaluación.
- Las principales variables a tomar en consideración a la hora de configurar una metodología evaluativa.
- Los elementos o componentes básicos de un modelo o propuesta evaluativa de programas educativos.

Vale la pena, no obstante, dejar previamente constancia, para que sirva de referencia al discutir sobre los contenidos del presente documento, del concepto de evaluación que mantenemos, que basado en la previa definición genérica de evaluación, se concreta en las siguientes líneas: «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» (Pérez Juste, 1995: 73-106)³².

4.1. El hecho educativo

Con la palabra Programa evocamos un concepto de muy amplio espectro; en el campo educativo, el término programa se concreta en un plan de acción, por tanto en una actuación planificada, organizada y sistemática, al servicio de metas educativas.

De tal consideración sobre el Programa se desprenden algunos importantes hechos que, obviamente, habrán de ser tomados en consideración a la hora de su evaluación; en concreto, todo programa:

- a) Debe contar con unas metas y objetivos que, obviamente, han de ser educativos. Tales metas y objetivos deben establecerse en función de una información relevante sobre las necesidades, carencias, demandas o expectativas de los destinatarios (o de los grupos encargados de su custodia y formación). Esa información puede ser genérica –estudios psicológicos, pedagógicos, sociológicos, médicos... sobre la etapa educativa de que se trate- o específica, a través de estudios empíricos (análisis y evaluación de necesidades).
- b) En los supuestos de que las necesidades, demandas, carencias o expectativas sean muy amplias y complejas, y no se disponga de posibilidades de abordar todas ellas, las metas y objetivos del programa representan las prioridades establecidas por los responsables del programa, previa la aplicación a las mismas de criterios como los de relevancia o gravedad, rentabilidad, inmediatez, eficacia presumible o responsabilidad de quienes han de tomar las decisiones.
- c) Metas y objetivos se establecen como logros que, de alcanzarse, supondrían una respuesta satisfactoria a las necesidades, carencias, demandas o expectativas de los destinatarios del programa. O lo que es lo mismo: cuando los responsables del programa fijan sus metas y objetivos, están anticipando lo que debería ocurrir para dar la debida respuesta al elemento desencadenante del programa, para solucionar el problema que plantea la existencia de las necesidades y carencias de los destinatarios del programa.
- d) 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.
- e) En cuanto variable tratamiento o 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.



- f) Ha de incorporar un conjunto de medios y recursos que, además de ser educativos, deben ser suficientes, adecuados y eficaces para el logro de metas y objetivos.
- g) Necesita de un sistema capaz de apreciar tanto si metas y objetivos se logran como de dar cuenta, en casos de que así no sea, o de que lo sea a niveles insuficientes o insatisfactorios, de dónde se producen las disfunciones y carencias y, a ser posible, de las causas de las mismas (a título de ejemplo: objetivos muy elevados, aplicación inadecuada de los medios, tiempos insuficientes, momentos inadecuados, efectos no planeados...).

4.1.1. El sentido de la evaluación «en» y «del» programa

Este último aspecto es una de las manifestaciones de la realidad que nos ocupa: la evaluación. Y decimos una de las manifestaciones porque la otra se encuentra incorporada al punto c), dado que la evaluación es un medio más, junto al resto, al servicio de metas y objetivos.

Conviene señalar que la evaluación puede presentar otras manifestaciones -nos referimos a las funciones³³- como pueden ser las de control, la sociopolítica o la administrativa³⁴; sin embargo, en tales casos, aunque los programas sean educativos, su evaluación se sale del ámbito de lo educativo para entrar en el correspondiente al control social o político, a la rendición de cuentas...

Dado que toda evaluación necesita de criterios y de referencias para valorar la calidad y los resultados o logros de un programa, un elemento fundamental a tomar en consideración es el contexto en el que se aplica. En tal sentido, el Proyecto educativo y el centro educativo, cuando de programas ordinarios se trate, son dos elementos básicos que deben tenerse en cuenta para la valoración del programa que, obviamente, debe ser coherente y encontrarse en sintonía, en armonía, con ambos contextos. 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 eficacia³⁵. 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.

4.1.2. El hecho educativo y su evaluación

Comenzaré con una cuestión con el carácter de principio: en la medida en que lo educativo sea el elemento rector y directriz, es preciso un enfoque integral de la evaluación, lo que requiere la coherencia y la armonía de todo el planteamiento evaluativo. La evaluación del aprendizaje alcanzado por los destinatarios de los programas no puede concebirse al margen de la evaluación de la enseñanza diseñada y llevada a cabo por los agentes (generalmente, profesores), ni siquiera de la evaluación de los centros educativos y de los contextos en que los programas se llevan a cabo. Otra cuestión será que se realicen evaluaciones integrales; pero, desde luego, el planteamiento debe serlo y, por tanto, cada una de ellas, cuando se realiza por separado de las demás, al menos ha de tomarlas como referencia.

Ahora bien, no siempre hemos hecho de lo educativo, que es lo que nos distingue y la realidad a la que servimos, el elemento central a la hora de configurar metodologías. Uno de los casos más evidentes lo encontramos en el poco y tardío desarrollo de la evaluación criterial frente a la normativa; otro, en la facilidad con que se traspasan modelos procedentes de campos muy diferentes al ámbito educativo (caso del modelo de calidad de la EFQM) o de países con tradiciones muy distintas de la nuestra (normas del *Joint Committee*).

En tal sentido, entre los elementos que deberían ser objeto de consideración al plantear una metodología específica para la evaluación de programas educativos, proponemos los siguientes:

a) La naturaleza de programa.

Sus objetivos han de ser educativos -en modo alguno manipuladores- y adecuados a las características de sus destinatarios y de los ámbitos o contextos en que se llevará a cabo. En relación con este punto, es preciso dejar constancia de la diferencia entre objetivos del ámbito cognoscitivo/cognitivo y los propios del denominado dominio afectivo, donde los valores sólo pueden florecer en marcos educativos respetuosos con la autonomía personal y la libertad, tanto de iniciativa como de opción. Por otra parte, el programa, entendido como variable independiente o tratamiento, debe estar claramente especificado, adecuadamente justificado y sólidamente fundamentado en alguna concepción teórica.

b) Las posibles interacciones del programa a evaluar con los programas ordinarios.

En muchas ocasiones, los programas objeto de evaluación se desarrollan en contextos educativos y en simultáneo con otros programas. La modalidad en que aquellos se desarrollén, sea como nuevas formas de abordar el logro de metas, sea como programas integrados con los ordinarios, sea como elementos añadidos a la actividad ordinaria, puede condicionar fuertemente su desarrollo y, desde luego, puede tener repercusiones para su evaluación.

c) La naturaleza de las actuaciones en que toma cuerpo el programa.

Las actividades han de ser, además de adecuadas, motivadoras y eficaces para el logro de los objetivos, educativas y respetuosas con la autonomía de los educandos.

d) La importancia del factor humano para la eficacia de los programas.

Este aspecto toma cuerpo, entre otros, en los siguientes aspectos:

- El papel especial que juegan los destinatarios de los programas, a la vez agentes más o menos activos de su propia formación y condicionantes de la formación de los demás.
- La influencia de la implicación y el compromiso de los agentes naturales del programa, lo que representa la necesidad de crear actitudes favorables y expectativas razonables sobre su bondad, adecuación y utilidad. Junto a ello, su capacitación para una aplicación eficiente del programa.
- La importancia de las relaciones humanas para el éxito de los programas por configurar un clima favorecedor, neutral u obstaculizador para su desarrollo.
- El papel condicionante del contexto, por lo general orientado por ¿valores? contrarios o no compatibles con los propios de los programas educativos. Entre tales contextos es preciso señalar desde el propio del ámbito del programa hasta el amical y el familiar, al que hoy en día hay que añadir el configurado por los medios de comunicación e información (contexto virtual).

e) El papel de la evaluación.

Es un importante medio para el logro de las metas educativas, por lo que es preciso que todos los implicados se comprometan -cada uno en su ámbito de responsabilidades- en su realización, difusión y uso. La evaluación, que puede ser, y debe serlo, un medio eficaz para el logro de objetivos y metas, puede llegar a convertirse en un elemento distorsionador del programa y, en particular, de sus metas y objetivos.

4.2. Variables a considerar

El término programa, señalábamos más arriba, evoca un concepto de muy amplio espectro; ante tal realidad, es preciso tomar en consideración una serie de factores o variables a la hora de configurar una propuesta evaluativa. Veamos las principales.



4.2.1. Tamaño o amplitud del programa

Un concepto de esta naturaleza permite encontrarnos ante programas muy concretos, reducidos en sus metas y de corta duración, y ante otros de gran amplitud, complejidad, profundidad y permanencia. Los extremos del continuo podrían ir desde el programa de un curso intensivo de fin de semana, centrado en conocimientos o habilidades muy específicas, destinado a un alumnado reducido, hasta el caso de todo un sistema educativo una vez implantado en su totalidad, como puede ser el nacido de la reforma educativa de 1990.

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: por ejemplo, en el primer caso, no parece tener sentido una selección de «informantes»; por el contrario, en el segundo, el muestreo es una cuestión esencial en el ámbito de la metodología, al igual que resulta imprescindible contar con un amplio equipo de profesionales de diversos campos y, probablemente, con colaboradores debidamente adiestrados en la recogida de información.

Otro hecho importante es el relacionado con los aspectos o contenidos a evaluar: mientras en programas limitados en amplitud y duración es admisible pensar que el desarrollo del programa se acomodará en gran medida a lo previsto o planificado, en los de larga duración es imprescindible estar abiertos a efectos colaterales o no planeados, tanto positivos como negativos, cuya importancia no radica únicamente en lo que cada uno de ellos pueda representar sino en sus posibles influencias e interacciones con el propio programa (de ahí las referencias clásicas a los efectos no planeados –Sawin³⁶– o las más recientes a modelos como los de evaluación iluminativa, de Parlett y Hamilton, por ejemplo).

Del mismo modo, la evaluación de la eficacia de un programa exige planteamientos sistemáticos de selección de contenidos, representativos y con la debida ponderación, en el caso de programas de gran amplitud, mientras puede abarcar la realidad toda en el extremo contrario.

4.2.2. Complejidad del programa

Un programa educativo puede centrarse en el logro de metas ligadas a una única dimensión o variable independiente o tratamiento; sin embargo, es muy frecuente que, dada la complejidad misma de la educación, se diseñen programas que pretenden cambios de muy diversa naturaleza.

Un ejemplo lo tenemos en los programas educativos –sean de aula, de ciclo, de centro– ligados al sistema educativo actual derivado de la reforma de 1990, con contenidos conceptuales, procedimentales y actitudinales.

Con independencia de las implicaciones que cada uno de ellos por separado tenga para el programa, es preciso hacer notar las posibles relaciones e interacciones entre uno y otro tipo de actuaciones. En particular, conviene dejar constancia de lo que representa el doble papel de juez y parte que juegan los agentes del programa, en especial los profesores, en todo lo que tiene que ver con las actitudes y con los valores.

4.2.3. Estatus del evaluador

Dos son las situaciones básicas, cada una con sus matices, en las que se puede encontrar el evaluador:

a) La de ser simultáneamente responsable o agente del programa y de su evaluación. Nos encontramos ante una evaluación interna, esencialmente orientada como formativa o para la mejora; habitualmente, este tipo de evaluaciones se da en programas ya establecidos cuyos elementos esenciales no se cuestionan; las evaluaciones presentan una cierta continuidad, se realizan de forma poco formalizada y permiten la toma de decisiones sobre la marcha. Sin embargo, cabe realizar

una evaluación desde dentro de carácter no continuo (evaluaciones finales formativas).

b) La de ser sólo evaluador, sea por propia iniciativa, sea por encargo o tras cierto proceso abierto de selección. Generalmente nos encontramos ante 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 correspondientes sobre la continuidad o no del programa, tal como está o previa la introducción de los cambios convenientes. Resultan adecuadas, y esa es la situación habitual, en la aplicación primera de un programa de cierta magnitud, y en contextos políticos, sociales o académicos.

Junto a tales situaciones básicas, cabe reseñar la del evaluador externo, en funciones de asesor, que acompaña al desarrollo del programa, proporciona información, la discute, ofrece perspectivas para la interpretación...

Como puede apreciarse, también aquí nos movemos en un continuo que va del evaluador responsable del programa, que realiza una autoevaluación interna y formativa más o menos continuada, destinada a su mejora profesional y al perfeccionamiento del propio programa, al evaluador profesional, experto, externo al programa, que realiza una evaluación final y proporciona un informe para que los responsables del programa hagan lo que consideren oportuno. Si la primera situación se presta más a lo que podríamos denominar programas educativos ordinarios, la segunda se da generalmente con programas de intervención novedosos, que responden a necesidades identificadas o demandadas, y en los que es preciso disponer de información para decidir (por lo general, a partir de valoraciones derivadas de la aplicación de a una serie de criterios y, en ocasiones, de referencias).

En la medida en que uno y otro tipo de evaluaciones puedan tener una dimensión de control social, parece necesario hacer compatibles ambas opciones: la segunda, para ganar en credibilidad, la primera porque permite una mayor relevancia de los datos, además, y sobre todo, de hacer posible un mayor grado de implicación y compromiso con las decisiones de mejora.

Un problema especialmente grave es el relativo a la independencia del evaluador externo con relación a los responsables del programa; en ocasiones, tales responsables pueden desear el control de la propia evaluación y del informe, lo que puede alterar el propio proceso evaluativo y, lo que es más grave, atentar contra los derechos de los interesados en el programa y de sus resultados. Este es uno de los puntos cruciales a tener en cuenta en una deontología profesional.

4.2.4. Unidades de análisis

Los programas educativos tienen su destinatario natural en las personas de los educandos. Sin embargo, en ocasiones es necesario, bien sea intrínsecamente bien por interés de los responsables del programa o de su evaluación, tomar en consideración otro tipo de unidades.

El primer caso se puede ilustrar en las ocasiones en que el programa puede verse favorecido -o perjudicado- en su desarrollo y efectividad por las condiciones del contexto en que se aplica. No tomarlas en consideración puede conducir a atribuir los resultados negativos a la baja calidad del programa en lugar de a las condiciones adversas (oposición de personas relevantes, estructuras inadecuadas, pasividad de ciertos responsables, filosofía general incoherente...) en que se aplica.

El segundo caso se da en aquellas ocasiones en que es importante el contexto o marco en el que se aplica el programa a los efectos de posibles comparaciones -el aula, el centro, el barrio, la ciudad, la comarca, la comunidad autónoma, el país, la etnia, el género, los niveles educativos...- y de establecer el grado de generalización de los resultados.



4.2.5. La relación entre el programa y su evaluación

La evaluación forma parte natural de cualquier programa entendido en forma integral³⁷, al menos en su función «formativa», ya que es consustancial a la idea de mejora inherente a todo programa educativo; por otra parte, hasta la propia evaluación en funciones «sumativas» debería ser exigible, dado que no tiene demasiado sentido desarrollar una serie de actuaciones –que representan esfuerzo, medios y recursos– si de ellas no se siguen niveles de logro razonables.

Sin embargo, nada más lejos de la realidad. La cultura evaluativa no está todavía en vigor y, con demasiada frecuencia, se suceden ediciones y ediciones de un programa –sobre todo de los que hemos denominado «ordinarios»– sin una mínima reflexión sobre su calidad, su eficiencia o su eficacia.

Por eso no cabe extrañarse de que en programas que hemos denominado novedosos, o, mejor, nuevos, no siempre la evaluación tenga la presencia y entidad esperables. Así, nos encontramos con situaciones en que los «programas» se diseñan sin un estudio previo de su necesidad o de las demandas; en que la evaluación «se sube al carro» una vez puesto en marcha durante bastante tiempo, e, incluso, que se realice una vez el programa ha concluido en su implantación.

El hecho de que la evaluación esté o no integrada en el programa condiciona notablemente la metodología, limitando su alcance y reduciendo sus posibilidades en caso negativo. Sin embargo, hay que señalar que una evaluación integrada no tiene por qué afectar al carácter de interna/externa ni al estatus del evaluador, aspectos ambos que admiten las diversas opciones existentes.

4.2.6. La información

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. La misma elaboración de indicadores a los efectos de comparaciones entre diversos países se ve afectada por limitaciones tales que los llevan a centrarse en aspectos que no captan la riqueza y profundidad de las metas educativas de mayor calidad.

Con mucha frecuencia, por motivos fácilmente explicables, los programas se formulan en términos tales que no resulta fácil encontrar evidencias sobre su eficacia o efectividad; los planteamientos sugieren la necesidad de datos duros y finales pero, con frecuencia, los recogidos son de opinión e intermedios.

En cualquier caso, y siempre que nos encontremos ante una evaluación externa, el evaluador deberá afrontar aquí una tarea tan técnica como de especial importancia y dificultad: de importancia, porque es a partir de los datos recogidos que se decide sobre los resultados del programa y, si fueran inadecuados, del grado y naturaleza de las carencias e, incluso, de las posibles causas; dificultad, porque aquí se encuentra implicada toda la problemática de traducir constructos inobservables a manifestaciones más o menos directamente accesibles y apresables por parte del observador. A ello es preciso añadir el problema que puede originarse cuando estas tareas no fueron planteadas por los responsables del programa o, siéndolo, se den discrepancias entre ellos y el evaluador.

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 aquellos, 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 recoge-

das 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.

4.3. Componentes básicos de un modelo evaluativo de programas educativos

Dada la entidad del concepto modelo, tal vez sea preferible acudir a uno menos fuerte, como puede ser el de propuesta. En cualquier caso, resulta imprescindible considerar diversos componentes, a algunos de los cuales nos vamos a referir enseguida.

4.3.1. Componentes

Abordaremos este aspecto desde dos enfoques diferentes, pero complementarios: el correspondiente al programa y a su evaluación.

a) En relación con el objeto: el programa.

- El programa en cuanto tal y, por tanto, con todos sus elementos 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.

c) En relación con la evaluación del programa.

- La función o funciones a las que sirve, en cuánto condicionantes del resto de decisiones.
- La metodología a utilizar, concepto amplio en el que caben, fundamentalmente, los siguientes aspectos:
 - La calidad técnica del propio diseño de evaluación del programa.
 - El diseño o plan general de la evaluación: preguntas a las que responder, objetivos a lograr, niveles a alcanzar (significación estadística, significación pedagógica).
 - Los métodos. Dadas las consideraciones anteriores, optamos por el principio de complementariedad metodológica, bien en forma simultánea bien de modo sucesivo. A título de ejemplo, puede plantearse una investigación cuantitativo/cualitativa a la hora de evaluar el programa en cuanto tal, la observación –participante o no– del proceso de aplicación del programa, y la utilización de algún diseño experimental o cuasiexperimental para valorar la eficacia del mismo.
 - La información que debe ser recogida en torno al contexto, al programa, a su implantación, a su aplicación, a sus resultados y a sus consecuencias.
 - Los momentos en que habrá de ser recogida.
 - Las técnicas e instrumentos que habrán de utilizarse, aplicando el mismo principio de complementariedad metodológica.
 - El sistema de registro y los procedimientos de análisis.



- Los criterios para valorar cada una de las dimensiones del programa, y las referencias que permiten una apreciación global del mismo.
- El tipo de decisiones previsibles, los procedimientos para su toma, concreción, seguimiento y evaluación.
- El informe, su formato, elaboración, debate, difusión.

La concreción básica de todo lo anterior puede tomar la forma que a continuación se detalla.

4.3.2. Líneas generales de la propuesta evaluativa

La siguiente propuesta se estructura en torno a cuatro momentos o etapas, 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.

a) 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 la realidad toda del programa- y sobre todo, por sus enormes prestaciones para la mejora y optimización del programa.

- 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: científicos, técnicos, pedagógicos, metodológicos,...
 - 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, pertinencia³⁸ 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.

-
- 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).
- b) 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.
- c) Tercer momento: evaluación 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 su 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 → progreso³⁹.
 - Niveles preespecificados (evaluación criterial) → logro del dominio.
 - Niveles de otros programas → superioridad⁴⁰.
- Decisiones:
 - Sumativas: mantener o suprimir el programa.
 - Formativizadas: mejorar el programa para una nueva edición.

d) 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.

4.4. Consideraciones finales

De lo hasta aquí dicho, parece desprenderse, de alguna manera, la existencia de dos situaciones claramente diferenciadas, pero ambas de gran importancia para la calidad de la educación en lo relativo a la evaluación de programas educativos:

La evaluación interna de programas educativos ordinarios -autoevaluación- llevada a cabo por los responsables de los mismos (frecuentemente profesores), sea uno o un equipo, como una actividad más (generalmente docente) al servicio de la mejora de los destinatarios y de su actuación profesional.

La evaluación externa de programas educativos -heteroevaluación- realizada por profesionales competentes expertos, bien desde la vida académica universitaria bien desde la actividad evaluativa profesionalizada. En este caso, la vinculación contractual con los responsables de los programas introduce elementos peculiares de notable complejidad que deben ser objeto de análisis.

En el primero de los casos, probablemente nos encontremos ante uno de los aspectos con mayores potencialidades de mejora de la educación que se da en nuestro sistema educativo y al que menos atención se dedique y hasta en menos estima se tenga. Y es que, en efecto, un planteamiento sistemático de la evaluación de programas ordinarios puede dar lugar a un conocimiento técnico de gran relieve y utilidad tanto para la mejora de la actividad educativa como de la metodología de la evaluación de los programas educativos.

Para la evaluación de los programas ordinarios es preciso elaborar planteamientos metodológicos sencillos que puedan ser asumidos por los profesores; de

otra manera, no incorporarán la evaluación a sus comportamientos ordinarios perdiéndose de este modo el potencial de mejora que acompaña a la evaluación.

Probablemente, la propuesta más sencilla consista en lo siguiente:

a) Elaboración de una pauta:

La pauta, con el carácter de inventario o de cuestionario, debería poder ser autoaplicada con cierto detalle en la primera ocasión, abarcando los tres momentos: inicial (previo a la aplicación del programa), procesual y final⁴¹.

b) Aplicación de la pauta en los tres momentos:

En la medida en que los centros educativos estén organizados en departamentos o seminarios, la parte correspondiente al momento inicial sería conveniente que fuera cumplimentada por el conjunto de los profesores del mismo, de forma que se contara con más elementos de juicio, además de hacer posible la coherencia de planteamientos de curso a curso, ciclo a ciclo y etapa a etapa.

c) Toma de decisiones sobre el programa para su segunda edición.

d) Institucionalización del proceso:

En cursos sucesivos, la evaluación se centraría, en especial, en los puntos referentes a las decisiones de mejora tomadas. Metodológicamente, la evaluación podría tomar como momentos clave los relativos a las pruebas – formativas y sumativas- de carácter parcial, la aparición de hechos llamativos, positivos y negativos, que aconsejen una reflexión y, por supuesto, la conclusión del programa con su evaluación final.

En un esquema de tal naturaleza, no parece necesario acudir a otros estándares de calidad que los de carácter técnico y, en concreto, a los relativos a la calidad de los datos, esto es: la fiabilidad, la validez, la valía y el contraste.

Si de la evaluación de programas ordinarios pasamos a la investigación evaluativa, la situación, obviamente, es muy diferente. Es en este marco resulta de gran utilidad acudir a los estándares de calidad aprobados por el *Joint Committee*⁴² sobre cuya aplicación a nuestro país mantengo algunas reservas.

Dos hechos llaman la atención de inmediato: a) la concentración de los estándares en la evaluación, sin referencias normativas a los programas, y b) su justificación en un contexto diferente, el americano, lo que lleva a que algunos de los estándares sean de difícil aplicación en la situación ordinaria de nuestro país. Junto a ello, señalaré lo que considero ausencias y haré mención de ciertas discrepancias en cuanto a la organización de algunos estándares. Veamos:

a) Separación entre programa y su evaluación

Es comprensible que se planteen unos estándares sobre evaluación y no sobre los programas objeto de atención, por varias razones: estamos ante una actividad técnica, en la que puede darse más fácilmente el acuerdo; la evaluación es concebida como una actividad profesional y, por tanto, ajena a las personas responsables de los programas. Pero no es menos cierto que eso tiene dos importantes consecuencias: un cierto formalismo en las propuestas, por un lado, y posibles disfunciones por otro⁴³.

Entiendo que debería plantearse un bloque de estándares sobre programas y, en nuestro caso, sobre programas educativos. A modo de propuesta, podrían ser, al menos, los siguientes:

- Calidad intrínseca del programa:

El documento en que tome cuerpo el programa debe concretar sus metas y objetivos, dando cuenta o razones de su elección en función de la relevancia y pertinencia de las mismas; en la medida de lo posible, ofrecerá información sobre la base teórica que lo sustenta.



- **Calidad técnica del programa:**

El programa debe reunir las condiciones técnicas para ser considerado tal. Entre ellas, la de especificar los objetivos; concretar la naturaleza de la intervención, debidamente fundamentada; plantear los medios y recursos e incluir un sistema de comprobación y mejora de sus resultados.

- **Calidad metodológica del programa:**

El programa debe formularse de forma que contenga información relevante para resultar evaluable y, consecuentemente, mejorable.

b) Definición de estándares en un contexto diferente.

Algunos de los estándares -en concreto F3, P2, P6, P8- están pensados especialmente para evaluadores profesionales, contratados como expertos, situación poco común en nuestro país, donde la mayoría de las evaluaciones es cubierta por profesores universitarios con formación metodológica.

c) Propuesta de modificaciones.

Algún estándar puede considerarse innecesario si se aceptan los planteamientos sobre la conveniencia, y hasta la exigencia, de complementariedad metodológica; en concreto, los denominados A8 -Análisis de la información cuantitativa- y A9 -Análisis de la información cualitativa- bien podrían quedar refundidos en uno, con la denominación o rótulo de Análisis de la información, sin adjetivaciones.

El último de los estándares, el A12. Metaevaluación, no parecería ser propio de cualquier evaluación de programas, sino una exigencia técnica para la organización, sistematización y mejora de nuevas evaluaciones y de los propios programas educativos. La metaevaluación podría ser entendida como una exigencia técnica que permite a los diseñadores de programas y a los evaluadores, realizar mejor sus tareas de diseño y de planificación de la evaluación.

Por otra parte, se echa de menos una referencia genérica a la metodología ya que los estándares que aparecen sobre este ámbito se limitan a los procedimientos -A3- o se centran en exclusiva sobre la información (A: Fuentes; A5, 6 y 7: información válida, fiable y sistemática; A8 y A9, como ya se ha señalado, sobre su análisis).

El estándar sobre metodología, a mi juicio, debería referirse a la necesidad de que fuera respetuosa con el principio de complementariedad, al que hemos aludido reiteradamente, siempre que se trate de evaluaciones integrales.

Notas

¹ El presente artículo es deudor de otras publicaciones y aportaciones del autor recogidas en la bibliografía. La naturaleza de su contenido no justificaría una reelaboración total de tales documentos para escribir básicamente lo mismo aunque en forma diferente.

² Es importante dejar constancia de que ambos hilos conductores no se oponen a otras concepciones evaluativas de diversa índole, derivadas de funciones de la evaluación de muy diferente naturaleza. Pero, en cualquier caso, los planteamientos aquí defendidos resultan de especial relieve si de lo que se trata es de una evaluación pedagógica.

³ Cuestión diferente es la de determinar sobre qué aspectos el juicio de los diferentes «sujetos» es el pertinente y adecuado y para que funciones resulta conveniente su utilización.

⁴ Y hasta religiosos, en el caso de las personas creyentes que asisten a centros confesionales.

-
- ⁵ Prueba que se realizaba tras el antiguo Bachillerato antes de poder entrar en la Universidad.
- ⁶ Taxonomía significa «ordenación», «clasificación». A Bloom y colaboradores se debe la más conocida de las taxonomías de los objetivos de la educación, organizada en torno a los cognoscitivos, los afectivos y los psicomotores. Los primeros constituyen una escala de seis amplias categorías, establecidas en forma jerárquica, desde el nivel más bajo (conocimiento) al más elevado (valoración), pasando por los niveles intermedios: comprensión, aplicación, análisis y síntesis.
- ⁷ La más importante de ellas, la conocida como investigación evaluativa. Junto a ella, el control por parte de la Administración o de los responsables de los centros y programas.
- ⁸ Es interesante subrayar el hecho de ayudar, acompañar, estimular..., porque nadie puede sustituir a otro en ese proceso. Nadie puede aprender por otro o realizar la formación de la mente de otro: aprendizaje, formación y educación son tareas propias de cada uno, si bien la competencia en el ejercicio de sus tareas por parte del profesor y del educador pueden resultar decisivas.
- ⁹ Puede acudir a las reflexiones de R. SANZ ORO (1990): *Evaluación de programas de Orientación*, Madrid: Pirámide, o a la aportación de GROSS y HUMPHREYS (1985): *Evaluating Intervention Programs*, Columbia University, New York: Teachers College Press.
- ¹⁰ NEVO, D.: «The conceptualization of educational evaluation: An analytical review of literature». En E.R. HOUSE (1986): *New directions in educational evaluation*, Lewes R.U.: The Falmer Press; 15-30.
- ¹¹ Así lo manifiestan, con datos y entusiasmo, Bloom y sus colaboradores. BLOOM, B.S. y cols. (1981): *Evaluación del aprendizaje*. Buenos Aires: Troquel; 90 y 198.
- ¹² Piense el lector que la falta de coherencia suele traducirse en la predominancia de la evaluación sobre los propios objetivos, lo que, además de no tener sentido, pervierte el sistema.
- ¹³ María Moliner, en su *Diccionario de uso del español* (1979). Madrid: Gredos, afirma: Justipreciar, tasar, valorar. Atribuir cierto valor a una cosa. APRECIAR, CALCULAR, ESTIMAR... JUZGAR.
- ¹⁴ Venimos defendiendo el concepto de *valía* de la información desde los primeros años 80, refiriéndolo a la necesidad de que la información sea matizada, rica, específica... algo imprescindible a la hora de la *toma de decisiones*. Vid. PÉREZ JUSTE, R. y otros (1991): *La Universidad Nacional de Educación a Distancia. Aproximación a la evaluación de un modelo innovador*. Madrid: CIDE, p.64
- ¹⁵ Además de los exámenes cabe acudir a los trabajos, a los debates en el aula, a la realización de diversas actividades, en presencia o no del profesor... Y cuando se trate de exámenes, no sólo en forma de ensayo sino acudiendo a conceptos, esquemas, pruebas objetivas, interpretación de gráficos, reconocimiento de imágenes... La variedad es muy elevada, y cada una de las modalidades resulta especialmente adecuada para algún tipo de objetivo.
- ¹⁶ No se debe despreciar la memorización de ciertos datos, bien por su importancia intrínseca bien por su utilidad para otro tipo de aprendizaje. Tampoco debe dejarse de lado la aplicación de ciertos principios, teoremas, leyes... y, desde luego, hay que cultivar las capacidades más elevadas del ser humano, ligadas a la comprensión profunda, la posibilidad de organizar de forma personal el saber, el razonamiento abstracto, el compromiso con ciertos valores socialmente reconocidos...
- ¹⁷ De hecho, conviene elaborar una tabla de especificaciones de la enseñanza y, en paralelo, la correspondiente a la evaluación. No hay mejor medio de asegurar la coherencia entre aquella y esta, de forma que si el programa es ya de calidad, la evaluación participará de la misma. En esa línea se manifiesta Hills (p. 306):
- ¹⁸ Dos son las dificultades técnicas fundamentales a superar: la primera se refiere a seleccionar un grupo -muestra- que sea plenamente representativo del grupo amplio o población. Esta dificultad es relativamente fácil de superar me-



diante las técnicas de muestreo. La segunda, más compleja, se refiere al mismo contenido de la prueba con la que se construirá el baremo, y resulta más difícil en estos momentos en los que, como consecuencia de ciertas características de nuestro tiempo, se tiende a evitar los currículos centralizados y uniformes y se prefiere adaptarlos a los contextos y a los centros. ¿Cómo construir una prueba común para jóvenes de 14 años si lo que se les enseña en un concreto territorio (país, comunidad autónoma) presenta una gran diversidad?

- ¹⁹ La denominada evaluación *criterial* se debe a Robert Glaser, y es especialmente adecuada para valorar los resultados de los procesos educativos. Además de la obra de Popham, se le recomienda el capítulo de J. Jornet en la obra de V. García Hoz, *Problemas y métodos de investigación en Educación Personalizada*; 419-443.
- ²⁰ El lector puede encontrar referencia a investigaciones en este sentido en BONBOIR, A. (1974): *La docimología. Problemática de la evaluación*. Madrid: Morata, segunda parte, cap. primero.
- ²¹ A diferencia de «los criterios» a los que ya nos hemos referido, denominamos aquí «criterio» al conjunto de conocimientos, destrezas, habilidades... que debe dominar el alumno para pasar a una unidad, curso, ciclo... posterior con posibilidades de éxito.
- ²² El profesor Yela afirmaba que el alumno rinde según «puede», según «sabe» y según «quiere». En el primer verbo incluye las aptitudes y capacidades, además del nivel de partida; en el segundo, la motivación y los intereses, en el tercero, las técnicas de estudio.
- ²³ El citado profesor Yela resumía las necesidades básicas en «valer algo, valer para algo y valer para alguien». El reconocimiento del esfuerzo –aunque no haya sido suficiente para alcanzar el nivel objetivo del éxito– es una forma de que el alumno perciba que «vale para alguien», un alguien tan importante como los padres en educación infantil o primaria, o como el profesor, en cualquier etapa educativa.
- ²⁴ Un matiz: en la enseñanza post obligatoria, predomina la referencia criterial, junto a la normativa en situaciones de selección. En una etapa obligatoria, donde, por principio, están todas las personas, con unas diferencias enormes en intereses, motivación, base, técnicas y capacidad, debería primar el «progreso», manifestación de la referencia personalizada.
- ²⁵ Convendría señalar que hay evidencias empíricas de que el mero hecho de informar a los alumnos sobre los aspectos en que no van bien, contribuyen a la mejora posterior del rendimiento. Vid. al respecto Page, E.B. (1958): «A Seventy-four classroom experiment in school motivation». En *Journal of Educational Psychology*, 49, pp. 173-181.
- ²⁶ No se olvide que los objetivos, generalmente referidos a capacidades, destrezas, estrategias... son, de alguna manera, transversales a las áreas de saber, a los conocimientos. A observar, a razonar, a clasificar, a argumentar, a demostrar... se puede –y se debe– aprender mediante el estudio, el trabajo, las actividades de las diferentes materias. Al respecto son de gran interés los trabajos del profeso García Hoz sobre el vocabulario. Vid. GARCÍA HOZ, V. (1977): *Estudios experimentales sobre el vocabulario*. Madrid: Instituto «San José de Calasanz» del CSIC.
- ²⁷ El lector podrá encontrar una sencilla y asequible propuesta en Pérez Juste, R. y otros (1994): *Centros educativos de calidad*. Madrid: CECE.
- ²⁸ Podrá encontrar una pauta para la evaluación de programas en los capítulos redactados por el autor en la obra de Villar y Medina (1998): *Evaluación de Programas educativos, centros y profesores*. Madrid: Universitas.
- ²⁹ Las ideas del autor en torno a este punto tan amplio se recogen en su obra *Diagnóstico, evaluación y toma de decisiones*, volumen 9 del *Tratado de Educación Personalizada*, editado en Madrid por Rialp en 1989.

- ³⁰ Bajo esta denominación queremos incluir cualquier tipo de objetivo educativo, desde los tradicionales conocimientos a la actual idea de procedimientos -habilidades, destrezas, estrategias- hasta llegar a los comportamientos de acuerdo con valores socialmente reconocidos.
- ³¹ E, incluso, como veremos más adelante, de la teoría pedagógica, si se acumularan ordenada y sistemáticamente los resultados de este tipo de evaluaciones.
- ³² 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: Univérsitas, 73-106.
- ³³ Vid. al respecto la aportación de D. Nevo: «The conceptualization of Educational Evaluation», en E.R. HOUSE (1986): *New Directions in Educational Evaluation*: Falmer Press.
- ³⁴ Otros autores presentan funciones diferentes. Vid. FERNÁNDEZ-BALLESTEROS, R. (1996): *Evaluación de Programas. Una guía práctica en ámbitos sociales, educativos y de salud*, Madrid: Síntesis; 30-32.
- ³⁵ En ocasiones, un buen programa puede no funcionar adecuadamente por falta de adecuación al contexto de referencia.
- ³⁶ Ya hace mucho tiempo se hacía mención de este tipo de efectos. Vid. SAWIN, E.I. (1969): *Técnicas básicas de evaluación*, Madrid: Magisterio Español, cap. 2.
- ³⁷ De todos es conocida la ya añeja cita de P. de Lafourcade (*Evaluación de los aprendizajes*, Buenos Aires: Kapelusz, 1969, p. 16): «sin indicación de objetivos, el proceso sería un barco a la deriva; sin un buen complejo metodológico, una acción insegura y azarosa; sin aprendizaje, un esfuerzo inútil; sin evaluación, una empresa de la cual se desconocería su eficiencia, pese a que la tuviera; y sin reajuste, una tarea a medias».
- ³⁸ 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*; 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».
- ³⁹ Estaríamos ante una referencia a la que antes hemos denominado «personalizada», al referirnos al alumnado, o «idiosincrática», si nuestro objeto son los centros o los programas educativos.
- ⁴⁰ La referencia en este caso se inscribe en el marco de la denominada previamente como «normativa».
- ⁴¹ El lector puede encontrar una propuesta del autor en la citada obra de Villar y Medina.
- ⁴² JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (1994). *The Program Evaluation Standards*. 2nd edition. Thousand Oaks, CA. Sage Pub.
- ⁴³ Un caso patente es el relativo a la «*evaluación de la calidad de las universidades*» sin previamente plantear algún tipo de orientación, pauta o propuesta sobre lo que es calidad, que queda al arbitrio de la «autonomía universitaria». Una de los, o ambos conceptos van unidos o no será posible establecer pautas comunes comparables... a no ser que consideramos que la calidad se encuentra subsumida de modo más o menos subrepticio en los instrumentos elaborados para la evaluación.

Referencias

Resulta muy difícil, por la enorme cantidad de obras al respecto, reflejar la bibliografía disponible. Como consecuencia, se reseña a continuación una propuesta de textos básicos, donde el lector podrá encontrar una amplia bibliografía específica. Se marcan con * obras especialmente recomendadas.



-
- ALVIRA MARTÓN, F. (1991): *Metodología de la evaluación de programas*. Madrid, C.I.S.
- BLOOM, B.S. y OTROS (1981): *Elaboración del aprendizaje*. Buenos Aires, Troguer.
- BUNGE, M. (1979): *La investigación científica*. Barcelona, Ariel.
- 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.
- * FERNÁNDEZ BALLESTEROS, R. (Edit.) (1996): *Evaluación de programas. Una guía práctica en ámbitos sociales, educativos y de salud*. Madrid, Síntesis.
- DE LA ORDEN HOZ, A. (1997): «Evaluación y optimización educativa», en SALMERÓN, H. y otros (Edits.): *Evaluación educativa*. Granada, Grupo Editorial Universitario.
- GUBA, E. y LINCOLN, Y. (1981): *Effective evaluation*. San Francisco, Jossey-Bass.
- GUBA, E. y LINCOLN, Y. (1989): *Fourth generation evaluation*. Newbury Park, CA. Sage.
- HOUSE, E.R. (Edit.): *New Directions in Educational Evaluation*. Lewes, Falmer Press.
- JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (1981): *Standards for Evaluations of Educational Programs, Projeets and Materials*. New York, McGraw-Hill Sage.
- * JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (1994): *The Program Evaluation Standards*. Thousand Oaks, CA. Sage.
- MINISTERIO DE EDUCACIÓN Y CULTURA (1997): *Modelo europeo de Gestión Calidad*. Madrid, Servicio de Publicaciones, Secretaría General Técnica.
- PÉREZ JUSTE, R. Y GARCÍA RAMOS, J.M. (1989): *Diagnóstico, evaluación y toma de decisiones*, Vol. 8 del *Tratado de Educación Personalizada*. Madrid, Rialp.
- PÉREZ JUSTE, R. y OTROS (1991): *La Universidad Nacional de Educación a Distancia. Aproximación a la evaluación de un modelo innovador*. Madrid, CIDE.
- PÉREZ JUSTE, R. MARTÍNEZ MEDIANO, C. y RODRÍGUEZ GÓMEZ, G. (1994): *Centros educativos de Calidad*. Madrid, CECE.
- * PÉREZ JUSTE, R. (1998): Evaluación de Programas educativos, en VILLAR, L.M. y MEDINA, A. (1998): *Evaluación de programas educativos, centros y profesores*. Madrid, Universitas; 73-106.
- PÉREZ JUSTE, R. (1999): «Nuevas perspectivas de gestión, control y evaluación», en *Bordón*, 51 (4); 449-478.
- PÉREZ JUSTE, R. (2000): «La evaluación como medio para la mejora de la eficacia y de la calidad del aprendizaje, de la educación y de las instituciones», en VARIOS: *Hacia el Tercer Milenio: Cambio educativo y Educación para el Cambio. Actas de XII Congreso Nacional y I Iberoamericano de Pedagogía*. Tomo I. Ponencias. Madrid, Sociedad Española de Pedagogía; 535-573.
- PÉREZ JUSTE, R. (2000): «La calidad de la educación», en PÉREZ JUSTE, R. y OTROS: *Hacia una Educación de Calidad. Gestión, instrumentos y evaluación*. Madrid, Narcea; 13-44.
- * PÉREZ JUSTE, R (Edit.) (2000): «Evaluación de programas educativos», en Monográfico de *Revista de Investigación Educativa*, 18, 2.
- POPHAM, W.J. (1983): *Evaluación basada en criterios*. Madrid, Magisterio Español.
- ROSSI, P. y FREEMAN, H. (1989): *Evaluación. Un enfoque sistémico para programas sociales*. México, Trillas.
- * ROSSI, P. FREEMAN, H. y LIPSEY, M. (1999): *Evaluation. A systematic approach*. Thousand Oaks, Sage.
- * STUFFLEBEAM, D.L. y SHINKFIELD, A.J. (1987): *Evaluación sistemática. Guía teórica y práctica*. Madrid, Paidós-Ministerio de Educación y Ciencia.
-

- *STAKE, R. y MABRY, L. (1997): *Advances in program evaluation*. San Francisco, Jossey-Bass Publishers.
- VARIOS (1992): Monográfico de *Bordón* sobre «Evaluación de programas». Vol. 43, 4.
- WILSON, J.D. (1992): *Cómo valorar la calidad de la enseñanza*. Barcelona, Paidós-Ministerio de Educación y Ciencia.

Ramón Pérez Juste
es Catedrático de *Métodos de Investigación en Educación y Diagnóstico*
de la *Universidad Nacional de Educación a Distancia*.
Correo electrónico: rperez@edu.uned.es