

UNED

Evaluación de centros y profesores

Eustaquio Martín Rodríguez *(Coordinador)*



Evaluación de centros y profesores

EUSTAQUIO MARTÍN RODRÍGUEZ (Coordinador)

JOSÉ FELIX ÁNGULO RASCO

JUAN FERNÁNDEZ SIERRA

SUSANA FERNÁNDEZ LARRAGUETA

UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA

EVALUACIÓN DE CENTROS Y PROFESORES

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del Copyright, bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamos públicos.

© *Universidad Nacional de Educación a Distancia
Madrid, 2015*

www.uned.es/publicaciones

© *Eustaquio Martín Rodríguez (Coordinador), José Felix Ángulo Rasco,
Juan Fernández Sierra, Susana Fernández Larragueta*

ISBN electrónico: 978-84-362-6817-1

Edición digital actualizada: junio de 2015

ÍNDICE

Presentación

Tema 1. LA EVALUACIÓN DE CENTROS Y PROFESORES. GÉNESIS DE LA CONFORMACIÓN DE UN NUEVO CAMPO PROFESIONAL, DE ESTUDIO E INVESTIGACIÓN (*Eustaquio Martín Rodríguez*)

Introducción

1. La diversificación de los planteamientos y ámbitos de la evaluación
 - 1.1. *Características más relevantes del contexto socioeducativo en que emerge la evaluación de centros y profesores*
 - 1.2. *Debates y deliberaciones sobre la evaluación*
2. El surgimiento y desarrollo de la evaluación de centros y profesores
 - 2.1. *La emergencia de la evaluación de centros y profesores*
 - 2.2. *El desarrollo y consolidación de la evaluación de centros y profesores*
 - 2.3. *Generalización e internacionalización de la evaluación de centros y profesores*
3. A modo de recapitulación

Bibliografía

Tema 2. CONCEPCIONES, VÍNCULOS, SIGNIFICADOS Y OTRAS CUESTIONES ESTRUCTURANTES DE LA EVALUACIÓN DE CENTROS Y PROFESORES (*José Félix Ángulo Rasco y Eustaquio Martín Rodríguez*)

Introducción

1. En torno al concepto de evaluación
 - 1.1. *«Evaluation» versus «assessment»*
 - 1.2. *Algunas definiciones de evaluación*

2. Acerca de la relación entre evaluación, investigación e investigación evaluativa
3. La evaluación en las distintas teorías y tradiciones científicas
 - 3.1. *Tradición «científico-racional»*
 - 3.2. *La tradición Interpretativo-Simbólica*
 - 3.3. *La tradición socio-crítica*
 - 3.4. *Posiciones postparadigmáticas*
4. Los indicadores de evaluación
5. La política y ética de la evaluación
6. A modo de recapitulación

Bibliografía

Tema 3. EL CENTRO EDUCATIVO COMO OBJETO DE EVALUACIÓN (*Eustaquio Martín Rodríguez*)

Introducción

1. Perspectivas de la evaluación de centros educativos
 - 1.1. *La evaluación de centros desde la perspectiva de la rendición de cuentas, eficacia y calidad*
 - 1.2. *La evaluación de centros desde la perspectiva de los procesos de mejora y desarrollo institucional*
2. Estrategias para la evaluación de instituciones educativas
3. Las modalidades de evaluación
4. La especificidad de los distintos ámbitos de evaluación
5. Metodología de evaluación
6. Criterios e indicadores de evaluación
7. A modo de conclusiones

Bibliografía

Tema 4. LA EVALUACIÓN DE CENTROS PROMOVIDA POR LAS ADMINISTRACIONES EDUCATIVAS (*Eustaquio Martín Rodríguez*)

Introducción

1. Evaluación y Administración
2. La evaluación externa de centros desde las iniciativas de las administraciones

- 2.1. *La evaluación de productos o rendimientos de los alumnos en el contexto de la evaluación de centros educativos*
- 2.2. *La evaluación de procesos y productos o evaluación de centros escolares*
- 2.3. *Propuestas y realizaciones españolas*
- 3. Estrategias de evaluación interna y combinada de evaluaciones interna y externa
- 4. A modo de conclusiones

Bibliografía

Tema 5. LA EVALUACIÓN BASADA EN EL CENTRO (*Eustaquio Martín Rodríguez*)

Introducción

- 1. Formas y tipos de autoevaluación institucional
- 2. Orientaciones y estrategias de autoevaluación
- 3. Propuestas y experiencias de autoevaluación
 - 3.1. *Análisis de algunas experiencias de autoevaluación promovidas por investigadores y proyectos internacionales*
- 4. Presupuestos y condiciones para la realización de la autoevaluación
- 5. Balance de las experiencias y propuestas de autoevaluación institucional

Bibliografía

Tema 6. LA EVALUACIÓN DE LAS INSTITUCIONES UNIVERSITARIAS (*Juan Fernández Sierra y Susana Fernández Larragueta*)

Introducción

- 1. El papel de los actores en la evaluación educativa
- 2. La Evaluación Institucional Universitaria en el ámbito europeo
 - 2.1. *La evaluación institucional en Gran Bretaña: The Quality Assurance Agency for Higher Education*
 - 2.2. *La evaluación institucional en Francia: Agence d'évaluation de la recherche et de l'enseignement supérieur*

2.3. *La evaluación institucional en Italia: Agencia Nazionale di Valutazione del Sistema Educativo e de lla Ricercay Nucleo di Valutazione*

3. Los Programas de Evaluación Institucional Universitarios Españoles

Bibliografía

Tema 7. EL PROFESORADO COMO OBJETO DE EVALUACIÓN (*Eustaquio Martín Rodríguez*)

Introducción

1. Acerca de la evaluación del profesorado .
2. Tendencias y características de la evaluación del profesorado
3. Los propósitos de la educación del profesorado
4. Perspectivas de evaluación docente
5. A modo de conclusiones

Bibliografía

Tema 8. LA EVALUACIÓN DEL DESEMPEÑO DOCENTE (*Eustaquio Martín Rodríguez*)

Introducción

1. Cuestiones en torno a la evaluación del desempeño docente .
2. Los contenidos de la evaluación del profesorado
3. La evaluación de la competencia y desempeño del profesor
4. La evaluación de los docentes a través de los rendimientos del alumnado
5. Los procedimientos de la evaluación de profesores .
6. A modo de conclusión

Bibliografía

Tema 9. LA EVALUACIÓN DOCENTE ENTRE IGUALES (*Eustaquio Martín Rodríguez*)

Introducción

1. La autoevaluación individual de los profesores
2. El profesorado y las diversas posibilidades de la autoevaluación docente
3. La autoevaluación docente y la investigación-acción
4. El uso del portafolio para la autoevaluación del profesorado
5. A modo de conclusión

Bibliografía

**Tema 10. LA EVALUACIÓN DE LOS DOCENTES POR LOS ALUMNOS.
EXPERIENCIAS CON EL PROFESORADO UNIVERSITARIO** (*Juan
Fernández Sierra y Susana Fernández Larragueta*)

Introducción

1. Las contradicciones de la evaluación de los docentes
2. La iniciativa y el control de los procesos evaluativos
3. Los cuestionarios de opinión, modelo hegemónico de evaluación del profesorado universitario
4. Una experiencia alternativa de evaluación de los procesos de enseñanza y aprendizaje en la universidad
 - 4.1. *Análisis de la experiencia* .

Bibliografía

PRESENTACIÓN

Uno de los objetivos prioritarios de la educación que aparece recurrentemente en reuniones, foros y discursos políticos de las últimas décadas es, sin duda, el de la calidad. Parece evidente que existe un cierto acuerdo en que una vez resuelto el problema del acceso universal de los alumnos, en edad escolar, al sistema educativo, lo urgente, lo que necesariamente tiene que abordarse sin dilaciones, es la mejora de la calidad de la educación.

Siendo éste un planteamiento loable y sobre el que existe, al menos en sus propósitos, un cierto consenso entre todas las fuerzas sociales y políticas, las discrepancias, controversias y desencuentros surgen tanto cuando se trata de definir qué es y en qué consiste la calidad, como cuando se planifican y concretan las medidas y procedimientos para acometer la optimización de los sistemas educativos. Desde cualquiera de las opciones posibles en estos casos, es innegable que la teoría y práctica de la evaluación institucional y del profesorado ha ido adquiriendo un protagonismo estelar ya sea como base sobre la que sustentar y legitimar las directrices de las políticas educativas que a nivel nacional se ponen en juego; para fundamentar las decisiones que necesitan tomarse en la conducción de los sistemas educativos; o para proveer a los centros y profesores de herramientas y datos que, además de facilitarles el conocimiento de sus quehaceres, les permitan profundizar en sus desarrollos institucional y profesional.

Teniendo en cuenta dichas perspectivas, esta publicación aborda el desarrollo temático de la teoría y práctica de la Evaluación de Centros y Profesores al objeto de dar respuestas a las necesidades y demandas formativas propias de los estudiantes de la titulación de Pedagogía, así como de los profesionales de la educación preocupados por seguir formándose en estos campos de estudio. A través de los diez temas que la integran, se presentan, en una primera parte, un marco de fundamentación en el que se recogen la conformación histórica y las bases teóricas y conceptos más sustantivos de la evaluación; en la segunda parte se acomete el estudio de la evaluación de

centros y, en la tercera parte, se pasa a tratar monográficamente la evaluación del profesorado.

En los dos primeros temas, que constituyen lo que hemos denominado como primera parte, se comienza con un estudio histórico de la evaluación en donde se rastrean sus orígenes e identifican las distintas etapas que conducen a su conformación como campo de estudio e investigación. Nos hemos fijado especialmente, al contrario de lo que hacen la mayoría de las publicaciones sobre este tema, tanto en las aportaciones generales que, en los primeros tiempos, cabe atribuir a la evaluación de proyectos y programas, como en las contribuciones particulares que a partir de los años 70 empiezan a surgir en la evaluación institucional y del profesorado. La consolidación disciplinar de la evaluación es fruto del rico patrimonio que se fue acumulando en unas y otras experiencias, aunque sin duda el legado de los estudios sobre proyectos y programas facilitó en gran medida la aparición y desarrollo de los nuevos ámbitos de evaluación.

Las bases conceptuales de la evaluación centran la exposición temática del segundo tema. Por un lado, se han analizado el concepto y los distintos significados de la evaluación educativa, categorizando las diferentes definiciones con las que trata de acotarse la misma, así como problematizando sus concepciones y significados. A este respecto, nos hemos detenido en la consideración de la evaluación como un proceso de diálogo, de enseñanza y aprendizaje social; en la especial vinculación entre la evaluación y la especificidad de su objeto como un modo de cuestionar y alterar la tradición ateorica con la que se plantea en no pocos estudios; en las, ya clásicas, dimensiones formativa y sumativa de la evaluación; en las modalidades de evaluación externa e interna; en las controversias entre evaluación cuantitativa y cualitativa; y en los aspectos político y ético de la evaluación.

Por otro lado, se ha tratado de clarificar la relación entre evaluación, investigación e investigación evaluativa discriminando las similitudes y diferencias entre las mismas y haciéndonos eco de las posiciones que surgen en torno a si la evaluación puede considerarse como una forma o modalidad de investigación o si, en cambio, ésta es una herramienta que la evaluación con su propia identidad, distinta a la investigación, utiliza en sus estudios.

Con el tema 3 nos adentramos en el estudio de la evaluación de las instituciones educativas. En él se resaltan los factores que han hecho, primero,

emerger y, después, que se desarrolle de una manera acelerada la evaluación institucional. La evaluación de centros se presenta como un campo complejo susceptible de no pocos dilemas desde las visiones y experiencias con las que nos acerquemos al mismo. Se contraponen, en este sentido, los propósitos de la evaluación tanto desde la perspectiva de la rendición de cuentas, eficacia y calidad, como desde la que se sustenta en favorecer los procesos de mejora y desarrollo institucional.

Las estrategias para la evaluación de instituciones educativas son tomadas como el elemento organizador que nos permite ir estudiando, en el tema 4, los distintos enfoques y propuestas. Hemos optado, a falta de otro criterio más inclusivo y que diera una cierta coherencia a nuestro discurso, ir analizando las estrategias externas de evaluación, en las que distinguimos las que son iniciativa de las administraciones de aquellas que son propuestas desde instancias independientes, las estrategias internas y/o de autoevaluación y las estrategias mixtas en las que se combinan modalidades externas e internas. En cada uno de los casos, hemos tratado de desentrañar los aspectos más relevantes de las propuestas como la consideración y concreción del objeto de evaluación, las perspectivas ontológica, epistemológica y metodológica adoptadas, los propósitos de la evaluación, los procedimientos y la participación de las audiencias, etc.

La evaluación basada en el centro o autoevaluación institucional se acomete en el tema 5. Como modalidad de evaluación interna, la autoevaluación está vinculada especialmente a la orientación de mejora y desarrollo profesional, aunque en no pocas ocasiones se ha utilizado y se sigue utilizando como un procedimiento de rendición de cuentas, *accountability*. Dada la diversidad de usos con propósitos diferentes se ha tratado de clarificar el surgimiento, la tipología y las distintas tradiciones académicas que proponen y defienden su incorporación a los quehaceres propiamente institucionales. Sin duda, estamos ante una modalidad de evaluación que debería ser la base de la supervisión interna de los centros, la plataforma de los procesos de innovación y mejora, y la herramienta con la que dar rigor y sustento a los proyectos y experiencias de comunidades autocríticas.

La evaluación de instituciones universitarias protagoniza el papel estelar del tema 6. Normalmente las publicaciones dedicadas a la evaluación de centros se ocupan casi exclusivamente de los centros educativos de educación primaria y secundaria, dejando de lado la Universidad como institución edu-

cativa. En nuestro caso, creemos necesario incorporar este tema para disponer de una visión completa de las instituciones formales del sistema educativo. Con la finalidad de ofrecer una panorámica de esta cuestión y dar pie al debate y reflexión, se ha comenzado describiendo tres ejemplos de procesos institucionales de evaluación llevados a cabo en tres países de la Comunidad Europea —Gran Bretaña, Francia e Italia—; para pasar después a analizar la evaluación institucional de la Universidad en nuestro país.

El tema 7, dedicado a la evaluación del profesorado, es la puerta de entrada al estudio de la tercera parte. Este es un ámbito de evaluación que, a diferencia de la evaluación institucional, está mucho menos desarrollado por las resistencias que los docentes han generado siempre a cualquier forma de control de su trabajo. Sin embargo, durante los años ochenta y noventa ha ido adquirido un gran impulso como consecuencia de los movimientos en favor de la profesionalidad de los docentes y de la reestructuración de la escuela.

Como en el tema dedicado a la evaluación de centros, analizamos en primer lugar los diferentes propósitos desde los que se puede abordar la evaluación del profesorado, para pasar, a continuación, a hacer explícitas las vinculaciones de la evaluación con las distintas concepciones de la enseñanza y la profesión docente. Tomando estos elementos como marco de referencia se plantean los distintos enfoques y propuestas de evaluación, considerando las experiencias que se están llevando a cabo por las administraciones educativas e identificando las demás modalidades de evaluación.

La evaluación del desempeño docente ocupa por completo el desarrollo del tema 8. En este tema, seguimos profundizando en el estudio de la evaluación del profesorado analizando, en primer lugar, algunas de las cuestiones relacionadas con la concepción y características de la evaluación del desempeño docente. Al objeto de clarificar cuáles son las fuentes de inspiración de los estándares docentes, analizamos a continuación los contenidos de evaluación, es decir, las distintas dimensiones en que habría que centrar la recogida de datos e información. Presentamos después la Supervisión Clínica como una modalidad de evaluación de desempeño que puede dar respuestas al control y al desarrollo profesional del profesorado. Recogemos, asimismo, las limitaciones y dilemas que implica el uso del rendimiento de los alumnos como criterio único de evaluación del profesorado, enfoque que ha recobrado su vigencia a raíz del desarrollo de las metodologías estadísticas que

hacen posible los sistema de medición del valor añadido Y, para terminar, nos hacemos eco de la necesidad de que la evaluación docente cuente con un marco normativo, ético y de procedimientos para acometer la misma.

En el tema 9 se desarrolla le evaluación entre iguales. Bajo el rótulo de autoevaluación del profesorado se pueden encontrar, sin embargo, en la literatura pedagógica enfoques, experiencias y procedimientos muy diferentes entre sí. Es verdad, que todos ellos defienden la autoevaluación como una propuesta de evaluación formativa dado que sus propósitos siempre se encaminan a potenciar el crecimiento y la mejora profesional del profesorado. Pero no es menos cierto, también, que los supuestos y concepciones de la profesión docente y la enseñanza que dejan traslucir se identifican, en ocasiones, con las competencias y destrezas que debe poseer un profesional técnico en función de los objetivos extrínsecos que ha de perseguir la enseñanza; en otros casos, se proyectan en el compromiso y responsabilidad individual de los profesores en el marco concreto de su actividad en el aula; y en algunas ocasiones, las menos, los docentes son considerados como prácticos reflexivos que investigan sobre su enseñanza ya que la enseñanza en sí es un proceso de investigación-acción.

Por último, en el tema 10 se aborda la evaluación del profesorado por los alumnos, centrando el desarrollo de este tema en las experiencias llevadas a cabo con el profesorado universitario. En este sentido, y como se afirma en la introducción al tema, de los elementos que intervienen en los procesos de enseñanza institucionalizada es sin duda el profesorado el que mayor incidencia tiene pues, como dice Wilson (1992), «la calidad del sistema educativo depende en definitiva de la calidad de las personas que lo desempeñan» (p. 126). Evidentemente los y las docentes son los encargados de introducir, consciente y racionalmente, cambios y matices en sus acciones y en la vida del aula e institución (metodología, modelos de evaluación, tipos de relaciones interpersonales e interprofesionales, etc.) que modulan el clima de trabajo y marcan el ritmo y el cariz de los aprendizajes de los alumnos; no podemos, pues, renunciar a la evaluación del elemento que mayor incidencia ejerce en el proceso de enseñanza.

Tema 1

La evaluación de centros y profesores. Génesis de la conformación de un nuevo campo profesional, de estudio e investigación

Introducción

1. La diversificación de los planteamientos y ámbitos de la evaluación
 - 1.1. Características más relevantes del contexto socio-educativo en que emerge la evaluación de centros y profesores
 - 1.2. Debates y deliberaciones sobre la evaluación
2. El surgimiento y desarrollo de la evaluación de centros y profesores
 - 2.1. La emergencia de la evaluación de centros y profesores
 - 2.2. El desarrollo y consolidación de la evaluación de centros y profesores
 - 2.3. Generalización e internacionalización de la evaluación de centros y profesores
 - A) Actuaciones en los niveles educativos no universitarios
 - B) Actuaciones en la educación superior
3. A modo de recapitulación

Bibliografía

TEMA 1

LA EVALUACIÓN DE CENTROS Y PROFESORES. GÉNESIS DE LA CONFORMACIÓN DE UN NUEVO CAMPO PROFESIONAL, DE ESTUDIO E INVESTIGACIÓN

Eustaquio Martín Rodríguez.

UNED

INTRODUCCIÓN

La comprensión de la evaluación de centros y profesores como campo profesional de estudio e investigación requiere del conocimiento de su devenir, de cómo se ha ido construyendo su práctica profesional a lo largo del tiempo en función de los cambios que se han ido generando tanto en los contextos sociopolíticos, económicos y culturales por los que ha transitado, como por los debates y deliberaciones que han contribuido a su configuración como disciplina académica. A través de su estudio podremos no sólo analizar su evolución, identificando las soluciones que en cada momento van apareciendo, ya sea superando a las que las anteceden o coexistiendo con las que ya existían, sino también disponer de marcos de referencia epistemológico, metodológico y ético-políticos que nos faciliten entender y encuadrar los distintos enfoques, opciones o modelos que se han ido experimentado en la práctica profesional, así como hacer un balance de los progresos que han posibilitado el desarrollo de su teoría y práctica.

La historia de la evaluación de centros y profesores está estrechamente unida a la evaluación de programas y proyectos hasta la segunda mitad del siglo XX. Su legado constituye un rico patrimonio compuesto por un corpus teórico-práctico amplio y diverso que proyecta las concepciones que se han ido manteniendo en campos como el currículum, la didáctica y la organización escolar y que, a su vez, trascienden e influyen poderosamente en la evaluación. El conocimiento de cómo se ha ido decantando la teoría y práctica de la evaluación es, por consiguiente, un espacio que necesita ser precisado y clarificado si realmente queremos, con posterioridad, partir de una fundamentación solvente y rigurosa en el estudio de la evaluación de centros y profesores.

En este tema, por tanto, se ha tratado de reconstruir la secuencia de cómo surge, se desarrolla y generaliza la evaluación de centros y profesores. Nos hemos situado cronológicamente entre mediados de los años sesenta y

principios de los setenta del siglo XX, porque es en esos momentos cuando se inicia la evaluación de instituciones y docentes. Y hablamos de diversificación de ámbitos de evaluación porque estas últimas vienen a ampliar y enriquecer las primigenias de alumnos y curricular, en la consolidación de la evaluación como campo de estudio e investigación. Simultáneamente, cabe señalar que además de la ampliación de ámbitos, la diversificación hace referencia al surgimiento de planteamientos, de nuevas propuestas y modelos que se erigen como alternativas a la tradición psicométrica de carácter cuantitativo y experimental.

A lo largo pues de las páginas siguientes se ofrece una panorámica general de los múltiples problemas y cuestiones que afectan a la evaluación de centros y profesores. En esta primera aproximación se describen las características de los contextos por los que avanza y se configura la evaluación, se contraponen las distintas perspectivas que definen sus enfoques y se identifican las distintas fases de su evolución. Con ello no se agotan, en modo alguno, las posibilidades de análisis de todas sus problemáticas que, por otro lado, se irán progresivamente ampliando, enriqueciendo y debatiendo en los diferentes temas de este libro.

1. LA DIVERSIFICACIÓN DE LOS PLANTEAMIENTOS Y ÁMBITOS DE LA EVALUACIÓN

Cuando nos adentramos en el estudio de la historia de la evaluación educativa se pueden encontrar no pocas clasificaciones de las distintas períodos o etapas en las que se secuencia su evolución desde 1.930 hasta los últimos años del siglo XX. En su momento optamos por considerar tres grandes períodos en la génesis y desarrollo de la evaluación. Una primera época, a la que hemos convenido en denominar como «De la psicometría a la evaluación educativa», en la que tendrían cabida los períodos pretylerianos y tylerianos propiamente dichos y que llega hasta finales de los años cincuenta; una segunda época a la que hemos llamado «De la evaluación de los alumnos a la evaluación curricular» que comprendería fundamentalmente el período de expansión de la evaluación, por ser en estos momentos cuando se potencia la evaluación de programas y proyectos curriculares en EE. UU. y el Reino Unido, y cuyos límites temporales son más difíciles de establecer, en función del contexto cultural y nacional al que nos refiramos;

y una tercera época a la que conoceremos como «De la diversificación de planteamientos y ámbitos de evaluación», que cronológicamente se situaría desde los primeros años de los setenta, en los que se inicia el período de la profesionalización, hasta nuestros días, y en la que es posible identificar una serie de fases en las que adquirirán un fuerte protagonismo otros ámbitos de evaluación, como los de centros y profesores, y la consolidación de un pluralismo conceptual y metodológico de la evaluación.

Los dos primeros períodos son más propios de la asignatura de Evaluación de Programas y Proyectos por lo que para entrar más directamente en la evaluación de centros y profesores nos situamos históricamente en el que hemos denominado «De la diversificación de planteamientos y ámbitos de evaluación», dado que hasta finales de los años sesenta, no se empieza a generalizar la evaluación de centros dentro del sistema educativo norteamericano (Nevo, 1997) y no es hasta principios de los años setenta, con un aumento considerable en los años noventa, cuando se ha producido «un ímpetu repentino y significativo de los modelos y enfoques de evaluación del profesor» (Shinkfield y Stufflebeam, 1995: 9).

La aparición de la evaluación de centros y profesores se debe al incremento de la demanda pública de información de lo que acontece en el interior de las instituciones educativas. Desde instancias políticas, en la medida en que los gobiernos asumen la orientación y financiación de la educación como un bien de naturaleza social y pública, se han ido reforzando los cauces y mecanismos de obtención de información sobre los logros y resultados escolares. La sociedad civil no ha demostrado, asimismo, menos interés por conocer los niveles de calidad de los centros, ya sea porque en ellos está en juego el capital cultural de las nuevas generaciones, o porque, como también ocurre, tiene una posición definida en torno a los dilemas a los que se enfrenta hoy la enseñanza escolar. Por otro lado, desde una perspectiva académica los centros educativos y el profesorado han emergido como un lugar y espacio privilegiados de investigación y evaluación, posibilitando el desarrollo de campos disciplinares como el de la organización escolar, así como la creación de nuevas líneas de indagación como las de la teoría del cambio, escuelas eficaces, mejora de la escuela, eficacia y mejora escolar, la reestructuración escolar, etc.

La teoría y práctica de la evaluación de las instituciones educativas se ve inmersa, por consiguiente, en un marco complejo de relaciones e influen-

cias diversas, en el que conviven concepciones y visiones muy distintas en los planos epistemológico, ontológico, metodológico y ético-político. Por ello, es normal que tanto desde las condiciones externas, contexto socio-económico, político y cultural, como desde las condiciones internas, esto es, espacios académicos en los que se despliega la evaluación como disciplina, se encuentren posiciones, a veces enfrentadas y con escasos puntos en común, de lo que significa la evaluación de centros y profesores. Veamos más en concreto dichas condiciones.

1.1. Características más relevantes del contexto socioeducativo en que emerge la evaluación de centros y profesores

La evaluación institucional y del profesorado está vinculada históricamente con el movimiento de la «*accountability*» o rendición de cuentas. La grave crisis económica que vivió el mundo occidental, en los primeros años de la década de los setenta, como consecuencia del incremento de los precios del petróleo tuvo un efecto inmediato en la reducción de los presupuestos públicos y llevó aparejada el cuestionamiento de la rentabilidad de las inversiones educativas que se habían realizado durante las décadas de los años 50 y 60. Desde ese momento, distintas iniciativas políticas y sociales van a ser los promotores de medidas encaminadas a que el control de los currículos escolares no quede sólo en manos de los profesores y a que los centros de enseñanza informen públicamente de sus logros y resultados.

El movimiento de la *accountability* de los años 70 en Estados Unidos no es sino la proyección al terreno educativo de la situación de crisis socio-económica, cultural y política que en esos momentos se vive, y que va a cobrar una especial importancia ante el clima social de desconfianza y frustración por los productos y transformaciones de escasa calidad que las instituciones educativas obtienen en favor del progreso de la sociedad. No es de extrañar, por tanto, que sea justamente en estos momentos cuando se inicia de manera sistemática la evaluación de centros y profesores o empiezan a tomar fuerza otros planteamientos de evaluación desde perspectivas científicas distintas a la clásica tradición positivista.

Las experiencias desarrolladas con la *accountability*, cuyo impulso más decisivo se fomentó con la puesta en juego de las políticas neoliberales que definieron los mandatos de Ronald Reagan y Margaret Thatcher, van a coin-

cidir, a finales de la década de los años 80, con el surgimiento de lo que Neave (1988) acuñó con la denominación del *Estado Evaluador*, en el campo de la educación superior. El cambio de relaciones entre Estado, Sociedad y Educación Superior se va a traducir, entre otros aspectos, en el paso de una evaluación de mantenimiento a una evaluación estratégica. Tratándose, mediante esta última, de llevar a cabo una revisión permanente de los objetivos de las reformas educativas realizadas para acometer los errores y deficiencias de las mismas.

Las características más relevantes de la evaluación estratégica son la de tratarse de una evaluación a posteriori en vez de a priori, y la de que no concede importancia al proceso cuanto a los resultados, es decir, hay un cambio de énfasis del proceso al producto, del estado inicial al resultado. Este cambio de la función y del concepto mismo de evaluación implica, de acuerdo con Neave (2001: 224), que aspectos «tales como el rendimiento estudiantil, el promedio de calificaciones, la productividad de la institución y de los departamentos, que antes sólo ocasionalmente eran sometidos a revisión y ‘gestión de crisis’, pasaron entonces a ser el objeto de mecanismos en alto grado complejos y críticos, regularmente aplicados, regularmente revisados e íntimamente vinculados con el propósito estratégico, no sólo de orientar la política educativa nacional, sino también, al mismo tiempo, de que la institución hiciera su propia autoevaluación a fin de que pudiera discernir qué lugar ocupaba en el escenario nacional y planificar las maneras de mejorarlo o considerarlo».

El concepto de Estado Evaluador, según Whitty, Power y Halpin (1999: 55-56), puede aplicarse también a las etapas obligatorias de los sistemas educativos. Ellos lo asocian, en el marco de la reestructuración, al «racionalismo económico», en el sentido de considerar a la educación como un bien de consumo y a la política educativa como un medio por el que puede regularse y distribuirse con eficiencia y eficacia. Desde su perspectiva hay dos elementos que explican la naturaleza de las reformas curriculares que se van a llevar a cabo en los sistemas educativos, así como la función de la evaluación en los mismos. Por una parte, la omnipresencia del discurso económico y su cambio de relación con el discurso político en el contexto de la globalización. Y, por otra parte, el reforzamiento de las tendencias neoconservadoras de las políticas educativas.

La insistencia en los objetivos económicos implica un desplazamiento de los propósitos sociales y culturales de la educación. El marchamo de la «calidad», y la parafernalia de las normas para su catalogación, no es sino

un intento de llevar hasta sus últimas consecuencias los procesos de producción y el control de los productos, adoptados en las organizaciones industriales, al ámbito educativo. Y junto a esta perspectiva economicista que se explicita en los aspectos de eficacia y eficiencia en el logro de los resultados, también se van poniendo en evidencia los rasgos políticos de las tendencias neoconservadoras que cuestionan y se replantean la ausencia en los currículos de determinados contenidos en la enseñanza de ciertas áreas y/o asignaturas.

La ampliación de las horas de docencia, por ejemplo en determinados campos y disciplinas culturales (Humanidades, Lengua y Literatura); la revisión de los contenidos de enseñanza en dichos campos, incorporando programas y temas que unifican y dimensionan las culturas nacionales que todos los alumnos deben conocer, no significa otra cosa que el reforzamiento de una visión tradicional de la enseñanza en la socialización e instrucción homogeneizantes del alumnado. Whitty, Power y Halpin (1999) señalan que en Inglaterra y Gales, en Australia y en algunos Estados de EE.UU. se observa que las reglamentaciones curriculares se dirigen también, de alguna manera, a crear y recrear determinadas formas de identidad nacional, incrementado la enseñanza de la historia, geografía, literatura, etc., de esos países o estados.

En el cambio de rumbo de este discurso educativo no puede perderse de vista, además, las políticas de descentralización que delegan responsabilidades a las instituciones y profesores, pero que, en definitiva, siguen manteniendo y reforzando la centralización de los mecanismos de control. Y ello no deja de ser una manera limitada de plantear el protagonismo de los centros escolares y de sus actores, ya que, por una parte, su margen de maniobra es mínimo en la capacidad de decidir sus trayectorias, pues sus cometidos son de carácter metodológico-instrumental pero no entran en la deliberación de los objetivos o propósitos, y, por otra, no se facilitan las condiciones para la creación de una cultura institucional en la que se primen modalidades de participación y colaboración que desde proyectos pedagógicos concretos conduzcan al desarrollo profesional e institucional. Su finalidad es otra, ya que como pone de manifiesto Carnoy (1999: 151):

«... las reformas descentralizadoras, que hemos situado en el marco del proceso de globalización, se apoyan por tanto en la *retórica* de incrementar la productividad dando un mayor control a los responsables locales en la

toma de decisiones educativas. Pero la verdad es que tales reformas están comprometidas de manera inexorable con la reducción de la contribución gubernamental a la educación pública».

En este contexto, la evaluación adquiere una relevancia nunca antes imaginada. No en vano, ya durante la década de los 80 en EE. UU., tal y como indica House (1992), la evaluación fue considerada demasiado importante como para dejarla en manos de los evaluadores y las grandes burocracias crearon sus propias agencias de evaluación. Ernest House identifica estas evaluaciones como una modalidad de evaluación interna, para diferenciarlas de las realizadas por evaluadores independientes, y resalta los problemas de credibilidad y procedimientos que se derivan de la falta de independencia política de los evaluadores, así como el hecho de que una gran parte de la teoría de la evaluación fue construida y fundamentada desde las relaciones contractuales entre los evaluadores externos y las organizaciones.

Neave (2001) señala, en esta misma línea, la aparición desde mediados de la década de los 80 de poderosos «intermediarios» especializados que acometen, como agentes de evaluación, la recopilación de datos e informaciones de la educación superior. Y otro tanto ocurre en la educación no universitaria, en las décadas de los años 80 y 90, con la creación de instituciones *ad hoc* para la evaluación de la calidad (Institutos o Centros de Evaluación y/o Calidad).

Esta nueva reedición, corregida y aumentada, de la *accountability* con una gran proyección internacional refuerza las formas de valoración tecnocráticas de las instituciones educativas. Sus manifestaciones más visibles pueden observarse en aspectos tales como: el resurgimiento de enfoques que enfatizan lo metodológico sobre lo conceptual; la adopción formal de metodologías «cualitativas» en sus requerimientos instrumentales pero que desconsideran los procedimientos de evaluación asociados a las mismas en cuanto a la participación de los actores sujetos a la evaluación; el énfasis en las evaluaciones terminales o de resultados; la generalización de indicadores estándar de rendimiento; el uso de tests y pruebas estandarizadas para medir los conocimientos de los alumnos; la aplicación de criterios economicistas de valoración, etc.

1.2. Debates y deliberaciones sobre la evaluación

Desde los inicios de la actividad científica en los distintos campos de las Ciencias Sociales, el modelo que ha marcado la pauta a seguir ha sido el de las Ciencias Experimentales fundamentalmente porque ostentaban la patente de científicidad. Es decir, amoldando la indagación a un esquema experimental se ha buscado alcanzar el estatus de saber científico. Las claves de la aproximación experimental —también conocida como técnica, positivista— son la verificación de teorías utilizando el trabajo de laboratorio y generando leyes generales susceptibles de aplicarse a la población de la muestra estudiada; leyes y teorías que recogen la realidad tal y como es. La exclusiva de científicidad ha venido avalada por la tradición histórica y por análisis teóricos de la evolución de la ciencia que han concebido este modo de elaborar conocimiento como el único con rango científico.

Como se tendrá oportunidad de comprobar en el Tema 2, al analizar la relación entre investigación y evaluación, el desarrollo de la segunda estuvo estrechamente vinculada a la primera. Sobre todo, porque el patrimonio acumulado por la investigación supuso un aporte sustantivo en la clarificación y adopción de las metodologías de evaluación hasta principios de los años setenta. A partir de entonces se empieza a cuestionar la idoneidad del método hipotético-deductivo que, por lo general, ha utilizado y sigue utilizando la experimentación y cuantificación como elementos de identidad de sus modos de proceder indagatorios.

Casi todas las obras y autores que analizan, en algún momento, la historia de la evaluación resaltan que la reunión celebrada en el *Churchill College* de Cambridge, el 20 de diciembre de 1972, fue un acontecimiento con una enorme proyección para el desarrollo de la evaluación. En dicha reunión, se llegaron a una serie de acuerdos en torno a la naturaleza de los estudios, a los diseños de la evaluación, a las técnicas e instrumentos de recogida de información y al papel y responsabilidad de los evaluadores que, de alguna manera, recogían las experiencias compartidas por un conjunto cada vez más pujante de estudiosos de la teoría y práctica de la evaluación. Como fielmente transcribieron en sus notas MacDonald y Parlett, 1973 (puede verse en Sáez Brezmes y otros, 1990: 104-113):

«[la] conferencia de Cambridge era un intento de repensar la evaluación abiertamente, buscando lo que podía aprenderse de estas otras investigacio-

nes [de historiadores, economistas, sociólogos y antropólogos] y valorando las implicaciones, dificultades y potencialidades de los nuevos estilos de evaluación» (Sáez Brezmes y otros, 1990:107).

Nuevos estilos de evaluación que, con respecto a los acuerdos alcanzados, destacaban la insatisfacción con las evaluaciones que se venían realizando de las prácticas educativas porque no aportaban pruebas de los efectos de las mismas, por la insuficiente atención prestada a los procesos educativos, por centrarse excesivamente en los cambios mensurables y por el predominio de un tipo de investigación educativa que priorizaba la precisión de la medida y la generalidad de la teoría.

Las siguientes reuniones de este grupo, sobre todo la celebrada en 1975, en el que el tema debatido fue el de los «Métodos de estudios de casos», supuso un avance en la utilización de un lenguaje común. La creación de un espacio compartido entre autores anglonorteamericanos empezaba a dar sus frutos, aun cuando existiesen en ciertos momentos diferencias y desencuentros, en la vigorización de las nuevas vías y enfoques de evaluación. Como señala, Simons (1999: 76), el «estudio de casos era una metáfora que pretendía atraer a quienes buscaban una forma de integrar las solitudes de datos globales que se derivaban de diversas revisiones críticas de la tradición de la evaluación». Y, sin duda, denominaciones y propuestas como «evaluación naturalista» (Guba, 1978; Guba y Lincoln, 1981), «evaluación constructivista» (Guba y Lincoln, 1989); «evaluación democrática deliberativa» (House y Howe, 1999; House y Howe, 2000); y «evaluación comprensiva y evaluación basada en estándares» (Stake, 2006) representan algunos de los jalones del camino seguido hasta llegar al pluralismo de la evaluación.

En el Reino Unido, la publicación en 1976 de los artículos, por un lado, de Parlett y Hamilton de *La evaluación como iluminación* y, por otro, el de MacDonald de *La evaluación y el control de la educación*, vendrían a refrendar los nuevos caminos emprendidos en este campo, cuyas formulaciones van ir madurando como consecuencia del legado del movimiento del desarrollo curricular de los años sesenta. No debe olvidarse que, por ejemplo, MacDonald ya en 1968 había empezado a evaluar el *Humanities Curriculum Projet*, dirigido por Stenhouse, a cuyo enfoque se le calificó de «holístico» y que en 1973, con financiación de La Fundación Ford, emprendió el proyecto *SAFARI*.

Tanto la crítica al paradigma clásico «agrícola-botánico» como la consideración de los evaluadores como figuras políticas que manifiestan de distinta manera sus actitudes ante la distribución del poder en educación, propiciaron la difusión del paradigma «socioantropológico» y la clasificación política de los distintos tipos de evaluación: burocrática, autocrática y democrática. El enfoque de evaluación democrática de MacDonald (1985: 475-76), se caracterizaba por ser:

«... un servicio de información a la comunidad entera sobre las características del programa educativo. El patrocinio del estudio de evaluación en sí, no concede un derecho especial sobre este servicio. El evaluador democrático reconoce el pluralismo de valores y busca la representación de intereses diferentes en la formulación del tema. El valor básico es una colectividad de ciudadanos informados y el evaluador actúa como un corredor que intercambia información entre grupos que desean conocimientos recíprocos. Sus técnicas de recolección y presentación de datos, tienen que ser accesibles para personas no especialistas. Su actividad principal es la recolección de definiciones del programa y reacciones frente a él. Garantiza el secreto a los informantes y les ofrece el control sobre la utilización de la información que ellos le proporcionan. El informe no proporciona recomendaciones y el evaluador reconoce la necesidad de prudencia en el uso de la información. Establece una negociación periódica de sus relaciones con patrocinadores y participantes del programa. El criterio de éxito es el número de personas que pueden ayudar. El informe aspira al estatus de “best-seller”. Los conceptos claves de evaluación democrática son “secreto”, “negociación” y “accesibilidad”. El concepto clave de justificación es el “derecho a saber”».

Estos planteamientos y, de un modo especial, las críticas que MacDonald hacía a las evaluaciones que por entonces se desarrollaban en EE.UU., catalogándolas como burocráticas, es decir al servicio incondicional de los patrocinadores (también conocida como de tendencia gerencialista), o autocráticas, como servicio condicional al patrocinador pero que encuentran su legitimidad en los círculos académicos y de investigación, van a hacer que aparezcan nuevos elementos claves que centren el debate ya iniciado. En un trabajo presentado por Robert Stake (1983: 96), en la conferencia del Consejo Escocés de Investigación, celebrada en 1974 en Edimburgo, hablaba, al comparar los diseños preordenado y de réplica, de que una «evaluación educativa es evaluación de réplica (“responsive”), 1) si

se orienta más directamente a las actividades del programa que a los propósitos de éste; 2) si responde a las exigencias de información de la audiencia, y 3) si al informar del éxito y del fracaso del programa alude a las diferentes perspectivas de valor de las personas (del) entorno. De estas tres formas distintas puede ser un plan de evaluación de réplica».

La evaluación de réplica, respondiente, responsiva o sensible de Stake —pues de todas estas maneras se ha traducido al castellano el término *responsive*—, supone una alternativa al modelo de objetivos-resultados, originariamente propuesto por Ralph Tyler, ya que defiende que ha de centrarse preferentemente en las actividades, es decir, más en los procesos que en los resultados. Asimismo, su aportación contribuyó a la democratización de la evaluación, dado que insistió en la necesidad de que la misma no estuviera exclusivamente al servicio de los patrocinadores y considerará las necesidades de los distintos grupos e instituciones concernidas por la evaluación, así como que no se ignorasen los distintos valores de personas y colectivos.

En este mismo sentido, Guba y Lincoln (1989) resaltan que los problemas endémicos de las primeras generaciones de la evaluación han sido fundamentalmente la tendencia al gerencialismo, la deficiente consideración del pluralismo de valores y la excesiva dependencia de un paradigma tecnológico o científico-racional de carácter positivista.

La tendencia hacia el gerencialismo se expresa, sobre todo, en el control y poder que ejercen los patrocinadores de la evaluación para, en última instancia, ser ellos quienes decidan las preguntas, propósitos y divulgación de los estudios por considerarse los dueños de los mismos. Este proceso es completamente injusto para aquellos grupos que no tienen el estatus del dirigente. También restringe los derechos del ciudadano, porque son las autoridades administrativas o los directivos los que controlan toda la información y la divulgación del informe, y en consecuencia conculcan el derecho a la información. Puede darse una connivencia entre patrocinador y evaluador de modo que sean ellos quienes, al margen de otras personas y grupos interesados, determinen todos los elementos relevantes de la evaluación.

El pluralismo de valores era una problemática ajena a la investigación y evaluación porque, hasta hacía muy poco tiempo, se había sostenido que la sociedad compartía valores y que todos sus miembros estaban aculturados y asumían dichos valores. Sin embargo, la inclusión del juicio en la evalua-

ción puso en evidencia que los objetivos eran instrumentos que incluían valores y ello implicaba un cierto acuerdo sobre éstos. La cuestión de qué valores predominan en una evaluación o qué valores pueden negociarse empieza a emerger como un asunto muy importante. La situación se complica ya que, por una parte, se ha defendido que aunque existen diferencias de valores, los hallazgos de una evaluación pueden ser válidos porque la metodología usada es científica y las ciencias están demostradamente libre de valores. Pero si se postula, en cambio, que la ciencia no está libre de valores eso quiere decir que no sólo los hallazgos están sujetos a diferentes interpretaciones sino que los hechos en sí mismos están determinados por la interacción con el sistema de valores del evaluador. Y el problema adquiere otras dimensiones, dado que entonces todos los hechos de evaluación se convierten en actos políticos. Es decir cada acto de indagación ya sea acto de investigación, acto de evaluación o análisis de políticas llega a ser un acto político en este sentido. La afirmación de la evaluación libre de valores está en consonancia con una tendencia gerencialista ya que, si no intervienen los valores, los hallazgos de una evaluación representan los hechos tal y como realmente son, y deberían ser por tanto aceptados como verdades objetivas. El hecho de que el patrocinador fije los parámetros y los alcances del estudio parece que no tiene entonces trascendencia porque los hechos parecen no estar influidos por estas decisiones.

Existe, asimismo, un compromiso excesivo con el paradigma positivista que entre otras tiene las siguientes limitaciones: una ausencia de contexto, salvo que esté cuidadosamente controlado para generalizar resultados; gran dependencia de las medidas cuantitativas, porque el rigor que promete el paradigma parece descansar en la dureza de los datos que son conseguidos a través de estos procedimientos; legitimación de la evaluación por el uso de «métodos científicos» que nos proporcionan un conocimiento de las cosas tal y como suceden, por lo que no podemos resistirnos a tales evidencias y no estamos autorizados a ponerlas en tela de juicio. Esto último es lo que se denomina la coerción de la verdad, la verdad es innegociable y si un evaluador la proporciona los clientes no tienen ninguna oportunidad de rechazarla como alejada de su experiencia. Pero es que, además, el uso de estos métodos impide utilizar claves alternativas de pensamiento sobre el objeto a evaluar, siendo impensable incluir los puntos de vista de los participantes. Para el investigador o evaluador que se adhiere a este paradigma, como la ciencia está libre de valores, personalmente conlleva no asumir

responsabilidad moral sobre sus acciones, no puede ser culpado porque está diciendo la verdad. Tampoco tiene ninguna responsabilidad sobre lo que sigue a la evaluación ya que su papel termina cuando el informe es entregado. Ni siquiera es responsable de los hallazgos porque él simplemente refleja lo que sucede en la realidad.

2. EL SURGIMIENTO Y DESARROLLO DE LA EVALUACIÓN DE CENTROS Y PROFESORES

La evaluación de las instituciones educativas y de los docentes se ha constituido, en el último tercio del siglo XX, en un campo de estudio en permanente auge en el que se entrecruzan perspectivas, visiones e intereses de muy diversa naturaleza. Parece claro que su importancia viene determinada por la necesidad, mayoritariamente aceptada, de contar con información que permita conocer y optimizar la contribución de los centros escolares y profesores a la formación del alumnado. Su protagonismo se debe a que son espacios sociales, dotados de sentido y significado, para hacer realidad un proyecto pedagógico. De su implicación y compromiso depende la viabilidad y las formas concretas de llevar a cabo las aspiraciones y expectativas educativas de la sociedad.

2.1. La emergencia de la evaluación de centros y profesores

Como se ha señalado en distintos momentos, la evaluación de centros y profesores se inicia e impulsa por imperativo de la *accountability*. Las experiencias llevadas a cabo, primero, en EE.UU. y, más tarde, en Inglaterra van a ir generalizando la obligación de rendir cuentas bajo distintos procedimientos y formas que no siempre son equiparables por las peculiaridades propias de cada uno de los países. Como señalan McCormick y James (1996: 31):

«[aunque] haya que considerar con precaución cualquier tipo de generalización, la dominancia de los modelos económicos de *input-output* en los Estados Unidos ha estimulado la dependencia de la evaluación o, de manera más precisa, de la valoración de los resultados o de los productos de la enseñanza. De ahí la omnipresencia de las pruebas de rendimiento y de los esquemas de evaluación que engloban la valoración de la actuación de profesores y estudiantes».

La necesidad de conocer la rentabilidad de las inversiones en educación se materializó en un mayor control de las instituciones y de los docentes. Sin embargo, mientras en EE.UU. se llevó a cabo a través de los grandes programas de pruebas de rendimiento que trataban de valorar los aprendizajes de los alumnos, es decir, los resultados o logros, en Inglaterra la evaluación de centros adoptó múltiples formas y no siempre se ocuparon exclusivamente de los resultados escolares.

Es más, en este último caso, el problema no se reducía al control propiamente dicho, ya que como afirmaba Lawton (1.987: 98):

«[la] obligación de rendir cuentas en sí no tiene por qué ser negativa; sólo puede representar un peligro si se traduce en la adopción de medidas basadas únicamente en el rendimiento, tales como la asignación de créditos en función de los resultados concretos o la evaluación de la competencia de los educadores por un criterio conductista del rendimiento. En todos los sistemas suele haber medios de aprovechar los recursos más eficazmente, pero el análisis de los costos nunca debe conducir a la distorsión del proceso pedagógico mediante una evaluación demasiado esquemática del rendimiento».

Por otra parte, según señalan Shinkfield y Stufflebeam (1995: 14-15), en un estudio llevado a cabo por la *National Education Association* (NEA) estadounidense, en 1964, ya se ponía de manifiesto que la mitad de los sistemas escolares seguían procedimientos formales en la evaluación de sus profesores y que se utilizaban escalas escritas en 3 de cada 4 escuelas para profesores, en prueba, y en 2 de cada 3 para profesores fijos. Casi invariablemente, el director era responsable del proceso de evaluación, pero ocasionalmente compartía esa responsabilidad con otros administradores. Pocos años más tarde, Stemnock (1969) no sólo halló, igual que el estudio de 1964 había descubierto, que los directores son casi siempre responsables de las evaluaciones, sino también fue capaz de concluir que los profesores estaban en gran medida de acuerdo en que el director era quien debería ser responsable de su evaluación profesional. Stemnock también descubrió que el 90 por ciento de las escuelas estudiadas en el ámbito nacional tenían procedimientos de evaluación de los profesores. Es significativo tener en cuenta que en 1972, en otro estudio del NEA, el 55 por ciento de los sistemas escolares habían revisado sus procedimientos de evaluación del profesor durante los tres años precedentes.

Otra cosa muy distinta fue la actitud que los profesores mantenían sobre su evaluación. Zelanak y Snider (1974) compararon la actitud de los profesos-

res que creían que la evaluación tenía propósitos administrativos con la de aquellos otros que asumían que su finalidad era la mejora de la enseñanza. Como era de esperar, mientras los segundos apoyaban este tipo de actividades y procesos, los primeros cuestionaban y percibían negativamente que la evaluación fuera utilizada para justificar despidos o promociones, modificar sus funciones o crear archivos permanentes de información.

De igual modo, la bibliografía de estos años recoge las controversias y dudas de los profesores sobre los criterios empleados para juzgar una enseñanza efectiva. Según Shinkfield y Stufflebeam (1995: 16-17):

«[había un alto] grado de acuerdo en que existía una tendencia a que el evaluador se centrara en los rasgos y características personales, en lugar de en los comportamientos, vinculados a la gestión eficaz de las condiciones de aprendizaje dentro de la clase, y de las propias destrezas pedagógicas [...] Algunas investigaciones hallaron que los administradores juzgaban muy a menudo la competencia del profesor basándose en (a) la aptitud docente, (b) la capacidad disciplinaria, (c) la erudición, y (d) la personalidad, y demostraban en sus trabajos que los instrumentos de evaluación del profesor y las valoraciones de los evaluadores, comúnmente utilizados en los sistemas escolares, eran poco fidedignos. Los estudios, por su parte, que se ocupan de la validez (ya sea de contenido, correlativa, o de constructo) de los instrumentos para la evaluación del profesor son a la vez poco comunes—exceptuando aquellos que tratan de la valoración del alumno sobre la competencia del profesor— e inconcluyentes».

El tema, sin embargo, más conflictivo de la historia de la evaluación y que ya en los años setenta sale a la luz con toda su virulencia, es sin duda la utilización de los resultados del aprendizaje de los alumnos como el criterio más importante de la eficiencia de los profesores. Las posiciones de los autores varían desde los que estaban totalmente en contra de este criterio hasta los que defendían el mismo. Rosenshine y Furst (1971) argumentaron que se conoce muy poco sobre la relación entre el comportamiento del profesor y el crecimiento de los alumnos, y que, al menos, existían cinco variables consideradas como destrezas de los profesores que estaban asociadas al rendimiento de los alumnos: claridad, variabilidad, entusiasmo, orientación a la tarea y oportunidades del alumno para aprender. Glass (1974) criticaba el uso de los tests estandarizados de logros por sus limita-

ciones para medir fielmente los efectos de los profesores sobre los avances de los alumnos en conocimientos durante un periodo de tiempo. Gage (1973) resalta que uno de los problemas más significativos de la investigación sobre la evaluación de la competencia de los profesores había sido el haberse centrado en un solo criterio o como muchos en dos o tres criterios de efectividad, desconsiderando la influencia que otras variables tenían como, por ejemplo, el contexto real en que se lleva a cabo la enseñanza.

Por el contrario, McNeil (1967) y Pophan (1971) intentaron aislar un indicador válido para evaluar las destrezas pedagógicas del profesor basándose en el logro de los objetivos educativos por parte de los alumnos. Con sus trabajos trataron de demostrar que organizando acontecimientos pedagógicos para producir los cambios deseados, y evaluando el logro de los alumnos de los objetivos educativos se pueden obtener índices selectivos de la actuación de los profesores relacionados con el rendimiento escolar de dichos alumnos. Pophan desarrolló una serie de tests de actuación docente, que variaban de unas materias a otras, a los efectos de medir, a través de procedimientos de pretest y postest, el logro de una serie de objetivos educativos por parte de los alumnos.

En España, es con la Ley General de Educación de 1970 cuando se introduce, por primera vez, el concepto de evaluación, proyectándose su uso no solamente para el aprendizaje de los alumnos sino también, de una manera más ambiciosa, referida al sistema educativo. Se establece, asimismo, tanto en la propia Ley como en el Decreto 664/1973, de 22 de marzo, que una de las funciones de la inspección será la de «Evaluar el rendimiento educativo de los centros docentes y profesores de su zona respectiva o de la especialidad a su cargo en colaboración con los Institutos de Ciencias de la Educación».

El desarrollo de esta función no fue más allá de su plasmación en este nuevo marco normativo durante bastante tiempo. Es posible que en esos años se concretasen algunas acciones aisladas en este sentido, pues ya, por ejemplo, en 1966 el Centro de Documentación y Orientación Didáctica de Enseñanza Primaria (CEDODEP) elaboró una escala de evaluación de centros, pero sin duda fueron escasamente relevantes y al día de hoy contamos con poca documentación como para asegurar que se emprendiese sistemáticamente ningún plan, a excepción de algunos informes sobre la evaluación de la implantación de la Ley del 70 que fueron realizados por un grupo formado *ad hoc*, un estudio del Servicio de Inspección Técnica de Educación (SITE)

que, en 1976, analizaba el rendimiento escolar de los alumnos a nivel nacional de cara a valorar la calidad de la enseñanza y la investigación emprendida por Muñoz Arroyo (1979) durante los años 1973 a 1976 en la que se propone un cuestionario para la observación y evaluación de centros de EGB.

En el contexto universitario, no obstante, durante estos años empiezan a aparecer algunas publicaciones como las de Rodríguez Diéguez (1973) con una propuesta de un cuestionario para la observación y valoración de la función del profesor; la de García Hoz (1975), sobre las pautas para la evaluación de centros educativos; la de Isaacs (1977), en la que se recogen instrumentos y procedimientos para la evaluación de profesores, objetivos educativos y centros escolares; y un artículo de Escudero Escorza (1979) que pasa revista a los estudios y realizaciones que en el campo de la evaluación educativa se estaban llevando a cabo entonces por la red del Instituto Nacional de Ciencias de la Educación y de los Institutos de Ciencias de la Educación (INCIE-ICE), y en el que se concluye

«[que la] simple realidad de lo sintetizado es una muestra evidente de que, con más apoyo de la Administración, con unos equipos de investigación estables, con unos evaluadores con capacidad de maniobra, la evaluación educativa en España puede alcanzar un desarrollo algo más que decoroso. El camino recorrido es suficiente como para abordar, con garantías, el proceso de análisis y mejora de nuestro sistema educativo».

2.2. El desarrollo y consolidación de la evaluación de centros y profesores

A finales de los años 70 y principios de los 80 los movimientos de evaluación de centros y profesores van a proseguir su desarrollo y consolidación. Por un lado, la publicación del informe *A Nation At Risk* en 1983, por la National Commission on Excellence in Education de EE. UU., refuerza la conciencia de amplios sectores de la sociedad americana, cuyo descontento se había ido agudizando en los años precedentes, de la necesidad de reformar la educación, sobre todo, en los niveles primarios y secundarios. Una parte importante de *Una nación en peligro* aborda la urgencia de mejorar la actuación del profesor, las cualificaciones de aquellos que acceden a la profesión y la retención de los mejores profesores. De nuevo el desarrollo de la *accountability* en el campo educativo tomaba nuevos bríos y el examen

riguroso de las intenciones y resultados educativos se convirtió en una cuestión de importancia nacional.

Por otro lado, y paralelamente, las propuestas de autoevaluación institucional, que habían surgido en el Reino Unido durante los 70, van a potenciarse especialmente en los años 80. Su extensión en esos momentos se debe, en palabras de Elliott (1.986: 238), a las siguientes razones:

1. «Las LEAS carecen de los recursos necesarios para supervisar la labor de los profesores en las escuelas para hacer un trabajo eficaz.
2. La autoevaluación es más aceptable para los profesores que la supervisión externa, toda vez que parece salvaguardar su autonomía profesional.
3. La supervisión externa choca con la idea del desarrollo profesional. Sólo a través de la autoevaluación pueden los profesores desarrollarse profesionalmente».

El movimiento de autoevaluación institucional se desarrolló rápidamente por el Reino Unido, Australia y estuvo presente en uno de los programas de la OCDE de los años ochenta, el *International School Improvement Project*. Helen Simons (1985:1) advertía que:

«[la] efervescencia que actualmente se fomenta y que un día se temió bajo el fantasma de la autoevaluación se ha acelerado dramáticamente en el Reino Unido durante los últimos ocho años (1977-1985). Las conferencias y los cursos de formación cada vez se hacen más eco de esta tendencia; incluso escuelas universitarias y universidades incluyen ahora la autoevaluación en sus planes de estudio. El Departamento de Educación y Ciencia (DES) ha potenciado los cursos y jornadas regionales, fundando el Consejo Escolar de Investigaciones y Proyectos, mientras que las Autoridades Locales de Educación (LEAs) han seguido la iniciativa tomada por las Autoridades Educativas Centrales en Londres (ILEA) en 1977 consistente en animar a los centros a que se autoevalúen».

Según esta misma autora, (Simons, 1999: 223), las actuaciones de autoevaluación escolar se llevaron a cabo como consecuencia de, al menos, tres contextos de influencia que se identifican concretamente con el legado del movimiento de reforma curricular; con las nuevas modalidades de formación permanente y con las peticiones de rendición de cuentas, siendo este último, según la mayoría de los autores, el que tuvo una mayor y más poderosa incidencia en su generalización durante los años 80.

El Departamento de Educación y Ciencia, a raíz de la nueva Ley de Educación de 1980, exigió a las escuelas la publicación de los resultados de los exámenes, solicitó, ya en 1981, que las LEA abordaran una revisión curricular y, en 1983, anunció que procedería a una supervisión continua de las actividades curriculares. De estas iniciativas del gobierno central surgen las propuestas y esquemas de autoevaluación de las escuelas hechas por las LEA como respuesta de éstas, en cierto modo, a la exigencia de sus responsabilidades curriculares por parte de aquel. Y un poco más tarde, como señalan Clift, Nuttall y McCormick (1989: 7), y conectado con la negociación de contratos y salarios de los docentes, el foco de atención se centrará en la evaluación del profesor, no en el currículum ni en la escuelas, aunque hacer «la distinción entre la evaluación del profesor y otras formas de autoevaluación no es fácil ya que los métodos empleados pueden ser idénticos».

Las directrices político-administrativas de la autoevaluación para la rendición de cuentas se encontraron entonces, en el Reino Unido, con un bagaje de experiencias y realizaciones provenientes tanto del legado del desarrollo curricular como del replanteamiento de la formación permanente del profesorado que configurarían un espacio y un clima muy diferentes para el debate sobre las consecuencias de este tipo de orientaciones. Como hemos tenido oportunidad de ver, las aportaciones de Stenhouse (1975) sobre la consideración del «profesor como investigador», las propuestas de «evaluación democrática» de MacDonald (1976) y de la «evaluación como iluminación» de Parlett y Hamilton (1976), entre otras, así como los intercambios de trabajos e ideas con autores norteamericanos como Ernest House y Robert Stake sobre las realizaciones de la accountability y los avances en el campo de la evaluación, propiciaron un contradiscurso, de carácter académico, desde el que se defendería la autoevaluación desde una perspectiva de desarrollo profesional de los docentes. Ello unido a las propuestas de «formación del profesorado centrada en la escuela» (Bridges y Eynon, 1983) fueron resaltando el protagonismo de los profesores y de las instituciones educativas en un intento de articular equilibradamente los desarrollos del currículum, de las escuelas y del profesorado.

De este modo y en estas circunstancias es como se abren paso formas muy diferentes de abordar la autoevaluación de centros y profesores, una asociada a las directrices gerencialistas y de control para la rendición de cuentas, y la otra vinculada con el desarrollo institucional y del profesorado

de cara a que estos procesos intrainstitucionales sean los que informen y fundamenten la mejora del aprendizaje, la enseñanza y la organización escolar. Ambas formas de enfocar y desarrollar la autoevaluación están bien documentadas en la literatura especializada como tenderemos ocasión de comprobar en diversos temas.

A estas tendencias de evaluación institucional y del profesorado no van a permanecer ajenos ninguno de los niveles de los sistemas educativos. La implementación de políticas y el desarrollo de actuaciones encaminadas a la mejora de la calidad de la educación superior han sido cada vez más frecuentes en el contexto de lo que Neave (1988) acuñó como el *Estado Evaluador*. Recuérdese que con esta expresión se hace referencia a la alteración que se produce en las relaciones entre Estado, Sociedad y Educación Superior en una gran parte de los países de Europa Occidental a finales de los años 70 y principios de los 80.

Aunque en otro momento tendremos oportunidad de hacer un análisis más detallado, Neave y Van Vught (1994) argumentan que el paso del *Estado Facilitador*, que posibilitó la expansión de la educación superior haciendo frente a los retos de la gran demanda de estudios mediante la continuidad de las relaciones entre las instituciones de enseñanza superior y los gobiernos, al *Estado Evaluador* pone de manifiesto un nuevo contexto de cambios sociales, económicos y culturales de un peso considerable. De entre ellos habría que destacar, aún cuando la influencia de cada uno de los factores puede ser diferente en los distintos países, los siguientes: las reducciones presupuestarias en las dinámicas de crisis del Estado de Bienestar; las reformas generalizadas en el sector público como consecuencia del inicio de las políticas neoliberales, el papel atribuidos a la ciencia y la tecnología en la inserción de las naciones y de los bloques regionales en un contexto internacional cada vez más competitivo, el fenómeno de masificación de la educación superior que se desencadena en la mayoría de los países en la década de los 60, y la búsqueda de una mejor adecuación de las relaciones con el mercado del trabajo y el mundo de la industria.

En este complejo y cambiante contexto social las iniciativas de evaluación son numerosas. Por ejemplo, en EE.UU. en 1983 el 98% de los distritos escolares tenían algún tipo de evaluación del profesorado en vigor, y, en 1980, William Sanders y Robert McLean de la Universidad de Tennessee comenzaron a explorar la viabilidad de utilizar una metodología estadística

de modelo mixto para superar los inconvenientes que surgían para situar los datos de los estudiantes en un sistema de evaluación basado en los resultados. Ellos crearon lo que ahora se conoce como el Sistema de Evaluación de Valor Añadido de Tennessee (TVAAS), que evalúa el impacto de los sistemas educativos, escuelas y profesores sobre los éxitos de los alumnos, comparando las puntuaciones obtenidas por estos en tests de logros durante al menos dos años.

La España de los años 80 tiene una serie de rasgos distintivos con respecto a estos otros contextos culturales que venimos analizando, tanto por su larga tradición como país centralista, aunque a finales de 1978 con la aprobación de nuestra Constitución se va a entrar en una nueva época con un modelo de Estado cada vez más descentralizado, como por las circunstancias sociopolíticas autoritarias que se vivieron durante muchos años. Aquí el movimiento de rendición de cuentas no tuvo ninguna proyección, el país estaba inmerso en un difícil camino hacia la transición política democrática y el único eco que la evaluación de centros y profesores tiene, en instancias asociadas con la administración educativa, es una publicación de la Inspección General de Educación Básica del Estado del año 1983 con el título de *Instrumentos de Supervisión*. En ella se hace una recapitulación de instrumentos de autoría institucional, salvo en los cuestionarios para la «observación y evaluación de Centros de EGB» y para la «observación y evaluación del profesor» elaborados por Muñoz Arroyo y Rodríguez Diéguez respectivamente en sus tesis doctorales, y que ya recogimos con anterioridad en el apartado dedicado a la evaluación de la década de los 70.

2.3. Generalización e internacionalización de la evaluación de centros y profesores

Con la llegada de los años 90 se van a consolidar definitivamente las tendencias y movimientos de evaluación de centros y profesores y aparecen, en otros muchos países, nuevas instituciones creadas *ad hoc* para acometer un ámbito de evaluación que, aunque no puede decirse que sea nuevo, sí se potenciará ahora de una manera extraordinaria, la evaluación de los sistemas educativos. Como en las décadas anteriores no se pueden perder de vista todo un conjunto de circunstancias socioeconómicas, políticas y culturales de estos años que modelan el devenir de la educación en general y de la evaluación en particular en su tránsito hacia el siglo XXI.

La irrupción de fenómenos como la globalización, con su correlato del triunfo del modelo de economía de mercado, el paso definitivo de la modernidad a la postmodernidad, con la desaparición y pérdida de legitimidad de los grandes relatos y discursos universales heredados de la Ilustración, y el progresivo avance de una nueva revolución, la tecnológica, que a diferencia de la industrial tiene como materia prima a la información en vez de la energía, van a ir desencadenado un nuevo orden y una estrategia geopolítica muy diferente en el concierto internacional.

Sin entrar en un análisis de todos estos fenómenos complejos, cuya problemática es tan amplia que nos desborda aquí y ahora, sí podemos destacar una serie de dimensiones que, en algunos casos, ya se venían gestando desde los últimos años de la década de los 70, y que, en otros, irrumpen ahora con fuerza, a saber: la acentuación de las políticas económicas neoliberales, definidas en muchos casos por poderes que no han sido legitimados democráticamente; la escasa capacidad de maniobra para mantener siquiera los programas sociales que las políticas socialdemócratas habían logrado extender en el apogeo del estado de bienestar, dadas las tendencias desreguladoras y liberalizadoras del mercado que achican el tamaño y la presencia del Estado en sectores sociales claves; el reforzamiento de los discursos de calidad de carácter eficientista, en los que ahora se resaltan los elementos de competitividad y de la relación costos-beneficios propios del mundo industrial y de la economía financiera.

A) Actuaciones en los niveles educativos no universitarios

La aparición de nuevas instituciones y el desarrollo de la evaluación de los sistemas educativos durante esta década, no cabe duda que es un intento de racionalizar las políticas y realizaciones educativas que están llevándose a cabo en el sector público, por lo que han de controlarse de cara a su mejora, dada la magnitud de los presupuestos que se dedican a estos menesteres y el malestar creciente que se va generando sobre la calidad de la enseñanza. En España, la Ley 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo (LOGSE) recoge en el artículo 62 —apartado 3—, que la «evaluación del sistema educativo se realizará por el *Instituto Nacional de Calidad y Evaluación*» (INCE) —en la actualidad, Instituto Nacional de Evaluación Educativa, INEE, <http://www.mecd.gob.es/inee/portada.html>, consultado 13/05/2013—, que asumió las siguientes funciones:

«... analizar el grado de eficacia y de eficiencia del sistema educativo; elaborar sistemas de evaluación para las diferentes enseñanzas reguladas en la LOGSE; elaborar un sistema estatal de indicadores; coordinar la participación española en los estudios internacionales sobre evaluación; colaborar y cooperar con las Administraciones educativas en materia de evaluación; proporcionar e intercambiar información con las Administraciones educativas para la toma de decisiones, proporcionar información al Consejo Escolar del Estado; informar a la sociedad sobre el funcionamiento y resultados del sistema educativo; difundir los resultados de sus evaluaciones e investigaciones» (Web del INCE, 2001).

En Italia, en julio de 1999, se procedió a la creación del *Istituto Nazionale per la Valutazione del Sistema dell'Istruzione* - <http://www.invalsi.it/invalsi/istituto.php>, consultado 13/05/2013 -, institución que asume a partir de ese momento las funciones que venía desempeñando desde el año 1994 el *Centro Europeo dell'Educazione*. Otro tanto ocurre en distintos países del mundo, que o bien ya disponían o bien crean nuevas instituciones o departamentos dentro de los centros existentes, como en los casos: de Francia, con la *Direction de la Programmation et du Développement* (DPD) - <http://www.education.gouv.fr> -; de Dinamarca, con *The Danish Evaluation Institute* (EVA) - <http://www.eva.dk/>, consultado 13/05/2013-; de Australia, con *The Australian Council for Educational Research* (ACER) - <http://www.acer.edu.au>, consultado 13/05/2013-; de EE.UU., con el *National Center for Education Statistics* (NCES) - <http://nces.ed.gov/> consultado 13/05/2013-; o de Portugal, con el *Instituto de Inovação Educacional* (IIE) - <http://www.portugal.gov.pt/pt/os-ministerios/ministerio-da-educacao-e-ciencia.aspx>, etc.

Lo primero que salta a la vista con la existencia y creación de este tipo de instituciones es la importancia creciente que va tomando la evaluación en el contexto internacional. Los distintos países dan un paso más, en el sentido de considerar que necesitan una información y un conocimiento más amplio y global de lo que está ocurriendo en la educación que completamente, de alguna manera, la obtenida a través de la evaluación de centros y profesores como elementos idiosincráticos del sistema. Un tema más problemático es cómo se llevan a cabo estos estudios evaluativos, las proyecciones que se hacen desde los datos recogidos sobre la calidad del profesorado y los centros educativos, así como su capacidad para crear un estado de opinión pública mediante la difusión de sus resultados a través de los medios de comunicación.

A la creación y diseminación de una cultura de la evaluación, en los niveles no universitarios de los sistemas educativos, van a contribuir tanto las organizaciones nacionales de cada país como ciertos organismos internacionales a través de la promoción de intercambios de experiencias y de la financiación de estudios regionales. Este sería, por ejemplo, el sentido de las Jornadas organizadas en Madrid en 1998, por el INCE, sobre el tema de *Instituciones para la evaluación del sistema educativo*, en las que participaron responsables de evaluación de las Comunidades Autónomas y del MEC; o las publicaciones y plataformas promovidas por la Organización de Estados Iberoamericanos sobre la *Evaluación de la calidad de la educación*, cuyas ponencias se recogen en el n.º. 10 de la *Revista Iberoamericana de Educación*, de 1996 (disponible en la versión digital de dicha revista, en <http://www.oei.es/index.php>).

De igual modo organismos como el *Centro para la Investigación e Innovación Educativas* (CERI), de la OCDE, viene trabajando desde los años 90 en el Proyecto *PISA* (Programme for Indicators of Student Achievement) en la elaboración, definición y experimentación de un conjunto de indicadores para la evaluación del rendimiento de los alumnos de los distintos sistemas educativos de los países participantes, así como para la comparación de los rendimientos de los alumnos de dichos países. La Comisión Europea publicó en el año 2000 uno de los proyectos que emprendió sobre la calidad de la educación en el que se proponen dieciséis indicadores, en campos como los logros en ciertas asignaturas o áreas curriculares, la continuidad y éxitos en los estudios, el seguimiento y control de la educación, y sobre la estructura y recursos educativos.

Dentro del programa Sócrates, la Comisión Europea también ha desarrollado el proyecto piloto europeo de evaluación de la calidad en la enseñanza escolar; INCE (1999), en el que han participado 101 escuelas de 18 países experimentando procesos de autoevaluación institucional de cara a «compartir y conseguir un conocimiento más profundo de la calidad de los centros escolares, y de cómo mejorarla».

Lo que parece claro es que, desde los primeros años de la década de los 90, la evaluación de centros y profesores se va a constituir en el tema estrella de las preocupaciones de las políticas educativas. En España, por ejemplo, durante los años 1990 y 1991 se va a gestar y elaborar el Plan de Evaluación de Centros Docentes de niveles no universitarios, *Plan EVA*, en

el que trabajó un grupo reducido de profesionales (pertenecientes al CIDE y a la Inspección Central de Educación) y cuyo programa piloto fue experimentado, en el curso académico 1991-92, en 40 centros educativos del territorio MEC. Esta iniciativa, junto con las propuestas para la *Valoración de la labor profesional de los docentes* que van a ver la luz en 1996, no es sino el reflejo de la filosofía de la administración educativa sobre el papel de la evaluación en la calidad de la educación, cuyo sustento normativo y documental se va a reflejar primero en la LOGSE, más tarde, en el documento *Centros Educativos y Calidad de la Enseñanza*, MEC (1994), y posteriormente en la *LEY ORGÁNICA 9/1995, de 20 de noviembre, de la participación, la evaluación y el gobierno de los centros docentes*.

Todos estos programas de evaluación, en el caso español, han venido desarrollándose progresivamente, aunque el Plan Eva desde el año 1995 fue perdiendo poco a poco su vigencia y seguidamente fue sustituido, coincidiendo con la llegada de la administración popular, por una nueva propuesta vinculada a la gestión de la calidad total de la educación. Dicha propuesta, para la que el MEC (1997, a y b, 1998) ha elaborado una serie de documentos y guías, opta por la modalidad de autoevaluación institucional en la medida en que (como se dice en la Presentación de la *Guía de Autoevaluación*):

«el Ministerio de Educación y Cultura puso en marcha un programa destinado a hacer de la calidad no sólo un objetivo, sino también —y en esto radica la novedad— un método para la mejora continua de esas organizaciones fundamentales, que son los centros docentes, en donde tienen lugar los procesos primarios de la calidad del sistema educativo» (MEC 1997a).

Con la transferencia de las competencias en educación a todas las Comunidades Autónomas, a partir del año 2000, la implantación y desarrollo de los programas de evaluación quedó en manos de dichas entidades territoriales. A día de hoy la situación de la evaluación de centros y profesores es muy diferente de unas Comunidades a otras. No obstante, tanto en la Ley Orgánica de Calidad de la Educación, de 23 de diciembre de 2002, como en la Ley Orgánica de Educación, de 3 de mayo de 2006, se recogen la evaluación voluntaria del profesorado y la evaluación del sistema educativo. En su momento, es decir, en los temas respectivos analizaremos los planteamientos y modalidades de evaluación recogidos en dichas leyes, así como su implantación en el sistema educativo.

En cualquier caso, hay que resaltar que la adopción de la modalidad de autoevaluación va a ser una constante que encontraremos, a lo largo de los años 90, en las propuestas que se van a ir haciendo para la evaluación de las instituciones de los distintos niveles del sistema educativo. Tal y como vimos en los años 80, hemos esbozado ahora y, seguidamente, analizaremos en la educación superior; las variedades que puede adoptar son muy distintas tanto desde los propósitos que se persiguen como desde las concepciones y metodologías en que se sustentan. Este hecho, es lo que justifica que Lesley Saunders (1999: 429) pueda encontrar respuestas a una de las preguntas con la que empieza su artículo en el que analiza para quién y para qué es la autoevaluación, en donde concretamente se interroga de «cómo y por qué es la autoevaluación escolar un 'animal' de una especie diferente según el hábitat donde se encuentre».

La evaluación del profesorado no sigue direcciones muy distintas. En este ámbito de evaluación, la rendición de cuentas se plantea fundamentalmente de cara a la habilitación y/o acreditación de los profesores para iniciar o proseguir su desempeño profesional, además de para su selección y el establecimiento de incentivos en su carrera profesional. Las características culturales de los países y de sus sistemas educativos son elementos que explican las modalidades de evaluación.

En los años 1993-94, la *Office of Educational Research and Improvement*, el *National Center for Education Statistics*, y el *Center for Research on Educational Accountability and Teacher Education* (CREATE) llevaron a cabo un estudio nacional en EE.UU. de los profesores de jardín de infancia hasta sexto grado de las escuelas públicas. En él, Shinkfield y Stufflebeam (1995: 25), se señala que el profesorado estaba de acuerdo con los siguientes puntos:

1. «La práctica de evaluar a los profesores de escuelas primarias está bien establecida en sus escuelas.
2. Sus evaluaciones están guiadas por normativas escritas.
3. Los profesores son informados de antemano de los criterios que se van a aplicar para evaluar su actuación.
4. La mayoría de los profesores son evaluados por el director de su escuela.
5. El método principal de evaluación utilizado es la observación de una clase.
6. Los profesores reciben respuestas verbales y escritas tras su evaluación.

7. Los profesores tienen la oportunidad de añadir sus alegaciones por escrito a la evaluación y/o presentar un recurso.
8. Los profesores apoyan las evaluaciones utilizadas para mejorar sus aptitudes docentes.
9. Los profesores ven los usos de las evaluaciones para despedir a profesores incompetentes o, especialmente, para conceder pagas por méritos como menos importantes que los usos de la evaluación para mejorar la enseñanza.
10. No obstante, los profesores indican que las evaluaciones deberían usarse más de lo que actualmente se hace para acabar con profesores incompetentes y determinar los niveles de pago de los profesores».

En España, las propuestas hechas para la valoración docente es, según se dice literalmente en los documentos ministeriales del año 1996 (MEC 1996a: 7),

«[un] requisito imprescindible hasta el momento para la acreditación de los candidatos a la dirección de centros docentes y para la obtención de una licencia por estudios y en el futuro lo irá siendo para acceder a otras vías de desarrollo profesional».

Este proceso es competencia de los servicios de inspección, quienes a través de informes obtenidos del propio profesor, el director, el jefe de estudios, el coordinador de ciclo (en educación infantil y primaria) o el jefe de departamento/seminario (en educación secundaria) —todos ellos elaborados siguiendo las pautas de unos cuestionarios— de entrevistas con cada una de estas personas, una observación en el aula en dos momentos distintos de la actividad docente y el análisis de documentación diversa, el inspector correspondiente procede a la redacción de un informe sobre la valoración de la labor profesional del profesor en el que se incluyen las puntuaciones obtenidas en cada una de las dimensiones y subdimensiones evaluadas. Los profesores pueden reclamar ante las instancias correspondientes en caso de desacuerdo con los resultados de dicha valoración.

B) Actuaciones en la educación superior

En Europa, en el campo de la educación superior, durante estos años se van a llevar a cabo dos estudios: el *Proyecto Piloto Europeo para la Evaluación de la Calidad de la Educación Superior* (European Commission,

1995) y el estudio complementario denominado *Análisis Comparativo de Iniciativas de Certificación y Evaluación de la Calidad en la Educación Superior* (European Commission. Sócrates, 1995). Ambos estudios marcan el inicio de la construcción de una cultura de la evaluación institucional europea de la que emergen planteamientos, propuestas metodológicas y la constitución de grupos de expertos y de redes que van a tener una influencia decisiva para la generalización de la evaluación de las universidades.

El primero de ellos se realizó en los 15 estados miembros de la Unión Europea, más Islandia y Noruega, entre los meses de noviembre de 1994 y junio de 1995. Participaron 46 establecimientos universitarios y no universitarios de los distintos países implicados en el proyecto, y se centró en el ámbito de la evaluación de la enseñanza, considerando de un modo especial el impacto que en ella tenían las actividades de investigación, en dos grandes áreas de disciplinas: la *ingeniería*, por una parte, y las *ciencias de la información y de la comunicación* o el *arte/diseño*, por otra. El objetivo del estudio, tal y como se manifiesta expresamente, no era el de establecer un ranking o clasificación de instituciones ni establecer un único modo de evaluación institucional, sino ir creando una cultura y un acervo común de experimentación de metodologías de evaluación. La metodología adoptada recogía los elementos comunes de las experiencias llevadas a cabo en los cuatro países más experimentados —Francia, Reino Unido, Países Bajos y Dinamarca—, planteándose que los procesos debían realizarse con independencia de los poderes políticos y de las propias universidades, así como que habrían de observar las siguientes fases: autoevaluación, evaluación externa por un grupo de pares —(grupo de expertos) con visitas *in situ*—, y publicación de un Informe de evaluación.

En el estudio *Análisis Comparativo de Iniciativas de Certificación y Evaluación de la Calidad en la Educación Superior*, se recogieron las experiencias gubernamentales y no gubernamentales existentes en Europa en el campo de la evaluación de la calidad de la educación superior, analizándose los objetivos y métodos utilizados en las mismas con el fin de conocer sus semejanzas y diferencias así como sus posibles complementariedades. Las conclusiones, partiendo del reconocimiento de las peculiaridades de las instituciones y, por tanto, de sus diferentes objetivos, trayectorias y acervo acumulado, resaltan que todos los participantes compartían que la evaluación debería seguir el siguiente proceso metodológico: autoevaluación preparada por las instituciones participantes o programas de estudio, revisión

de pares por parte de expertos en el campo evaluado y presentación de un informe. Dicho proceso metodológico se aplica en diferentes actividades, tanto si se trata de actividades de evaluación de la enseñanza y el aprendizaje como de los mecanismos establecidos en las instituciones.

También, por esta época, distintos organismos internacionales emprenden una serie de iniciativas y proyectos cuyo tema central va a ser la evaluación de la educación superior. Así, en 1992 la OCDE inicia el proyecto *Evaluación y proceso de toma de decisiones* con la participación de Francia, España y Alemania en el que se trataba de potenciar la evaluación como soporte de la toma de decisiones en pro de la mejora de la calidad de la enseñanza (OCDE, 1994). Mas tarde, en el año 1995, se lleva a cabo el proyecto *Gestión de Calidad, Evaluación de Calidad y proceso de toma de decisiones* —en el que participaron 27 instituciones de Australia, América del Norte, Europa Mediterránea, países nórdicos, noroeste de Europa y Europa Central—, que no es sino un estudio exploratorio de un conjunto de estudios de caso nacionales con el fin de conocer los objetivos, métodos y resultados de sus sistema de evaluación, así como de indagar sobre su influencia en la toma de decisiones.

Desde 1994, el Centro para la Investigación y la Innovación en la Enseñanza (CERI) de la OCDE, dentro de su programa IMHE (Institutional Management in Higher Education), y con la colaboración de la ACA (Asociación de Cooperación Académica) puso en marcha el proyecto piloto sobre el *Proceso de Revisión de la Calidad de la Internacionalización (IQRP)*, cuyos objetivos eran prestar asistencia a las instituciones de educación superior para evaluar y mejorar la calidad de su dimensión internacional, centrándose especialmente en: las políticas de mejora y estrategias de desarrollo institucionales para la internacionalización, la integración de la internacionalización en los proyectos y quehaceres institucionales y la consideración de la internacionalización como un componente de la evaluación de la calidad de las instituciones, (OCDE,1999). En cuanto a los procedimientos de evaluación, el IQRP establece, al igual que los proyectos europeos vistos con anterioridad, las fases de: autoevaluación, evaluación externa e informe de evaluación.

La UNESCO es otro de los organismos internacionales que ha asumido un protagonismo relevante en la elaboración de políticas y proyectos de evaluación institucional. A través del Centro Europeo para la Enseñanza

Superior (CEPES), la UNESCO ha venido desarrollando distintas actuaciones en la evaluación y acreditación de la educación superior en los países de Europa Central y Oriental. Y dentro del grupo de países de este ámbito geográfico creó el Grupo Europeo de Evaluación Académica (EGAA) para establecer marcos de intercambio y cooperación con los países de Europa Occidental, con el fin de: emprender estudios comparativos sobre procedimientos y criterios de evaluación de la educación superior; seleccionar y definir criterios de excelencia adecuados a las características de los sistemas europeos de educación superior; y avanzar en la aplicación de la convención para el reconocimiento de los estudios, diplomas y títulos existentes en la educación superior europea.

Tanto en el *Documento de Política para el Cambio y el Desarrollo en la Educación Superior* de 1995, como en los documentos de trabajo de la *Conferencia Mundial sobre la Educación Superior* celebrada en París en 1998, o en el número 107 del año 1998 de la revista *Perspectivas*, monográfico dedicado a la Educación Superior en el Siglo XXI, la UNESCO insiste una y otra vez en la importancia de la evaluación para el logro de una educación superior pertinente y de calidad. Por ello, uno de los desafíos más importantes para la educación superior en el siglo XXI será el de articular equilibradamente la cultura de la rendición de cuentas, mediante procesos de evaluación y regulación, con la cultura de autonomía responsable de las instituciones, en donde el elemento relevante será el mérito (UNESCO, 1998: 35).

Durante todos estos años se van creando organismos nacionales, en diferentes regiones del mundo, entre cuyos cometidos se recoge especialmente la evaluación de la calidad. Así, por ejemplo, en Europa Occidental: Francia dispone del *Comité Nacional de Evaluación de los establecimientos públicos con carácter científico, cultural y profesional* (CNE), desde el año 1984, Dinamarca cuenta desde 1992 con el *Centro de Evaluación* (Evalueringscenteret); el Reino Unido creó en 1991 el *Consejo de Calidad de la Educación Superior* (HEQC), y el *Consejo Nacional de Cualificaciones Profesionales* (NCVQ), y en España el *Consejo de Universidades* ha emprendido desde al año 1992 distintos programas de evaluación de la calidad de las universidades españolas. En América Latina contabilizamos, entre otros, los siguientes organismos: la *Comisión Nacional de Evaluación* (CONAEVA) de México y la *Comisión Nacional de Evaluación y Acreditación* (CONEAU) de Argentina, fundadas en 1992 y 1995 respectivamente; el

Consejo Superior de Educación de Chile que viene desarrollando sistemas de evaluación desde 1988; la *Coordinación del Perfeccionamiento del Personal de Nivel Superior* (CAPES) de Brasil, creado en 1951, y que es el encargado en la actualidad del Programa de Evaluación Institucional de las Universidades Brasileñas (PAIUB); etc.

Junto a estas organizaciones nacionales e internacionales, en el campo de la educación en general, y en el de la educación superior en particular, van a ir conformándose también todo un conjunto de redes cuya preocupación es el intercambio de conocimientos y experiencias sobre el tema de la evaluación de la calidad. Entre las primeras, por ser muy cercanas a nuestro contexto cultural, cabe resaltar las iniciativas emprendidas por la Organización de Estado Iberoamericanos que, a raíz de la Cumbre Iberoamericana de 1995, promueve el programa de Evaluación de la Calidad de la Educación en el que están implicados todos los Ministerios de Educación, <http://www.oei.es/linea7.htm>, y en los primeros meses del año 2000 nace la «Red Iberoamericana de Investigación sobre Eficacia Escolar y Mejora de la Escuela» (RIEME) con la finalidad de ayudar a incrementar la Calidad y Equidad de la Educación en Iberoamérica. Entre las de educación superior señalaremos, entre otras, las siguientes: la Red Internacional de Agencias de Certificación de la Calidad en la Educación Superior (INQAAHE), (Craft, 1992, 1994), cuyo origen se remonta a 1991, la Red Europea para el Aseguramiento de la Calidad en la Educación Superior que celebró su primera Asamblea General en Bruselas, en marzo de 2000; la Red Universitaria de la Evaluación de la Calidad (RUECA), fundada en 1995; y la Red de Evaluación Institucional de la Educación Superior (RAIES) creada en 1997.

3. A MODO DE RECAPITULACIÓN

En este tema hemos presentado la conformación de la evaluación de centros y profesores como un campo profesional de estudio e investigación. A tal efecto se ha analizado como dicho campo es fruto de la incidencia de factores exógenos, en concreto identificamos ciertas condiciones socioeconómicas y políticas que son el desencadenante de la emergencia de la *accountability* y del surgimiento del Estado evaluador, así como de fenómenos endógenos, es decir, de los debates y deliberaciones que en la comu-

nidad académica se llevan a cabo entre planteamientos vinculados con diferentes formas de concebir el conocimiento, con las opciones metodológicas posibles para generarlo y con los roles asumidos por los actores educativos en la misma evaluación.

Situados en un contexto caracterizado por la rendición de cuentas y por la diversidad de propuestas para acometer la evaluación institucional y de la docencia, en la segunda parte del tema se aborda la génesis del desarrollo y consolidación de estos dos nuevos ámbitos de evaluación hasta la actualidad.

BIBLIOGRAFÍA

- CARNOY, M. (1999). «Globalización y reestructuración de la educación», *Revista de educación*, n. 318, pp. 145-162.
- CEDODEP, (1966). Técnicas de valoración de personal docente. Organización y supervisión de escuelas. Madrid, CEDODEP.
- CLIFT, P.S.; NUTALL, D.L. y McCORMICK, R. (1989). *Studies in school self-evaluation*. Lewes, The Falmer Press, reprinted, 1st publ. 1987.
- CRAFT, A. (1992). *Quality Assurance in Higher Education. Proceedings of an International Conference Hong Kong, 1991*. London, The Falmer Press.
- (1994). *International Developments in Assuring Quality in Higher Education*. London, The Falmer Press.
- ELLIOTT, J. (1986). «Autoevaluación, desarrollo profesional y responsabilidad», en M. Galton y B. Moon, (Eds.): *Cambiar la escuela, cambiar el curriculum*, Barcelona, Martínez Roca, pp. 237-259.
- ESCUDERO ESCORZA, T. (1979). «En torno a la evaluación educativa en España», en INCIE (Ed.): *Temas de investigación educativa*. Madrid, MEC-INCIE, pp. 197-223.
- EUROPEAN COMMISSION (1995). *Projet Pilote Européen pour l' évaluation de la qualité dans l' enseignement supérieur*. Rapport Européen. Education, Training & Youth.

- EUROPEAN COMMISSION SÓCRATES (1995). *Initiatives of Quality and Assurance and Assessment of Higher Education in Europe*. Education Training & Youth.
- GARCÍA HOZ, V. (1975). «Una pauta para la evaluación de centros educativos», *Revista Española de Pedagogía*, (130), pp.117-150.
- INSPECCIÓN GENERAL DE EDUCACIÓN BÁSICA DEL ESTADO (1983). *Instrumentos de Supervisión*. Madrid, Ministerio de Educación y Ciencia.
- ISAACS, D. (1977). *Cómo evaluar los centros educativos. Instrumentos y procedimientos*. Pamplona, EUNSA.
- LÓPEZ RUPÉREZ, F. (1994). *La gestión de la calidad en educación*. Madrid, La Muralla.
- MARTÍN RODRÍGUEZ, E. (1988). «Profesión docente y autoevaluación institucional», *Revista de Educación*, n. 285, pp. 33-43.
- MACBEATH, J.; SCHRATZ, M.; MEURET, D., y JAKOBSEN, L.B. (2000). *Self-Evaluation in European Schools. A story of Change*. London, Routledge/Falmer.
- MACDONALD, B. y PARLETT, M. (1990). Repensar la evaluación: notas sobre la conferencia de Cambrigde», en M. J. Sáez Brezmes, y otros: *Evaluación de programas, proyectos e instituciones en educación*. Alcalá de Henares, Universidad de Alcalá, Instituto de Ciencias de la Educación.
- MCCORNICK, R. y JAMES, M. (1983). *Curriculum Evaluation in Schools*. London, Croom Helm (*Evaluación del currículum en los centros escolares*. Madrid, Morata, 1996).
- MEC (1976). *Ley General de Educación y Financiamiento de la Reforma educativa y disposiciones complementarias*. Madrid, Ministerio de Educación y Ciencia, Colección Compilaciones.
- (1992). *Plan de Evaluación de Centros Docentes, niveles no universitarios. Programa Piloto para el curso 1991-92*. Madrid, MEC.
- (1994). *Centros Educativos y Calidad de la Enseñanza. Propuesta de actuación*. Madrid, Ministerio de Educación y Ciencia.
- (1996a). *Valoración de la labor profesional de los docentes. Función docente, educación infantil y primaria*. Madrid, Ministerio de Educación y Ciencia.
- MUÑOZ ARROYO, A. (1979). «Evaluación de Centros de Educación General Básica», *Boletín de Inspección Técnica de Educación*, año 1, n. 0, pp. 7-26.

- NEAVE, G. y VAN VUGHT, F. A. (1994). *Prometeo encadenado. Estado y educación superior en Europa*. Barcelona, Gedisa.
- NEVO, D. (1997). *Evaluación basada en el centro: un diálogo para la mejora educativa*. Bilbao, Mensajero.
- OCDE (1994). *Evaluation et processus de décision dans l'enseignement supérieur: Expériences en Allemagne, Espagne et France*. Paris, OCDE, Serie Documentos.
- (1999). *Qualité et internationalisation de l'enseignement supérieur*. Paris, OCDE-IMHE.
- PÉREZ JUSTE, R.; LÓPEZ RÚPEREZ, F.; PERALTA, M.^a D. y MUNICIO, P. (2000). *Hacia una educación de calidad. Gestión, Instrumentos y Evaluación*. Madrid, Narcea.
- RODRÍGUEZ DIÉGUEZ, J.L. (1973). *La función de control en la Educación*. Madrid, CSIC (Consejo Superior de Investigaciones Científicas).
- SANDERS, W. y McLEAN, R. (1995). «The Tennessee Value-Added Assessment System (TVAAS): Mixel Model Methodology in Educational Assessment», in Shinkfield, A. and Stufflebeam, D. L. (Eds.): *Teacher Evaluation...* op. cit., pp.337-350.
- SAUNDERS, L. (1999). «Who or What is School 'Self-Evaluation for?», *School Effectiveness and School Improvement*, v. 10, n. 4, pp. 414 – 429.
- SHINKFIELD, A. y STUFFLEBEAM, D. L. (1995). *Teacher Evaluation: Guide to Effective Practice*. Boston, Kluwer Academic Publisher.
- SIMONS, H. (1985). «Against the Rules Procedural Problems in Self-evaluation», *Curriculum Perspectives*, vol. 5. n. 2, October, pp. 1-6.
- (1999). *La evaluación democrática de las instituciones escolares*. Madrid, Ediciones Morata.
- SITE (1976). «Programa de evaluación de la calidad de la enseñanza», *Vida Escolar*, n. 177-78, marzo-abril.
- STAKE, R. E. (1983). «Evaluación de programas: en especial de Evaluación de Réplica», en W. B. Dockrell, y D. Hamilton, (Eds.): *Nuevas Reflexiones sobre la Investigación Educativa*. Madrid, Narcea, pp. 40-52.
- UNESCO (1995). *Documentos de Política para el cambio y el desarrollo en la Educación Superior*. París, UNESCO.

— (1998). *Conferencia Mundial sobre la Educación Superior en el Siglo XXI. Visión y Acción*. París, UNESCO.

Tema 2

Concepciones, vínculos, significados y otras cuestiones estructurantes de la evaluación de centros y profesores

Introducción

1. En torno al concepto de evaluación
 - 1.1. «Evaluation» versus «assessment»
 - 1.2. Algunas definiciones de evaluación
2. Acerca de la relación entre evaluación, investigación e investigación evaluativa
3. La evaluación en las distintas teorías y tradiciones científicas
 - 3.1. Tradición «científico-racional»
 - 3.2. La tradición Interpretativo-Simbólica
 - 3.3. La tradición sociocrítica
 - 3.4. Posiciones postparadigmáticas
4. Los indicadores de evaluación
 - a) Características técnicas que deben poseer los indicadores educativos
 - b) Tipos de Indicadores
 - c) Modelos de Indicadores / Modelos de Escolaridad
 - d) Indicadores y Toma de Decisiones
5. La política y ética de la evaluación
6. A modo de recapitulación

Bibliografía

TEMA 2

CONCEPCIONES, VÍNCULOS, SIGNIFICADOS Y OTRAS CUESTIONES ESTRUCTURANTES DE LA EVALUACIÓN DE CENTROS Y PROFESORES

José Félix Ángulo Rasco

Universidad de Cádiz

Eustaquio Martín Rodríguez

UNED

INTRODUCCIÓN

Como el Diccionario de la Lengua recoge, *evaluar* consiste en estimar, apreciar, calcular el valor de una cosa. Sin duda, la evaluación es una actividad constitutiva de nuestra conducta como seres humanos. Normalmente, nuestros cursos de acción están influidos por las decisiones que tomamos como consecuencia de la comparación y/o valoración de la información disponible. Sin embargo, este tipo de actuaciones se suelen hacer de manera natural e informal, es decir, considerando racionalmente los elementos con los que se cuenta en un momento dado, y, además, en un ámbito privado, esto es, sólo conciernen a nuestros problemas personales.

Cuando la evaluación se aborda de modo sistemático y en un contexto que trasciende la esfera privada e individual del comportamiento humano, entramos de lleno en su consideración desde una perspectiva formal y en el terreno propio de las dinámicas sociales, es decir, el público. Este es el campo en el que toma cuerpo la evaluación como un tipo de estudio, indagación o investigación que se lleva a cabo en un contexto social determinado.

En este tema se hace un recorrido por algunas de las cuestiones más importantes y sustantivas de la evaluación. Pretendemos, primeramente, desentrañar en toda su complejidad el concepto de evaluación presentando sus distintas acepciones, así como la evolución que el mismo ha ido teniendo en el transcurso del tiempo. Las perspectivas conceptuales y las diversas definiciones proyectan, como tendrá oportunidad de verse, distintas maneras de abordar la evaluación, en general, y la de centros y profesores, en particular.

Posteriormente se analizan las relaciones entre evaluación, investigación e investigación evaluativa para tratar de determinar la entidad de la evaluación, los distintos elementos desde los que se la vincula con la investigación y sus señas de identidad. La denominación de investigación evaluativa, propuesta y utilizada por importantes evaluadores en sus escritos y

obras, indica bien a las claras los aspectos comunes que ambos procesos de indagación comparten, pero sin duda existen aspectos diferenciales como los propósitos, los contextos de relaciones en que se sitúan los evaluadores e investigadores y los ritmos y tiempos con los que se trabaja en la entrega o devolución de la información.

En este marco en el que se explicitan los vínculos entre evaluación e investigación, en lo que a la adopción de epistemologías realista e idealista se refiere, se entra a plantear también sus posiciones en las tradiciones científicas científico-racional, interpretativo simbólica y sociocrítica. Desde ellas se trata de poner de manifiesto que la opción por unas u otras determina no sólo el tipo de conocimiento que la evaluación aporta, sino también las metodologías puestas en juego para generarlo, así como desde qué valores y al servicio de quién se produce.

La última parte del tema se dedica, por un lado, a presentar el uso de los indicadores de evaluación por la creciente importancia que han ido adquiriendo, así como por ser una cuestión totalmente pertinente con las problemáticas epistemológica y las tradiciones científicas planteadas. Y, por otro lado, se desarrollan las dimensiones política y ética en tanto que referencias ineludibles, como decíamos antes, de al servicio de quién se realiza el estudio y de los principios y valores que tácita o explícitamente están presentes en todo proyecto de evaluación.

1. EN TORNO AL CONCEPTO DE EVALUACIÓN

En castellano es frecuente utilizar el término evaluación para toda una gama de procesos: evaluación curricular, de materiales, de centros, de alumnos y de profesores, e, incluso, de lo que en psicología se denomina «evaluación conductual». Además cuando se hace referencia al alumnado, encontramos otros términos asociados al mismo, utilizados en la misma práctica, como los de «control», «pruebas», «calificación», «rendimiento» y «exámenes».

Dados los distintos tipos de usos, que en general se vinculan a diferentes ámbitos, y las diversas asociaciones que se hacen de la evaluación con otros muchos términos, deberíamos revisar si existe más de un sentido atribuible al mismo.

1.1. «Evaluation» versus «assessment»

Los ingleses hacen referencia a los términos *assessment* y *evaluation* para referirse a dos formas distintas de entender la evaluación.

Por un lado, la palabra *assessment*, que podría traducirse por tasación, está directamente vinculada en la cultura de la evaluación a la tradición psicométrica de utilización de tests y procedimientos correlacionales de tratamiento de datos. Scriven (1991: 60) identifica el «assessment» con los juicios a través de datos numéricos, como procesos que se centran en este tipo de datos y menos en los enfoques de juicios valorativos, precisando que:

«... normalmente se utiliza (el término *assessment*) como sinónimo de evaluación, pero para quienes el producto es más importante que el proceso hacen esfuerzos denodados para diferenciarlos, quizá para evitar las connotaciones peyorativas que para ellos tiene el término evaluación. Ninguno de estos esfuerzos tiene mucho valor, ya que al carecer de lógica interna resultan inaceptables».

Sobre la base de la distinción que hace MacDonald (1976) entre «evaluation» y «assessment», en la que reserva a la primera el papel de la formulación de afirmaciones o valoración de un servicio educativo, mientras que identifica a la segunda con las valoraciones que se hacen no sobre el servicio sino sobre los logros de los receptores del mismo, Angulo (1990: 115) concluye que:

«... en este sentido, es claro que “assessment” está relacionado con la “medición” de los aprendizajes alcanzados por los alumnos en relación a los objetivos propuestos, pero este proceso, sea o no necesario, no es nunca una evaluación. En todo caso esta información puede ser una parte, no necesariamente privilegiada, de la información total que puede manejarse en una evaluación».

No obstante, la clarificación de este aspecto se facilita si situamos a ambos conceptos en dos planos de significación interrelacionados: uno con respecto al sentido y otro con respecto al objeto de ambos conceptos. Según el primer plano —el sentido—, «evaluation» hace referencia a todo proceso orientado al conocimiento de la *calidad del servicio educativo prestado*, mientras que «assessment» lo haría con respecto al «impacto» que dicho servicio ha tenido sobre los *receptores*. Según el segundo, la distinción se

sitúa en el «objeto» del proceso: la evaluación se ocupa de la influencia en la calidad educativa que puedan tener los distintos y diversos componentes de los procesos de enseñanza/aprendizaje (currículum, metodología, experiencias de aprendizaje, organización de los centros de enseñanza, administración educativa —provincial, autonómica, estatal—, los docentes y los mismos estudiantes); mientras que «assessment» se orientaría única y exclusivamente a los individuos o grupos de individuos receptores del servicio (estudiantes).

Por tanto, podemos acordar desde estas diferencias dos acepciones del término evaluación: una **acepción amplia** que correspondería al proceso por el cual conocemos y valoramos la calidad del servicio y el papel de los distintos componentes en el mismo y otra **acepción estrecha** o más restringida que señalaría el proceso o procedimiento a través del cual averiguamos su «calidad» únicamente a través del impacto de dicho servicio en individuos o grupos de individuos receptores del mismo, restringiendo la comprensión de calidad del servicio educativo a su incidencia sobre los receptores (alumnos). Hay que recordar, además, que, la acepción estrecha de evaluación, subyace a algunos términos usados cotidianamente como los de exámenes y controles, y a otros más técnicos como los de pruebas nacionales, indicadores de calidad educativa, mediciones de rendimiento, sistemas de tests, etc.

1.2. Algunas definiciones de evaluación

En la literatura profesional pueden encontrarse innumerables definiciones y descripciones de qué es y en qué consiste la evaluación. En todas ellas subyacen concepciones que, como ocurre habitualmente en los campos de las ciencias sociales, no están exentas de valores, de referencias a unas u otras formas de conocimiento y a distintas opciones metodológicas. A lo largo de su historia, los significados de la evaluación han ido cambiando progresivamente, emergiendo nuevas propuestas y enfoques que, en líneas generales, han dado lugar a una situación caracterizada por un pluralismo conceptual y por la aceptación de métodos, criterios e intereses múltiples. Nevo (1990: 2587) propone una clasificación de las definiciones de evaluación en tres grupos diferenciados: las que se basan en metas, las descriptivas no enjuiciables y las enjuiciables.

El referente fundamental de las del primer grupo es la definición de Ralph Tyler (1950: 69), para quien la evaluación es «el proceso de determinar hasta qué punto se están realizando los objetivos educativos».

La aparición de la figura de Ralph Tyler supuso un avance en los planteamientos que van a ir configurando el campo de la evaluación durante las décadas de los años treinta y cuarenta, o *época tyleriana*, aún cuando su influencia se proyectó hasta bien pasados los años sesenta. Un primer aspecto que hay que señalar es que sus aportaciones no pierden nunca de vista la dimensión educativa de la relación entre currículo y evaluación. La definición y delimitación de los objetivos que debían alcanzarse tenían que considerar el currículo que se iba a llevar a cabo, así como el rendimiento del alumnado determinado por éste. La tecnología de los tests de desarrollo sería la base sobre la que se construirían los tests referidos a los objetivos a alcanzar, con los relacionados con las diferencias individuales y con las normas nacionales o del estado. En su investigación sobre *Eight-Year Study* (Smith y Tyler 1942), en la que se puso a prueba la eficacia del desarrollo de los currículos y estrategias didácticas que estaban realizándose en un centenar de escuelas americanas, tuvo la oportunidad de experimentar y difundir sus concepciones de evaluación educativa.

En dicho estudio, a diferencia de la tradición de los experimentos psicológicos con grupos de control y experimentales, Tyler procede a la comparación interna entre objetivos y resultados para verificar los logros obtenidos por los estudiantes. De esta manera el grado de éxito se mide directamente por las puntuaciones obtenidas en los tests específicamente contruidos a tal efecto, en vez de a través de tests estandarizados o mediante indicadores indirectos como el equipamiento y materiales de los que disponen los centros o la participación de la comunidad en las instituciones.

Las definiciones descriptivas no enjuiciables son todas aquellas que hacen énfasis en la evaluación en su variedad de proceso que informa la toma de decisiones, como por ejemplo la de Norris (1990: 101) que explícitamente habla de un «proceso cuyo propósito es concebir, obtener y dar información para permitir la búsqueda de soluciones (toma de decisiones) a cuantos están implicados y/o tienen alguna responsabilidad sobre el tema evaluado».

Una parte muy importante de las definiciones de evaluación que se encuadran en este segundo grupo, insisten en que la evaluación es una fuente de información que sirve para fundamentar la toma de decisiones, y que, asimismo, éstas pueden vincularse a la valoración terminal que se hace una vez concluida la acción o mientras se está llevando a cabo la misma.

Junto a esta orientación de toma de decisiones, ha ido abriéndose paso también la función de la evaluación como mecanismo que contribuye a ilustrar, iluminar o a mejorar el conocimiento de las dinámicas educativas y/o formativas. Desde luego, parece claro que ésta es una de las finalidades primarias de la evaluación, tanto porque en sí misma la evaluación es una forma de conocimiento, como porque cuando se incorpora a los quehaceres institucionales se hace para disponer de una visión sistemática que de otro modo no es posible tener. Otra cosa distinta es lo que posteriormente se haga con esta información. Pero, en cualquier caso, difícilmente podrán tomarse decisiones si, por una parte, no existe un conocimiento informado de lo que está ocurriendo en los ámbitos formativos de que se trate, y si, por otra, los distintos actores de la formación no se sensibilizan o toman conciencia, precisamente a través de la evaluación, de los logros y carencias de sus quehaceres profesionales e institucionales.

Las del tercer grupo, es decir las enjuiciables, se caracterizan por considerar que la evaluación implica valoración; el Joint Committee on Standards for Educational Evaluation (1988: 166) acota que es una «Investigación sistemática del valor o mérito de un objeto, por ejemplo, un programa, proyecto o material de enseñanza». En este mismo grupo se encuadran también las posiciones de aquellos autores, como en el caso de Guba y Lincoln (1989) que, aún sin dar una definición de evaluación, defienden la combinación de la descripción y el juicio en sus propuestas. Un ejemplo, de esta última posición, se recoge en la definición del Centro Australiano de Desarrollo Curricular para los que la evaluación es el «proceso de obtener, delimitar y aportar información útil para tomar decisiones y establecer juicios de valor sobre los programas de estudios y el currículum» (citado por Kemmis (1986: 117).

La evaluación como *juicio* surge también como resultado de las flaquezas de la evaluación basada en objetivos, debilidades que se pusieron de manifiesto a finales de los años 50. En este momento se ve la necesidad de implementar proyectos en los que se cuestiona que se pueda trabajar con

objetivos previos, programas experimentales y utilizando estrategias de evaluación aplicadas al finalizar los programas, porque en ese momento si la evaluación muestra deficiencias es demasiado tarde para mejorarlas. Estos problemas fueron documentados en la evaluación entendida como mejora (Cronbach, 1963, *Course Improvement Through Evaluation*).

El promotor de este nuevo rumbo que emprende la evaluación es Robert Stake y la clave es el juicio. Lo que tiene que hacer la evaluación es enjuiciar. Para él la descripción y el juicio son los actos básicos de la evaluación. Otros autores, como por ejemplo Scriven, profundizan mucho en este tipo de orientación y defienden que hay que trabajar con objetivos que se consideren en sí mismo como problemáticos, tanto las metas como las actuaciones tienen que estar sujetas a evaluación, se requieren estándares con los que hacer el juicio. Sin embargo, la inclusión de patrones de juicio repugnaba a muchos evaluadores que pensaban que siguiendo las pautas de evaluación científica la evaluación tenía que ser una empresa libre de valores. No obstante, las circunstancias sociales les obligaron a asumir el tema del juicio, aunque no se sentían con competencia. Después de 1967 ha quedado claro que todos los modelos, aunque con diferencias, coinciden en que el juicio es una parte clave de la evaluación. En todos varía como se aborda el juicio, desde aquéllos en los que únicamente es competencia del evaluador, como en el de crítica artística, a los que negocian los patrones de juicio con los promotores o con las audiencias.

La clasificación presentada recoge, de alguna manera, el devenir histórico de la evaluación. Las definiciones ya sean con respecto a metas, las descriptivas o las enjuiciables proyectan como se ha ido configurando la evaluación como campo de estudio y práctica profesional. Pero esta tipología de definiciones, como la mayor parte de las clasificaciones que se hacen, aun suponiendo un esfuerzo loable de sistematización, resulta engañosa o cuanto menos parcial y limitada. Porque es cierto que la evaluación ha sido, y sigue siendo, por la coexistencia de enfoques diversos, un modo de determinar la adecuación entre objetivos y resultados, una forma de conocimiento informado sobre el que sustentar la toma de decisiones, o un juicio de valor sobre la naturaleza y/o características de algo. Sin embargo, no es menos verdad que la evaluación es todo eso y algo más, y que las clasificaciones, por su nivel de generalidad, admiten e integran, en un mismo grupo, definiciones que suponen una teoría y práctica de la evaluación totalmente diferentes.

2. ACERCA DE LA RELACIÓN ENTRE EVALUACIÓN, INVESTIGACIÓN E INVESTIGACIÓN EVALUATIVA

La evaluación mantiene una estrecha vinculación con la investigación en la medida en que se ha beneficiado del patrimonio acumulado por ésta, en todo lo relativo a sus avances y logros en metodologías, procedimientos, técnicas, instrumentos, etc. La modalidad pionera de evaluación de programas, en su aún corta historia de un siglo, no tuvo que empezar de cero, sino que inició su andadura prolijada por la investigación hasta el punto de que, en no pocos momentos, se han identificado y asimilado ambos procesos de indagación. Es por ello, fundamentalmente, por lo que resulta complicado tratar el tema de la evaluación sin hacer referencias permanentes a la investigación.

Normalmente asociamos la investigación a la generación del conocimiento y al desarrollo de la teoría científica. Sin duda, los distintos campos del saber se consolidan y avanzan conforme se van desentrañando y conociendo los principios y causas de su objeto de estudio. Cuando los investigadores se preocupan fundamentalmente por ampliar el conocimiento científico, al margen de su posible aplicación, se suele hablar de investigación básica; y, por el contrario, cuando lo que se prioriza es la búsqueda de soluciones a problemas o temas particulares y/o prácticos nos estamos refiriendo a la investigación aplicada. En uno y otro caso, la investigación es una herramienta de la que las ciencias se sirven para ir elaborando un conjunto organizado y sistemático de conocimientos sobre sus respectivos objetos de estudio.

Las similitudes y diferencias entre evaluación e investigación ha sido uno de los temas que, durante un cierto tiempo, ha centrado la atención de casi todos los autores y obras de evaluación. A la confusión entre ambos tipos de actividades no fue ajena el empleo casi generalizado de la expresión *evaluation research* o investigación evaluativa que, como tal denominación, encabezó los títulos de algunas de las publicaciones de los pioneros de la evaluación. Las controversias, a este respecto, surgieron en torno a si la evaluación puede considerarse como una forma o modalidad de investigación o si, en cambio, ésta es una herramienta que la evaluación con su propia identidad, distinta a la investigación, utiliza en sus estudios.

Para Schuman (en Stufflebeam y Shinkfield, 1987: 113), no existen dudas de que

«la evaluación es un proceso científico... [por lo que conceptualiza la investigación evaluativa] como “la primera y principal *investigación*, que debe acercarse cuanto le sea posible a las normas comúnmente aceptadas de la metodología investigativa”... [y puntualiza que] la investigación evaluativa es investigación aplicada, y su propósito es determinar hasta qué punto un programa específico ha conseguido el resultado deseado».

La consideración de la evaluación como un proceso científico y de investigación de carácter aplicado es, sin duda, una de las tradiciones mantenidas desde la **epistemología objetivista o realista** que defiende, en los campos pedagógico, psicológico y sociológico, que la ciencia tiene por finalidad el descubrimiento de verdades universales, ajenas a los contextos sociales de producción. Las verdades que han de formularse en leyes, con capacidad para predecir y regular los fenómenos estudiados, deben conducir a una ciencia definitiva y objetiva y, en última instancia, a un sistema de leyes aplicables universalmente. Se encuadra dentro de la tradición científico-racional, conocida también como empírico-racional, positivista, hipotética-deductiva, conductista, etc. Se caracteriza por su defensa a ultranza de la «objetividad» científica y por mantener una concepción de la ciencia y del conocimiento científico rigurosamente formal. Separa los hechos de los valores, manteniendo una visión universalista de la ciencia al margen de las implicaciones sociales y contextos humanos en que ésta se produce.

La investigación se centra en todo aquel tipo de fenómenos observables, medibles y cuantificables que pueden ser abordados a través del método hipotético-deductivo, experimental, que mediante diseños preordenados trata de establecer relaciones de asociación y/o causalidad entre las variables objeto de estudio. Ello exige la categorización de las variables discretas que se seleccionen, la cuantificación de las variables continuas y el registro y medición de unas y otras para su tratamiento estadístico. Son, por tanto, investigaciones de carácter empírico que utilizan muestras representativas de la población con la que se trabaja de cara a poder generalizar los resultados obtenidos.

La caracterización que se hace de la evaluación desde la investigación aplicada, en contraposición con la investigación básica, se refiere sobre todo a la limitación de la generalización de los resultados, por cuanto su cometido son los estudios particulares o idiosincrásicos que o bien se derivan de las posibles teorías y leyes formuladas a través de la investigación básica o bien han de ser contempladas por ésta de cara a universalizar los

hallazgos. Las metodologías no van a ser, por el contrario, distintas en uno u otro tipo de investigación.

Esta forma de conceptualizar y entender la investigación aplicada, en tanto que **idiosincrásica**, dista mucho de ser la mantenida por autores que utilizan metodologías más cualitativas y que defienden nuevos acercamientos al campo de la investigación como, por ejemplo, la investigación-acción, la democrática, la descriptiva, la ilustrativa, la fenomenológica, etc. (Walker, 1989: 211) manifiesta que, en el campo de la educación, la investigación debería:

«... interactuar con la experiencia y la práctica, con vistas a ampliar el alcance y la aplicación de innovaciones, aumentar la eficacia y centrar su atención en la relación entre diseño y rendimiento. Esto no supone en modo alguno una defensa de la investigación aplicada como alternativa a la investigación pura o convencional. Existe una necesaria relación entre las investigaciones aplicada y científica, dado que ambas deben estar en mutua dependencia».

Desde estas posiciones se hacen más visibles las similitudes entre evaluación e investigación aplicada y, especialmente, cuando la experiencia y la práctica no sólo son los motivos desde los que se generan los estudios, sino también los referentes más inmediatos que los investigadores han de considerar de los «prácticos» que participan en un proceso de indagación. No obstante, el investigador a diferencia del evaluador se refugia en una comunidad académica, quedando a salvo tanto de las presiones de las audiencias como de la inmediatez de la información. Aunque, como dice Walker (1983: 48-49),

«[la] diferencia entre investigación y evaluación es esencialmente de grado porque el investigador, al optar por responder a los intereses y necesidades de un grupo en vez de otro, penetra también en un entramado de relaciones políticas, quizá menos acuciantes pero tan reales como las que rodean al evaluador».

Llegados a este punto, y retomando el empleo de la expresión investigación evaluativa podría entenderse que su origen está vinculado a la relación entre investigación aplicada y evaluación que, desde posiciones paradigmáticas positivistas, tienen un mismo significado, mientras que desde los planteamientos de la epistemología idealista, a pesar de sus similitudes, no son totalmente identificables. Es más, para algunos autores, la evaluación no puede considerarse una modalidad de la investigación sino

que habría que entender ésta como una rama de aquélla. McDonald (1985: 473) tiene, en este sentido, una posición muy definida al advertir que

«[el peligro] de conceptualizar la evaluación como una rama de la investigación es que los evaluadores llegan a estar atrapados en los tentáculos restrictivos de la respetabilidad de la investigación... Cuanto más productivo sería el hecho de determinar la investigación como una rama de la evaluación, una rama cuya tarea es solucionar los problemas tecnológicos encontrados por el evaluador».

A diferencia de la anterior, la epistemología idealista o subjetivista no considera la existencia de un conocimiento «objetivo» y generalizable para todas las circunstancias y situaciones humanas. Las peculiaridades de los contextos y la singularidad de las dinámicas humanas y sociales hacen que no pueda hablarse de soluciones objetivas y universalmente válidas. El conocimiento científico está condicionado social e históricamente por lo que es necesario ensayos y propuestas que partan de la realidad concreta que se quiera estudiar y que ésta a su vez sea considerada como algo dinámico y en devenir, además de cargada de valores, opciones ideológicas, símbolos, tradiciones, etc.

El conocimiento de los procesos y mediaciones que se dan en todos los fenómenos sociales y humanos demanda un tipo de metodología alternativa a la usada por el modelo hipotético-deductivo, capaz de fijarse preferentemente en las interacciones, elementos procesuales y dinámicas que conducen a determinados productos y resultados. Dicha metodología prima la participación de los actores de la educación en el proceso de investigación, ya sea en su consideración de audiencias concernidas, como en el caso de la investigación democrática, o asumiendo un mayor protagonismo como desde la conceptualización del profesor como investigador promovida por la investigación acción (Stenhouse, 1984).

Otro aspecto fundamental, en el que también existen diferencias importantes entre unas y otras orientaciones, son los criterios de **validez** de las evaluaciones. Para House (1994: 231) la validez consiste en la «cualidad de merecer el reconocimiento». Para él, en la evaluación de un servicio público, los estudios no sólo tienen que ser veraces y creíbles, como en los casos de las evaluaciones personales o interpersonales, sino también correctos desde un punto de vista normativo, es decir, ajustados a derecho. Los enfoques objetivistas basan sus criterios de validez en la utilización de

métodos explícitos que permiten la replicabilidad de los estudios, de manera que distintos investigadores puedan llegar a resultados similares. La **fiabilidad** de la medida es otro concepto clave, en la medida que permite el contraste y la concordancia de los instrumentos. Los enfoques subjetivistas, en cambio, apelan a la experiencia y no al método científico para fundamentar sus criterios de validez. La correspondencia entre las experiencias de los participantes y los resultados de la evaluación es el criterio básico de validez.

Estas visiones tan diferentes traslucen las divergencias que mantienen estas tradiciones de investigación a la hora de conceptualizar y utilizar criterios de validez en sus respectivos estudios. Los estudios de corte «cientifista» trasladan a la investigación aplicada, que suelen equiparar con la evaluación, los principios y procedimientos que han servido de base para formalizar la investigación básica de carácter experimental. Y, en este sentido, la única nota distintiva de la evaluación es su carácter local o idiosincrásico con respecto a la investigación básica que asume sus estudios desde una perspectiva de universalidad; pero no existen diferencias de ningún tipo en cuanto a las metodologías y procedimientos utilizados. La investigación naturalista, interpretativa o cualitativa se encuentra más cercana a las características y condiciones desde las que se llevan a cabo los estudios evaluativos. Los diseños de los estudios, el papel de los investigadores, las metodologías y procedimientos empleados, la participación de los implicados, etc., no son diferentes a los problemas y desafíos que acometen las evaluaciones; a excepción, quizá, de que las audiencias trascienden los intereses de los profesores e investigadores, de los juegos políticos en los que se ven inmersos los evaluadores y de la inmediatez de información y resultados que se demanda normalmente a la evaluación.

3. LA EVALUACIÓN EN LAS DISTINTAS TEORÍAS Y TRADICIONES CIENTÍFICAS

La concepción del conocimiento científico, la relación entre teoría y práctica, la separación o no de hechos y valores, la posición ante la investigación, los métodos y técnicas de investigación, la función de teorías y modelos, etc., son algunos de los aspectos diferenciales, entre otros, que definen y avalan a las distintas tradiciones científicas. La función y cometido de la evaluación en cada una de éstas varían sustancialmente.

3.1. Tradición «científico-racional»

También se la conoce como empírico-racional, positivista, hipotética deductiva, conductista, etc. Se caracteriza por su defensa a ultranza de la «objetividad» científica y por mantener una concepción de la ciencia y del conocimiento científico rigurosamente formal. Separa los hechos de los valores, manteniendo una visión universalista de la ciencia al margen de las implicaciones sociales y contextos humanos en que ésta se produce.

El investigador debe preocuparse por encontrar las relaciones causa efecto que se dan entre los fenómenos humanos, tratando de descubrir los principios y reglas que explican la conducta de los individuos y organizaciones, así como la dinámica de los fenómenos y acontecimientos. Una vez conocidas estas regularidades científicas podrán hacerse las predicciones oportunas y se podrá conocer objetivamente como afrontar los mismos problemas en parecidas situaciones.

De acuerdo con Proppe (1990: 327), esta tradición recoge sus fundamentos del positivismo como

«... concepción filosófica de la ciencia, la realidad y la vida: En su forma más pura, considera la ciencia como el único camino válido del conocimiento y las *realidades* como los únicos objetos posibles de conocimiento. Se opone a cualquier tipo de metafísica y, en general, a cualquier tipo de investigación que no pueda ser reducido a un método científico. La concepción positivista sólo reconoce, fundamentalmente, dos enfoques legítimos: el empírico (como la ciencia natural) y el de las disciplinas formales (como la lógica y las matemáticas). Da mucha importancia al empirismo inductivo (recoger datos y tratarlos estadísticamente) y al análisis deductivo (la construcción de sistemas de explicaciones hipotéticas y nomológicas)».

Las teorías y modelos científicos surgen de los problemas y temas que generan la reflexión, intuición y experiencia del investigador que, una vez formalizados, tratará de que se experimenten en la realidad. La práctica confirmará en parte o totalmente las regularidades científicas propuestas por investigadores y expertos, pero aquella será realizada por personas cuya función principal es la actuación en la realidad. Existe, pues, una separación funcional y real, dada la cualificación necesaria para investigar y planificar, entre investigadores, expertos y practicantes.

El objeto del conocimiento de la tradición científico-racional es todo aquel tipo de fenómenos observables, medibles y cuantificables, que pue-

den ser abordados a través del método hipotético-deductivo, experimental, propio de las ciencias físico-naturales o ciencias fuertes de las que esta tradición toma sus fundamentos y metodología de investigación para las ciencias sociales.

3.2. La tradición interpretativo-simbólica

A diferencia de la anterior, la tradición interpretativo-simbólica no considera la existencia de un conocimiento «objetivo» y generalizable para todas las circunstancias y situaciones humanas. Las peculiaridades de los contextos y la singularidad de las dinámicas humanas y sociales hace que no pueda hablarse de soluciones objetivas y universalmente válidas. El conocimiento científico está condicionado social e históricamente por lo que es necesario ensayos y propuestas que partan de la realidad concreta que se quiera estudiar y que ésta a su vez sea considerada como algo dinámico y en devenir, además de cargada de valores, opciones ideológicas, símbolos, tradiciones, etc.

Al decir de Proppe (1990: 328),

«[el] paradigma relativista, en cambio, está relacionado con el interpretacionismo, esto es, con la hermenéutica y algunos aspectos de la fenomenología (transcendental o ideal). La investigación interpretativa trata de saber qué sucede, o ha sucedido, y qué significa, o ha significado, para diferentes individuos o grupos; estudia los procesos, pero se mantiene “leal” al dualismo de tema y objeto. Su principal objetivo es la descripción pura y completa. La ciencia interpretativa puede ayudarnos a encontrar significados más profundos y a ser conscientes de una amplia gama de reglas culturales y distorsiones ideológicas que afectan a la comprensión por el hombre de la realidad social».

La presencia de elementos totalizadores en los temas y problemas de investigación antes que un obstáculo es uno de los ingredientes, y de los más importantes, para poder comprender y explicarse los procesos que intervienen en la configuración de la realidad. Y ésta no se limita ni se agota con la aprehensión de los fenómenos observables que puedan medirse y cuantificarse directamente.

El conocimiento de los procesos y mediaciones que se dan en todos los fenómenos sociales y humanos demanda un tipo de metodología alternati-

va a la usada por el modelo hipotético-deductivo, capaz de fijarse preferentemente en las interacciones, elementos procesuales y dinámicas que conducen a determinados productos y resultados.

En esta perspectiva la racionalidad no es inmanente a las organizaciones como sostienen las teorías positivistas, ni tan siquiera un elemento de juicio en base al cual se juzga el comportamiento de sus miembros. Para estas teorías la racionalidad no es un concepto unívoco puesto que está sujeto al paradigma cultural que lo interpreta. Así, las definiciones de lo que es «normal», «racional», «real», siempre están enmarcadas dentro de un paradigma construido dentro de la organización que es a su vez objeto de resistencias y superación.

3.3. La tradición sociocrítica

La tradición sociocrítica comparte con la interpretativo-simbólica el cuestionamiento de la concepción de la ciencia mantenida por la tradición científico-racional. Pero se diferencia de aquélla, a su vez, por su visión sociopolítica de la educación y el papel que asigna a ésta como mecanismo de liberación y emancipación de los seres humanos dentro de una dinámica socioescolar colectiva y solidaria.

Para la tradición sociocrítica, el conocimiento científico, en el que se entremezclan hechos y valores, opciones personales, ideológicas y científicas, no puede limitarse a la comprensión y explicación de los procesos personales, culturales y simbólicos sino que además desde una perspectiva crítica y de toma de conciencia política ha de contribuir a la creación y construcción de una sociedad no reproductora de las desigualdades, hándicaps y manipulación política características de la actual.

La investigación surge directamente de la práctica y son los mismos practicantes quienes abordan su realización a través de metodologías esencialmente cualitativas y sin la diferenciación de roles entre investigadores especialistas y practicantes como en distinto grado aparecen en las dos tradiciones anteriores. Son los propios practicantes quienes abordan la reforma e innovación institucional y quienes en todo momento controlan este proceso.

Tal y como plantea Proppe (1990: 330),

«[trascender] el objetivismo y el relativismo supone la necesidad de un nuevo paradigma: Es aquí donde encaja la propuesta de un paradigma dialéctico. Este tercer paradigma propugna la acción, en forma de evaluación dialéctica, considerada como una serie de actividades que ayudan a los participantes (incluido el evaluador) a percibir, comprender y *reaccionar* —con mayor autoconciencia, autenticidad y responsabilidad— a las consecuencias de un proceso que resultan de su interacción con su ambiente. En este paradigma es esencial recurrir al contexto histórico y cultural de cada caso».

De acuerdo con la perspectiva reformista de House (1980), los desacuerdos en cuanto a evaluación entre aquellos con poder e interés en una organización tienen lugar dentro de un consenso general con respecto a los valores democráticos fundamentales. Él hace alusión a los modelos críticos de evaluación emergentes, reconociendo que cualquier persuasión de dicho tipo concerniente a la evaluación quedaría fuera de su propio marco para comprender la situación política de la evaluación¹. Desde luego, está claro que esta tradición no asume el marco de la democracia liberal en el que se movían las tradiciones anteriores.

3.4. Posiciones posparadigmáticas

Los posmodernistas parten de que las filosofías y epistemologías surgidas desde la Ilustración están acabadas, han llegado a su fin. Niegan la vigencia del proyecto emancipador de la modernidad. Es más, no creen que sea la razón la que nos permita solucionar nuestros problemas, sino más bien tenderá a enquistarlos y a agravarlos. Por ello es necesario *deconstruir*, *desmontar* y *desnormalizar* los discursos que se han ido elaborando. Cualquier gran proyecto emancipador es sospechoso y, de hecho, para ellos la democracia liberal y el marxismo no han sido sino dos formas políticas de controlar y marginar al pueblo. La postura teórica posmoderna, Hargreaves (1996: 65)

«supone negar la existencia de un conocimiento fundamental sobre la base de que no existe una realidad social cognoscible más allá de los signos del

¹ HOUSE (1994: 43) citaba en 1980 a Apple, Lundgren, Kallos y Paulton como representantes de un nuevo enfoque y decía que aunque «estos autores presentan diferencias sustanciales entre sí, su enfoque global consiste en asumir una postura teórica que tiene en cuenta de forma deliberada a la sociedad».

lenguaje, la imagen y el discurso. En consecuencia, no puede existir acuerdo preestablecido sobre las maneras de comprender cosas que llamamos sistemas sociales, o incluso otras identidades humanas, porque también ellas carecen de una esencia interior más allá del lenguaje, la imagen y el discurso».

Su posición epistemológica surge desde la consideración de que los paradigmas que en los dos últimos siglos habían servido como soporte epistémico y conceptual al conocimiento han entrado en crisis, por lo que consideran necesario situarse en un espacio diferente que, pese a su dificultad y complejidad, resulta irrenunciable. Rechazan las posiciones empiristas y las positivistas, enmarcándose más dentro de una epistemología interpretativo-constructivista, en el sentido de intentar crear un espacio epistémico fruto de la *deconstrucción* de las teorías y metodologías precedentes, cuyos elementos recuperables los *articula* de manera significativa y congruente en la configuración teórica que se asume.

A este respecto, conviene recordar, que el postmodernismo desde su defensa de la metodología dialéctica entiende que el conocimiento sobre todo en el campo de las ciencias sociales está fundamentado histórica y culturalmente, cargado de valores políticos y morales, y al servicio de determinados propósitos e intereses. Rechazan su carácter acumulativo y progresivo, así como que en su adquisición se proceda a través de formas pasivas. Mediante análisis de discurso (de Alba, 1996), emprenden tareas de deconstrucción y articulación significativa, de cara a desmontar los elementos de las distintas capas de significación, de componentes arqueológicos o de elementos lingüísticos y extralingüísticos en los juegos de lenguaje; y se analizan en el contexto de su constitución genealógica y de su desarrollo o cambios de significación, articulándose en la configuración analítica que se haya decidido trabaja en consonancia con los intereses y fines del análisis.

La deconstrucción parte del conocimiento y el análisis de los contextos discursivos sociales amplios, considerando como herramienta tanto a las hipótesis iniciales de trabajo, como a los elementos de mayor fuerza irruptiva, dislocadora u organizadora de la configuración discursiva que se analiza, para conocer la significación de dicha configuración, sus desplazamientos y condensaciones más importantes y las potencialidades y tendencias de sus imaginarios, bloqueos discursivos, históricos, culturales, epistémicos, etc. Y, por su parte, la articulación significativa recupera el trabajo realizado en la deconstrucción, con el fin de ofrecer una lectura verosímil y articulada del discurso analizado.

4. LOS INDICADORES DE EVALUACIÓN

En los últimos años y, sobre todo, a raíz de la proliferación de estudios internacionales se ha ido extendiendo el uso de los indicadores de evaluación con la finalidad de lograr una información más objetiva y fiable en la comparación de la situación de los servicios públicos en distintos contextos nacionales. Normalmente los indicadores hacen referencia a series temporales de datos que tienen un carácter unívoco y homogéneo en diferentes situaciones y contextos y que, por ello, ponen de manifiesto los cambios que se producen en las dimensiones de evaluación.

Un indicador educativo es un estadístico sobre el sistema educativo que revela algo sobre su actuación y su «salud». Al igual que la temperatura (que utiliza un termómetro de mercurio), un indicador aporta información sobre el funcionamiento del sistema educativo en un momento dado, señala cómo se está desarrollando su progreso, y, según sus defensores, previene de problemas futuros. No toda estadística es un indicador. Para serlo necesita un punto de referencia, es decir, un estándar frente al que juzgar al estadístico.

Para ser útil en un contexto político y de toma de decisiones, un indicador debería aportar uno de los siguientes tipos de información:

- Información que describa la actuación (*performance*) del sistema educativo, en el logro de condiciones y resultados deseables. (Por ejemplo, puntuaciones altas en los tests de logros; decrecimiento del número de abandonos escolares).
- Información sobre acontecimientos, instancias y problemáticas del sistema que mantengan conexión conocida con los resultados deseados. (Por ejemplo, factores docentes y escolares relacionados con el aprendizaje del alumnado).
- Información que describa elementos centrales del sistema, como por ejemplo, cantidad de recursos financieros disponibles, «cantidad» de trabajo docente y currícula escolares ofertados.
- Información orientada a problemas; es decir, centrados en problemas actuales o potenciales en el sistema educativo (por ejemplo, demanda de docentes, circunstancias escolares en zonas social y económicamente deprimidas).

- Información que sea relevante para la toma de decisiones política; es decir, que describan condiciones educativas particularmente relevantes para la administración y pertinentes para el cambio de orientación en las políticas públicas sobre educación.

a) Características técnicas que deben poseer los indicadores educativos

Los indicadores han de poseer una serie de características técnicas para resultar operativos como elementos de información sobre el sistema.

- Los indicadores deben medir hechos y cuestiones ubicuas de la escolaridad, es decir, aquéllas que pueden ser encontradas en alguna forma a lo largo del sistema.
- Los indicadores deben medir hechos y cuestiones perdurables del sistema, por ejemplo, el número de cursos en matemáticas requeridos para la graduación universitaria.
- Los indicadores tienen que poder ser comprendidos fácilmente por una amplia audiencia de educadores y administradores.
- Los indicadores tienen que ser «posibles» en términos de tiempo, costo y conocimiento experto requerido para recoger la información relevante.

b) Tipos de Indicadores

Los dos tipos de indicadores básicos son las *estadísticas simples* como la ratio, salarios de los docentes, o el cero por ciento destinado a la enseñanza en los presupuestos del Estado. También pueden ser indicadores las estadísticas compuestas, que miden y registran combinaciones de acontecimientos y características relacionadas. Estas estadísticas se suelen expresar en una tabla de doble entrada, en la que se combinan, por ejemplo, comarcas (o provincias) con nivel de pobreza, y dominio lectoescritor. Las estadísticas compuestas más elaboradas son aquéllas que utilizan índices de correlación.

c) Modelos de indicadores / modelos de escolaridad

Algunos autores plantean, correctamente, que los modelos de indicadores (o sistemas de indicadores) señalan un modelo conceptual de cómo funciona el sistema educativo, pero esta puntualización es insuficiente si no se indica, además, que los indicadores suponen esencialmente una forma de concebir la escolaridad; es decir, al seleccionar indicadores, seleccionamos lo que en un momento determinado y por unos agentes concretos, se consideran elementos, acontecimientos, problemas fundamentales y definitorios de la escolaridad. No podemos separar, al igual que ocurre con la selección cultural que todo currículum realiza, ciertas estadísticas de la visión que representan sobre el sistema educativo, y lo aceptado como valioso, perdurable y necesario del mismo. En este sentido, es posible afirmar que los indicadores son un elemento político, antes que instrumental; susceptible, por lo tanto, a las preferencias ideológicas y valorativas que significan, dan sentido, conforman y dominan momentos históricos concretos e instituciones particulares.

A continuación presentamos algunos modelos formales que identifican elementos de la escolaridad en las figuras 2.1, 2.2 y 2.3.

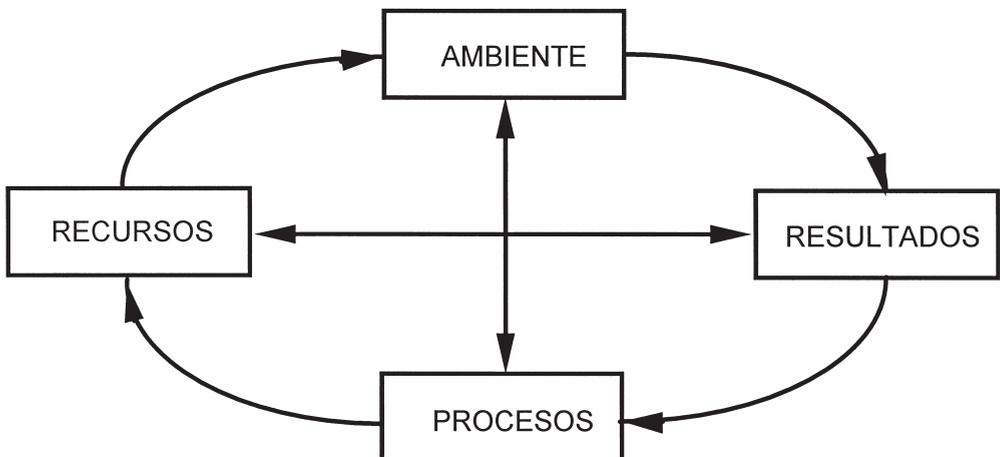
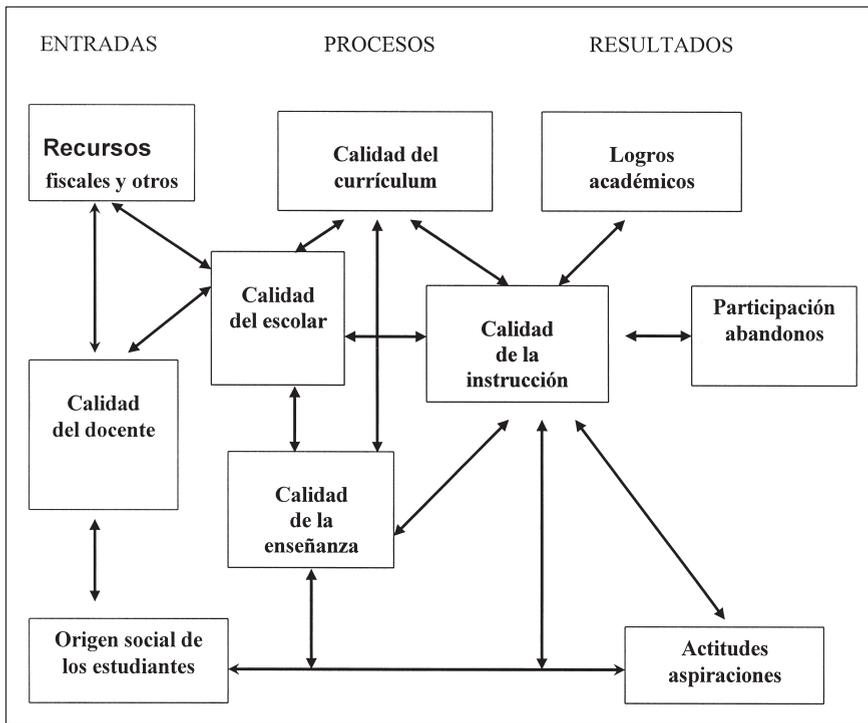
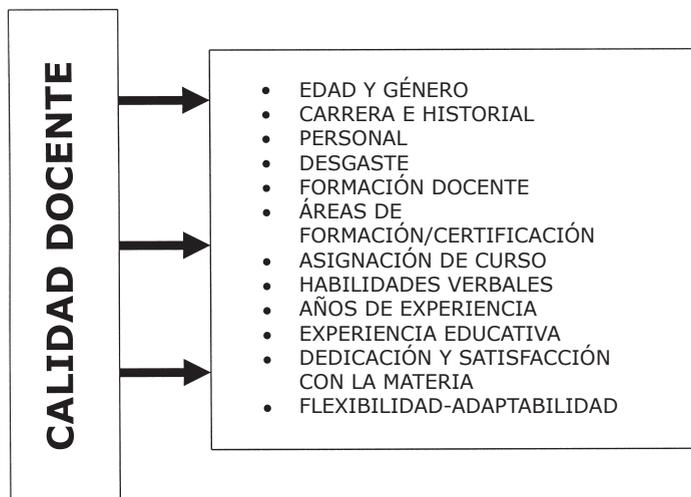


Figura 2.1. Modelo genérico de los elementos en que se centran los indicadores.



Fuente: Oakes (1987), pág.9.

Figura 2.2. Un modelo comprensivo de sistema de enseñanza.



Fuente: Oakes (1987), pág.9.

Figura 2.3. Componentes de la calidad docente.

Contemplando las distintas figuras, podemos apreciar la creciente complejización, tanto en lo referente a las relaciones como a los mismos elementos o componentes. El último modelo es el más elaborado, aunque respete los tres elementos básicos de los primeros (entradas, procesos y resultados). En dicho modelo, cada uno de los elementos seleccionados representa un conjunto de indicadores, que a su vez, puede ser subdivididos (por cuestiones de operatividad, por ejemplo) en otros indicadores. Este es el caso de la calidad del docente, que según su autora –Oakes (1987)- se encuentra relacionada con cuestiones como la edad, los años de experiencia, etc.

Es muy importante apreciar que incluso en el modelo más complejo, el papel de la administración se encuentra reducido a los recursos fiscales o de otro tipo.

d) Indicadores y toma de decisiones

Se afirmaba antes que los indicadores representan posiciones —políticas— sobre lo que se considera valioso en un sistema educativo, en razón de que tomados en su conjunto suponen un modelo de escolaridad concreto. Pues bien, los indicadores además plantean otros dos problemas valorativos de suma importancia: en primer lugar, la publicación (cuantitativa) de los indicadores ha de ser interpretada por los receptores de la misma; interpretación que, inevitablemente, se realiza desde un marco valorativo concreto. En ocasiones, los defensores de planteamientos conservadores y tecnócratas, no suelen percatarse de dos cosas: una, que los indicadores, incluidos los más elaborados, son aproximaciones, índices de la realidad, pero no la realidad misma; dos, que la distancia entre un juicio sobre la calidad de un sistema educativo y los datos cuantitativos que los indicadores presentan, es más considerable de lo que se suele admitir.

En segundo lugar, los indicadores, al menos por las dos razones que acabamos de aducir, no pueden *ni deben* sustituir el juicio de los receptores de los mismos. Este punto es muy importante porque los indicadores se elaboran para tomar decisiones políticas sobre el presente y el futuro del sistema educativo, es decir, sobre el trabajo y la vida escolar de los que en él participan, y, claro está, sobre el tipo de educación que una sociedad y una nación pretende potenciar y desarrollar.

Como afirma Oakes (1987:23):

«Los indicadores generarán conocimiento sobre el sistema educativo que alimentarán los debates sobre la escolaridad. Los indicadores influyen en lo que piensan la administración y el público en general sobre la escolaridad... Pero aunque construyamos los mejores indicadores, nunca identificarán las causas, ni aportarán soluciones inequívocas a los problemas educativos. Al final, las interpretaciones sobre lo que signifiquen los indicadores y las decisiones sobre qué políticas aplicar estarán influidos tanto por valores como por conocimientos. Los indicadores no pueden hurtar las decisiones fundamentales sobre la escolaridad del proceso político, de la reflexión y el debate entre los administradores y el público, responsables, todos ellos, de la salud de su funcionamiento».

No obstante, existe otro matiz importante relacionado con la toma de decisiones y con la construcción de los indicadores mismos. Quiero decir que los procesos de selección de los mismos suelen estar ubicados en agencias e instituciones nacionales e internacionales gubernamentales o privadas, compuestas por expertos y administradores, en las que no suelen participar muchos de los agentes y sujetos que se verán afectados por los mismos y que viven o son receptores y «usuarios» del sistema educativo. Esta cuestión crea una brecha a través de la cual, es mucho más factible que las valoraciones tecnocráticas, las presiones políticas nacionales e internacionales particulares, primen ciertas cuestiones sobre el sistema educativo sobre otras igualmente importantes pero marginadas. Al mismo tiempo ello ha permitido la extensión de cierto secretismo y la consagración de unas barreras artificiales, legitimadas por razones «técnicas» en la superficie, pero políticas en el fondo, entre los que conciben, crean y aplican los indicadores y los que tienen que organizar su trabajo escolar al dictado de las interpretaciones y las valoraciones que la publicidad de los indicadores introduce en la arena pública.

El problema de quién decide qué indicadores se ha intentado solucionar identificando los niveles de actuación de los mismos. Así, por ejemplo, Oakes (1987: 35) identifica tres niveles a los que les corresponden conjuntos de indicadores propios: indicadores estatales, indicadores autonómicos e indicadores comunes. Sin embargo, esto no hace más que trasladar a un terreno «técnico» la discusión política de fondo. La pregunta «¿quién decide qué indicadores?» no queda resuelta de esta manera, sino aplazada

cuando no ocultada. Como han afirmado Bryk y Hermansson (1993: 466), los indicadores tendrían que aportar información que estimulase y animase a los participantes escolares a reflexionar y analizar a fondo sus propias actividades. Los indicadores deberían invitar al público a comprender con profundidad las instituciones y los procesos educativos, allí donde ocurran y no sólo en las escuelas». Para ello, sería necesario no sólo mejorar técnicamente los indicadores, sino seleccionar también aquéllos que estén relacionados con las necesidades cotidianas tanto de los docentes como de los que indirectamente estén relacionados con los fenómenos educativos. Todo esto, supone, al menos, cambiar los parámetros vigentes para que la pregunta formulada tenga una respuesta claramente distinta a la que actualmente podría hacerse. Pero éste, como tantos otros, es un tema de debate político antes que técnico.

5. LA POLÍTICA Y ÉTICA DE LA EVALUACIÓN

La evaluación es una actividad que se lleva a cabo en un determinado contexto social y político. House (1980) analiza los distintos enfoques que surgen y se desarrollan en el marco de la concepción política de la democracia liberal. Todos ellos aceptan el campo y las reglas de juego de dicha concepción. No obstante, los modelos que trabajan desde la óptica de la epistemología subjetivista, al reconocer la pluralidad de valores e intereses de los implicados, plantean la necesidad de la responsividad (Stake, 1967) de la evaluación a las distintas audiencias, cuestionando los principios de la epistemología objetivista y su uso de una ética utilitarista.

Entendida de esta manera, la evaluación cobra una dimensión política diferente ya que el objeto de estudio debe ser definido teniendo en cuenta los problemas y necesidades de todos los implicados, y no sólo de los patrocinadores, y la diseminación de información dota de conocimientos a las audiencias concernidas para afrontar un debate democrático sobre las diferentes visiones en juego y, en definitiva, sobre las formas de distribución y legitimación del poder. A este respecto, convendría tener muy en cuenta las palabras de Cronbach (1987: 35), cuando nos recuerda que:

«El evaluador está al servicio de una o más fuerzas políticas. Como ya señalé en *Toward Reform*, los evaluadores ayudan a que los políticos se perpetúen en el poder al proporcionarles aquella información de la que

carecen los otros participantes en el proceso político, lo cual puede contribuir a la concentración del poder. Sin embargo, debería ser posible que el evaluador se erigiera en consejero de toda la comunidad educativa, y esto sucedería si pudiera diseminar su información y encarar las inquietudes de todas las partes que tienen intereses en juego».

Además de otras consideraciones que deben ser tenidas en cuenta, la democratización de la evaluación, planteada en estos términos, es una consecuencia del paso de la relación entre evaluadores y patrocinadores a la establecida entre evaluadores y audiencias, y, principalmente, de su consideración como servicio a una comunidad en la que existen una pluralidad de valores e intereses.

Desde esta conceptualización de la evaluación como un servicio a una comunidad plural, MacDonald (1976) propuso su modelo de evaluación democrática en el que el evaluador asume el papel, como agente neutral, de poner en contacto y facilitar el conocimiento mutuo de grupos con posiciones, valores e intereses diferentes.

Dicho enfoque, no cabe duda que supone un avance con respecto a las tendencias gerencialistas dominantes de evaluación burocrática, pero no está exento de una cierta ingenuidad, al presuponer que las audiencias pueden alterar las relaciones de poder mediante el conocimiento de la evaluación, y de una visión un tanto idealizada de lo que son las relaciones humanas entre grupos con posiciones asimétricas en la jerarquía administrativa y desiguales en su capacidad de autodeterminación. Ya hemos adelantado en otro momento, las críticas que Elliott (1986) hacía a MacDonald en relación a su defensa de que los individuos tienen el «derecho a estar informados» como miembros de grupos de interés particular más que como miembros del público en general.

De una manera explícita, Elliott (1986: 231) afirma que:

«El evaluador *democrático* no puede ser un “agente de cambio” neutral, que transmite información a diversos grupos de interés. Su responsabilidad es para con un sólo auditorio, el público en general, y no para con una pluralidad de auditorios. Si esto es así debe evaluar según los criterios que definen el “bien común” o el “interés general”. No puede permanecer neutral en cuestiones de uso de la información. MacDonald se equivoca en su afirmación de que la evaluación democrática carece de la noción de “mal

uso de la información”. El evaluador no puede rehuir la responsabilidad de enjuiciar según criterios públicos si quiere mantenerse en una postura genuinamente democrática».

Como consecuencia de todos estos aspectos y, sobre todo, por no abordar los supuestos y desigualdades en que los distintos grupos se sitúan para, de alguna manera, abordar su transformación, la propuesta de evaluación democrática de MacDonald ha sido calificada, en no pocas ocasiones, como políticamente conservadora. Quizá, por ello, y desde el reconocimiento de lo que supuso, pero intentando superar sus deficiencias, House & Howe (1999, 2000) han dado un paso más con su concepto de evaluación democrática deliberativa que, como también hemos señalado con anterioridad, implica considerar conjuntamente las exigencias de inclusión, diálogo y deliberación.

Reconociendo que los evaluadores llevan a cabo su tarea en circunstancias sociales concretas y admitiendo que la perspectiva democrático-deliberativa está demasiado idealizada como para poder ser aplicada al pie de la letra en un mundo como el nuestro, nuestros autores proponen que, al menos, se tome como una guía de actuación. Su puesta en práctica según ellos, House & Howe (2000: 10-11), debe hacerse planteándose las diez preguntas siguientes:

1. *«¿Qué intereses están representados?»* En la evaluación se deberán considerar los intereses de todas las partes relevantes. Normalmente, se entiende por ello las opiniones y los intereses de todos los que tengan una participación activa en el programa o en la medida que se somete a evaluación.
2. *«¿Están representados los principales involucrados?»* Lógicamente, no es posible incluir a cada particular involucrado. Por muy poco preciso que sea, los evaluadores deberán conformarse con los representantes de varios de los grupos involucrados. No hay duda de que hay ocasiones en que no todos los involucrados pueden quedar representados. Tal representación puede suponer que los evaluadores lleven al estudio los intereses de los involucrados en su ausencia.
3. *«¿Ha quedado excluido algún involucrado?»* A veces quedan excluidos importantes grupos, y muy frecuentemente son los que menos voz o poder tienen, es decir, los pobres, los desvalidos y las minorías. Los evaluadores tienen la tarea de representar dichos grupos de la mejor forma que puedan. Después de todo, ¿quién representa el interés público?

4. *¿Se dan desequilibrios importantes de poder?* Durante la realización de los cuestionarios, a menudo surgen intereses demasiado poderosos que, por tanto, amenazan a la imparcialidad de las conclusiones. A menudo, proceden de los propios clientes o de grupos de involucrados particularmente poderosos y consiguen dominar las condiciones de realización del estudio o las opiniones en él reflejadas.
5. *¿Hay alguna forma de controlar tales desequilibrios?* Entra dentro del ámbito de actuación del evaluador el controlar los desequilibrios de poder. Así como los profesores son responsables de crear las condiciones para una discusión eficaz en las clases, los evaluadores deben establecer las condiciones para una recolección de datos, un diálogo y una deliberación eficaces. Hemos de admitir que esto requiere una capacidad de juicio particularmente refinada.
6. *¿Cómo participa la gente en la evaluación?* La forma de participación suele resultar fundamental. La participación directa lleva mucho tiempo y dinero y, además, es potencialmente peligrosa para la imparcialidad. De todos modos, para obtener información de manera correcta, es necesario que los grupos involucrados participen con seriedad.
7. *¿Con qué grado de autenticidad participa?* Los participantes están acostumbrados a rellenar cuestionarios que no les interesan lo más mínimo. Conforme se van empapando de responsabilidad, van prestándole menor atención a sus respuestas. Se trata de un grave problema cuyas dimensiones pueden terminar aumentando con rapidez. Un uso artificial, falta de sinceridad de las evaluaciones puede llevar a conclusiones falsas.
8. *¿Con qué grado de involucración participa?* Una vez más, aunque la interacción resulta fundamental, también puede terminar siendo excesiva. ¿Deberían participar los grupos involucrados en análisis de datos altamente técnicos? Podrían llegar a tergiversar las conclusiones. A la inversa, una involucración superficial o artificial puede resultar igual de negativa.
9. *¿Hay deliberación reflexionada?* Por norma general, los evaluadores acaban sus estudios con retrasos, apresurándose para llegar a un plazo fijado. Las conclusiones no se meditan lo suficiente ni se incluye toda la deliberación que se debería en el estudio. En esta fase del estudio, resulta tentador reducir el grado de involucración de los demás para poder llegar a tiempo a los plazos.

10. *¿Qué consideración y extensión le han dado a la deliberación?*

Normalmente, cuanto más extensa sea la deliberación, mejores serán las conclusiones que podamos extraer. En la mayoría de los casos, la evaluación, más que ser excesiva, peca de escasa. La perspectiva democrático-deliberativa es muy exigente a este respecto, aunque a esto se junta el hecho de que el error más común de las evaluaciones es que las conclusiones no se ajustan muy bien a los datos».

Desde una posición más crítica, tanto a la concepción política de la democracia liberal como al papel de la evaluación profesional, Kemmis (1986: 131) refuerza el protagonismo de los participantes en el desarrollo curricular afirmando que:

«Enfatizar el valor de la autocrítica no es defender el aislamiento de un programa. Al contrario, es pretender desarrollar el programa en base a un debate donde los participantes puedan adoptar puntos de vista más amplios que los propios, llegando a la visión global del programa, dentro de un amplio contexto histórico. Pero también es recordar que una vez que un programa cede la autocrítica para dejar paso a la autoridad exterior, está perdiendo su autonomía como comunidad intelectual».

Para él, las aportaciones de evaluaciones externas, independientes y metasevaluaciones pueden contribuir enormemente a ilustrar, enriquecer y ampliar el debate sobre el que se sustente el desarrollo de un programa, en la medida que supone incorporar visiones externas que favorecen la comprensión y el diálogo, pero la «autoevaluación por parte de la comunidad del programa es la base de la evaluación de un proyecto o programa; otros tipos de evaluación pueden enriquecerla pero no sustituirla». En la postura de Kemmis, subyace un cuestionamiento de los principios de «verdad», «justicia», «interés público», «intereses de los diseñadores o promotores» cuando son planteados al margen de la pluralidad de valores y de la comunidad de intereses que concurren en cualquier actividad educativa. Todos ellos no pueden ser establecidos y definidos desde una sola de las instancias presentes, ya sea a través de las legitimidades de los evaluadores, de los patrocinadores o de la administración educativa. Es necesario que se conozcan y contrasten la diversidad de intereses y que se tienda a una construcción social negociada de los valores y principios que deben presidir el desarrollo del programa y su evaluación.

La evaluación es también un proceso y una actividad ética en la medida que requiere que los procedimientos con los que se lleva a cabo observen y respeten una serie de principios, de obligaciones o normas morales. MacDonald (1976), Elliott (1986), House (1980, 1990, 1993), Norris (1990), Simons (1984, 1985, 1987), etc., se hacen eco de la importancia de este tema dentro de la evaluación educativa.

En algunos casos (House, 1990: 2478), advirtiendo de los cuatro problemas éticos fundamentales que aparecen con más frecuencia en la evaluación:

- «a) ocultar a los participantes la naturaleza de la evaluación, o hacerles participar sin que lo sepan;
- b) exponer a los participantes a actos que podrían perjudicarles o disminuir su propia estimación;
- c) invadir la intimidad de los participantes; y
- d) privar a los participantes de los beneficios».

Morris (1999: 19) completa las observaciones hechas por House (1990), estableciendo las más frecuentes y más serias violaciones de las percepciones de los evaluadores. Véase el Cuadro I.

Cuadro I: Las más frecuentes y más serias violaciones de las percepciones de los evaluadores.

Más frecuentes	Más serias
El evaluador selecciona preferencialmente pruebas o tests con las que está familiarizado.	El evaluador cambia las preguntas de la evaluación para igualar el análisis de datos.
El evaluador pierde interés en la evaluación cuando el informe final está entregado.	El evaluador promete confidencialidad cuando no puede garantizarla.
La evaluación responde más a las preocupaciones de un grupo de interés que a las de otro.	El evaluador, una vez llegado a un acuerdo, toma decisiones sin volver a consultar con el cliente.
El evaluador deja de buscar qué valores son válidos para la audiencia.	El evaluador conduce la evaluación sin tener suficiente destreza o experiencia.
El evaluador escribe un informe sumamente técnico para una audiencia técnicamente no sofisticada.	El informe de evaluación es elaborado de forma partidista para ensombrecer las debilidades de grupos de interés.

Fuente: Morris (1999)

En otros casos, House (1980) y Norris (1990: 106), proponiendo los valores básicos que deben servir como principios morales de la evaluación, a saber:

1. «Igualdad moral: que todas las personas se deben tomar como miembros del mismo grupo de referencia y por lo tanto deben ser tratados con igualdad.
2. Autonomía moral: que nadie debe imponer su voluntad a otros por fuerza o coerción o por otros medios ilegítimos.
3. Imparcialidad: los conflictos de intereses se deben resolver imparcialmente, es decir teniendo todos los intereses representados y no favoreciendo a ninguno mediante cualquier procedimiento que se emplee de decisión.
4. Reciprocidad: que no debe haber posiciones de privilegio entre la gente, tratando a otros como nos gustaría que nos trataran».

Helen Simons (1982) en su defensa de la evaluación democrática, sugiere una serie de principios que deben ser respetados tanto por los evaluadores como por los miembros de los centros educativos:

- «a) Imparcialidad: es necesario que el investigador-evaluador tome recaudos contra las opiniones o percepciones sesgadas, empezando por las propias. Su rol es exponer las distintas alternativas que pueden presentarse en la política de la escuela sin oponer los propios juicios o recomendar determinados cursos de acción.
- b) Confidencialidad-control: la información que puede considerarse como privada del individuo debe ser tratada confidencialmente. Aquella que tiene un estado de conocimiento público dentro de la escuela debe ser controlada en cuanto a su difusión al exterior de la institución. Es importante por una parte evitar el cotilleo y por otra que los participantes tengan control sobre sus propios datos. En ese caso existen más posibilidades de que estén predispuestos a ofrecer percepciones más exactas y responder honestamente acerca de los problemas que están en discusión. Por otra parte el evaluador no debe acceder a los documentos, informes o reuniones sin la autorización de los implicados. Los informes deben ser orientados hacia el problema en cuestión, siempre debe conservarse el anonimato de los informantes de modo que la evaluación tenga un carácter despersonalizado.

- c) Negociación: este principio es central en la evaluación democrática ya que protege a los participantes en el uso y control de la información. La negociación debe establecer: 1) los límites del estudio, 2) la relevancia de las necesidades institucionales y de las innovaciones curriculares, 3) el contenido, audiencia y alcances de la difusión del o los informes de la evaluación, 4) control de los posibles sesgos o distorsiones de la información. En síntesis, debe fomentar una evaluación justa y participativa.
- d) Colaboración: Las prevenciones iniciales, fundamentalmente basadas en la supuesta amenaza a la privacidad del desempeño del profesor en su clase pueden superarse si se asegura el compromiso de compartir la información relevante para discutir honestamente la política de la institución. La información confiable debe servir para proporcionar una base de autorreflexión a aquellos que deseen colaborar, sin presiones de ningún tipo. La información debe ser contrastada por uno o dos observadores externos.
- e) Dar cuenta de la acción de la escuela (accountability). Deben encontrarse vías para incrementar la información acerca de la institución al exterior, pero teniendo en cuenta que la evaluación y crítica interna es mucho más importante. La información al exterior debe proporcionarse con ciertos recaudos, dependiendo del ambiente local y de la naturaleza de los problemas que fueron objeto de evaluación».

La propuesta y desarrollo de la ética descansa básicamente en la consideración de que la evaluación proporciona información y que, obviamente, la información es poder. No siempre, por tanto, hay acuerdos en este tema, ya que para muchos autores, las categorías y principios que utilizan los evaluadores asumen las del *status quo* vigente o bien tratan asimétricamente los valores de los distintos grupos presentes en la evaluación. Por ejemplo, House (1990: 2479) señala, a este respecto, que la propuesta de evaluación democrática de MacDonald acentúa la autonomía individual de los participantes hasta el extremo de que tienen el derecho de veto con respecto a lo que se dice de ellos en el informe de evaluación. A su juicio, esto supone aceptar que «el conocimiento de la acción social es propiedad privada de los profesionales, y la verdad depende de los diferentes marcos interpretativos que guían la conducta de los agentes sociales», ya que a cada uno le pertenecen los hechos de su vida.

Más concretamente, la compaginación de principios como los de confidencialidad y negociación, en lo que se refiere en esta última a la publicidad de los resultados de la evaluación, resulta en ocasiones bastante problemática, en especial cuando se defiende genéricamente como el «derecho de todos a estar informados». Jenkins (1986), por ejemplo, criticó la consistencia de los principios éticos utilizados en el SAFARI al argumentar que la ética de acceso (confidencialidad) entraba en contradicción con la ética de publicidad (negociación), pues si mediante la primera, el evaluador se ve limitado para recoger los datos e informaciones necesarias para la evaluación, cuando llega el momento de difundir el informe, es decir, cuando se tiene que aplicar el segundo, el evaluador a través de la negociación utiliza todo tipo de subterfugios para limitar el control de los participantes y ampliar el suyo. En el fondo, para Jenkins, este problema trasluce las relaciones de poder entre evaluador y participantes. Helen Simons, también es consciente de estos problemas, y aunque no llega a considerar incoherentes dichos principios, sí plantea que dentro de estos dilemas el evaluador se ve obligado en bastantes ocasiones a relativizar el derecho al control de los participantes de cara a dar publicidad a la evaluación.

Elliott (1986) es mucho más tajante al afirmar que los procedimientos del SAFARI se basan en dos puntos de vista éticos inconsistentes, dado que si el «derecho a estar informado» tiene que basarse en el principio ético de la negociación como procedimiento, ésto no justifica éticamente la promesa de confidencialidad. Para Elliott (1986: 232):

«El concepto justificativo clave para esta última, sería “el derecho a la intimidad”. MacDonald, en su descripción de los conceptos claves de la evaluación democrática, prefiere ignorarlo, y, sin embargo, es el único argumento mediante el que se puede justificar *éticamente*, y no sólo *técnicamente*, la *promesa de confidencialidad*. Si el “derecho a estar informado” es la única justificación —como afirma MacDonald—, la “promesa de confidencialidad” carece de intención ética real».

Si la evaluación responde a los distintos intereses en juego de las audiencias, es decir si es responsiva, o por el contrario se centra exclusivamente en las necesidades de los patrocinadores, el cometido de los evaluadores será muy distinto en lo que se refiere a la misma delimitación del objeto de evaluación (será necesario conocer y negociar los alcances del

estudio), a la identificación de las fuentes de recogida de información (de quiénes y de qué vamos a recabar datos) y al establecimiento de los procedimientos con los que se va a realizar el estudio (contrato de evaluación).

La finalidad o para qué de la información, y en definitiva de la evaluación, no parece que ofrezca muchas dudas. Una parte muy importante de las definiciones de evaluación insisten en que la evaluación es una fuente de información que sirve para fundamentar la toma de decisiones, y que, asimismo, éstas pueden vincularse a la valoración terminal que se hace una vez concluida la acción o mientras se está llevando a cabo la misma, es decir, al finalizar o en el proceso mismo de desarrollo si nos atenemos a la ya clásica distinción entre evaluación sumativa y formativa (Scriven, 1967).

Junto a esta finalidad de la toma de decisiones, ha ido abriéndose paso también la función de la evaluación como mecanismo que contribuye a ilustrar, iluminar o a mejorar el conocimiento de las dinámicas educativas. Desde luego, parece claro que ésta es una de las finalidades primarias de la evaluación, tanto porque en sí misma la evaluación es una forma de conocimiento, como porque cuando se incorpora a los quehaceres institucionales se hace para disponer de una visión sistemática que de otro modo no es posible tener. Otra cosa distinta es lo que posteriormente se haga con esta información. Pero, en cualquier caso, difícilmente podrán tomarse decisiones si, por una parte, no existe un conocimiento informado de lo que está ocurriendo en los ámbitos educativos de que se trate, y si, por otra, los distintos actores educativos no se sensibilizan o a toman de conciencia, precisamente a través de la evaluación, de los logros y carencias de sus quehaceres profesionales e institucionales.

La evaluación es, asimismo, como han indicado no pocos autores una forma de persuasión. Precisamente House (1994: 71) habla de que «la evaluación aspira a persuadir a un público concreto del valor de algo o de que ese algo es lo que interesa, apelando a la razón y comprensión del público». Y, por ejemplo, Shaw (1999: 19) subtítulo el capítulo de su libro dedicado a las teorías de la evaluación y evaluación cualitativa, con la expresión *persuasiones y persuasores* para analizar las teorizaciones y autores de esta modalidad de evaluación. Lo que parece cada vez más claro es que una de las fuentes de legitimidad de los discursos políticos y de la acciones de los gobiernos, según House (1998: 124),

«procede de los profesionales de la ciencia, que han asumido parte de la autoridad que antes se confería a la religión. Por su identificación con la ciencia, la evaluación profesional legitima las actividades gubernamentales estimando el valor social de los programas y políticas de los gobiernos».

6. A MODO DE RECAPITULACIÓN

En este tema hemos tratado de precisar el concepto de evaluación contraponiendo, en primer lugar, los dos términos que habitualmente se utilizan para denominar la evaluación, “evaluation” y “assessment”. Tanto en el sentido como respecto al objeto, los significados y acepciones de los mismos son diferentes. El primero, *evaluation*, hace referencia al conocimiento de la calidad del servicio educativo prestado y, por consiguiente, con respecto al objeto se centra en los diferentes elementos que directa o indirectamente intervienen en el proceso de enseñanza-aprendizaje, curriculares, organizativos, metodológicos, administración educativa, profesorado, estudiantes, comunidad escolar, etc. El segundo, *assessment*, se centra en cambio en el impacto del servicio en los receptores, es decir, en el alumnado, fijándose por tanto, en lo que se refiere al objeto, en los resultados, rendimiento o calificaciones escolares.

Se han analizado asimismo los tres tipos de definiciones que explícitamente se posicionan acerca de en qué consiste o qué es la evaluación. En concreto, traemos a colación las definiciones basadas en metas, las descriptivas no enjuiciables y las descriptivas enjuiciables. En función de las características de cada una de ellas, es posible encuadrarlas en las dos grandes tradiciones de evaluación, que son analizadas casuísticamente en uno de los últimos libros de Rober Stake (2006), *la evaluación comprensiva* basada en la experiencia y vinculada al paradigma interpretativo (la segunda y la tercera), y *la evaluación basada en estándares* asociada a la medición y al paradigma positivista (la primera).

Una vez planteadas las concepciones y definiciones de evaluación, creímos que era el momento de tratar la relación entre evaluación, investigación e investigación evaluativa a los efectos de establecer sus aspectos comunes, en tanto que ambas están comprometidas con la generación del conocimiento, y las diferencias entre estas dos formas de indagación en cuanto a que el conocimiento producido, en la investigación, contribuye a

avanzar en el desarrollo de las disciplinas y campos científicos, mientras que en la evaluación sirve prioritariamente para fundamentar la toma de decisiones.

En un tema como el que nos ocupa, en el que se abordan elementos de estructuración y fundamentación de la evaluación de centros y profesores, no se podía obviar el estudio de las distintas tradiciones científicas, incluida las posiciones postparadigmáticas, así como el uso de los indicadores en la evaluación y los aspectos éticos y políticos de la evaluación.

BIBLIOGRAFÍA

- ANGULO RASCO, J. F.; CONTRERAS DOMINGO, J. y SANTOS GUERRA, M. A., (1991). «Evaluación Educativa y Participación Democrática», en *Cuadernos de Pedagogía*, n. 195, pp. 74-79
- (1994). «Evaluación Educativa y Participación Democrática», en J. F. Angulo y N. Blanco, (Comp.): *Teoría y Desarrollo del Currículum*. Málaga, Ediciones Aljibe, pp. 343-354.
- CRONBACH, L.J. (1987). *Designing Evaluation of Educational and Social Programs*. San Francisco, Jossey-Bass Publishers, third printing.
- (1991). «Nuestras noventa y cinco tesis», pág. 76, en J. Calvo de Mora, (1991): *Evaluación educativa y social*. Granada, Servicio de Publicaciones de la Universidad, pp. 75-87.
- DE ALBA, A. (1996). «Análisis de discurso educativo curricular: una propuesta de evaluación», documento policopiado (inédito).
- EISNER, E. W. (1998). *El ojo ilustrado. Indagación cualitativa y mejora de la práctica educativa*. Barcelona, Paidós Educador.
- ERICKSON, F. (1989). «Métodos cualitativos de investigación sobre la enseñanza», en M. C. Wittrock, (Ed.): *La investigación de la enseñanza, II. Métodos cualitativos y de observación*. Madrid, Paidós- MEC, pp. 195-301.
- HARGREAVES, A. (1996). *Profesorado, cultura y posmodernidad (Cambian los tiempos, cambia el profesorado)*. Madrid, Morata.
- HOUSE, E.R. (1994). *Evaluación, ética y poder*. Madrid, Morata.

- (1998). «Acuerdos institucionales para la evaluación», *Perspectivas*, vol. XXVIII, n. 1, marzo, pp.123-131.
- LEWY, A. (1990). «Evaluación formativa y sumativa», en T. Husen, y Neville Postlethwaite, T. (Dir.): *Enciclopedia Internacional de Educación*. Barcelona, Vicens Vives-MEC, vol. V, pp. 2550-2553.
- MACDONALD, B. (1985). «La evaluación y el control de la educación», en J. Gimeno y A. Pérez Gómez, (Eds.): *La enseñanza: su teoría y su práctica*. Madrid, Akal, 2ª ed., pp. 467-478.
- MARTÍN RODRÍGUEZ, E. (1988). «Profesión docente y autoevaluación institucional», *Revista de Educación*, n. 285, pp. 33-43.
- MILES, M. B. y HUBERMAN, A. M. (1984). «Drawing valid meaning from qualitative data: Toward a shared craft», en *Educational Researcher*, 13, pp.20-30.
- NEVO, D. (1998). «La evaluación mediante el diálogo: una contribución posible al perfeccionamiento de la escuela», *Perspectivas*, XXVIII, n. 1, marzo, pp. 87-100.
- NORRIS, N. (1990). *Understanding Educational Evaluation*. London, Kogan Page.
- PROPPE, O. (1990). «La investigación de la evaluación como una forma de potenciar el desarrollo en las escuelas y el profesionalismo de los profesores», *Revista de Educación*, n. 293, p. 325-343.
- SANTOS GUERRA, M. A. (1993). *La evaluación: un proceso de diálogo, comprensión y mejora*. Málaga, Aljibe.
- SCRIVEN, M. (1991). *Evaluation Thesaurus*. Newbury Park, Sage Publications, first printing.
- SHAW, I. F. (1999). *Qualitative Evaluation*. London, Sage Publications.
- SCHUMAN, E. A. (1967). *Evaluative research: Principles and practice in public service and social action programs*. New York, Russell Sage Foundations, citado por Stufflebeam y Shinkfield (1987): *op. cit.*
- STAKE, R. E. (1983). «Evaluación de programas: en especial de Evaluación de Réplica», en W. B. Dockrell y D. Hamilton, (Eds.): *Nuevas Reflexiones sobre la Investigación Educativa*. Madrid, Narcea, pág. 96.
- TIANA FERRER, A. (1998). *Tratamiento y usos de la información en evaluación*. Madrid, OEI, (serie Documentos-5), en <http://www.oei.es/>.

WALKER, R. (1983). «La realización de estudios de casos en educación: Ética, teoría y procedimientos», en W. Dockrell y D. Hamilton, (Eds.): *Nuevas reflexiones sobre la investigación educativa*. Madrid, Narcea, pp. 42-83.

— (1989). *Métodos de investigación para el profesorado*. Madrid, Morata.

Tema 3

El centro educativo como objeto de evaluación

Introducción

1. Perspectivas de la evaluación de centros educativos
 - 1.1. La evaluación de centros desde la perspectiva de la rendición de cuentas, eficacia y calidad
 - 1.2. La evaluación de centros desde la perspectiva de los procesos de mejora y desarrollo institucional
2. Estrategias para la evaluación de instituciones educativas
3. Las modalidades de evaluación
4. La especificidad de los distintos ámbitos de evaluación
5. Metodología de evaluación
6. Criterios e indicadores de evaluación
7. A modo de conclusiones

Bibliografía

TEMA 3

EL CENTRO EDUCATIVO COMO OBJETO DE EVALUACIÓN

Eustaquio Martín Rodríguez

UNED

INTRODUCCIÓN

La evaluación de las instituciones educativas se ha constituido, en el último tercio del siglo XX, en un campo de estudio en permanente auge en el que se entrecruzan perspectivas, visiones e intereses de muy diversa naturaleza. Parece claro que su importancia viene determinada por la necesidad, mayoritariamente aceptada de contar con información que permita conocer y optimizar la contribución de los centros escolares a la formación de las generaciones a las que acoge. Su protagonismo se debe a que son espacios sociales, dotados de sentido y significado, para hacer realidad un proyecto pedagógico. De su implicación y compromiso depende la viabilidad y las formas concretas de llevar a cabo las aspiraciones y expectativas educativas de la sociedad.

No es extraño, por tanto, que con el paso del tiempo se haya incrementado la demanda pública de información de lo que acontece en el interior de las instituciones. Desde instancias políticas, en la medida en que los gobiernos asumen la orientación y financiación de la educación como un bien de naturaleza social y pública, se han ido reforzando los cauces y mecanismos de obtención de información sobre los logros y resultados escolares. La sociedad civil no ha demostrado, asimismo, menos interés por conocer los niveles de calidad de los centros, ya sea porque en ellos está en juego el capital cultural de las nuevas generaciones, o porque, como también ocurre, tienen una posición definida en torno a los dilemas a los que se enfrenta hoy la enseñanza escolar.

En este tema se entra de lleno a plantear la evaluación de centros. Hemos considerado necesario que se parta del conocimiento de las perspectivas desde las que se puede afrontar este reto, para que una vez que se clarifiquen las orientaciones y propósitos con que se lleva a cabo la evaluación, hagamos explícitas las diferentes estrategias, modalidades, metodologías y criterios que específicamente son propias de este ámbito de evalua-

ción. A este respecto, conviene no perder de vista que todo proyecto de evaluación de centros de manera explícita o tácita se sustenta desde las posiciones adoptadas en los siguientes planos: ontológico, —concepción de qué es un centro escolar—; epistemológico, —cómo asegurarnos de que conocemos lo que afirmamos conocer—; metodológico, —cuáles son los distintos modos de proceder para llegar al conocimiento, a través de qué diseños, métodos y técnicas; y ético y político, —desde qué valores y principios evaluamos y al servicio de quién o quiénes está la evaluación—.

1. PERSPECTIVAS DE LA EVALUACIÓN DE CENTROS EDUCATIVOS

Desde una perspectiva académica los centros educativos han emergido como un lugar y espacio privilegiados de investigación y evaluación, posibilitando el desarrollo de campos disciplinares como el de la organización escolar, la creación de nuevas líneas de indagación como las de la teoría del cambio, escuelas eficaces, mejora de la escuela, eficacia y mejora escolar, la reestructuración escolar, etc., así como la ampliación de los ámbitos de evaluación como los referidos a los profesores, centros y sistemas educativos.

La teoría y práctica de la evaluación de las instituciones educativas se ve inmersa, por consiguiente, en un marco complejo de relaciones e influencias diversas, en el que conviven concepciones y visiones muy distintas en los planos ideológico, axiológico, epistemológico, etc. Por ello, es normal que tanto desde las condiciones externas, contexto socioeconómico, político y cultural, como desde las condiciones internas, esto es, espacios académicos en los que se despliega disciplinariamente la evaluación, se encuentren posiciones, a veces enfrentadas y con escasos puntos en común, de lo que significa la evaluación de centros.

Esta diversidad de perspectivas viene a reafirmar, una vez más, que la evaluación no es una actividad atórica ni consiste, fundamentalmente, en un acto puramente técnico o instrumental. La pluralidad de valores sociales; las diferentes formas de entender y concebir los centros educativos como organizaciones; la coexistencia de distintas tradiciones y paradigmas científicos; las variadas metodologías que pueden utilizarse para la recogida y tratamiento de la información; conforman la evaluación de centros como un campo complejo susceptible de problematizarse desde las visiones y experiencias con las que nos acercamos al mismo.

Para su análisis y estudio, hemos optado por plantear en primer lugar las dos grandes orientaciones desde las que se puede abordar la evaluación de centros educativos. De una parte, la que hemos denominado de rendición de cuentas, eficacia y calidad, y, de otra, la encuadrable dentro de los propósitos de desarrollo institucional y de mejora de la escuela. Se trata de dos orientaciones con énfasis muy distintos en cuanto a los elementos y valores en que centran la evaluación de las instituciones, a sus bases conceptuales y teorías de las organizaciones que mantienen, a las líneas de investigación que les sirven de apoyo, a los enfoques metodológicos que utilizan, etc. Aún cuando analizamos y resaltamos sus características diferenciales que, en la mayoría de los casos, son difícilmente conciliables, es verdad que existen espacios compartidos en los que cada una de las orientaciones reconoce las aportaciones de la otra.

El punto de encuentro más relevante es, sin duda, la consideración de los centros educativos como el núcleo y eje de los procesos de cambio desde los que acometer la mejora de la calidad de la educación. Existe un cierto acuerdo en valorar la importancia de los resultados escolares, aunque, en unos casos, se defiende como el único indicador de los objetivos y calidad extrínseca de la escolarización, y, en otros, se plantea como un elemento más no desvinculado de los objetivos intrínsecos de los procesos educativos en los que se tiene muy en cuenta los demás componentes del servicio prestado. La información y conocimiento proporcionados por la evaluación sirven también, en ambas orientaciones, para retroalimentar los distintos componentes de los procesos de enseñanza-aprendizaje. Sin embargo, son distintos los momentos, formas e instancias desde las que se integran los resultados de la evaluación en el quehacer institucional.

1.1. La evaluación de centros desde la perspectiva de la rendición de cuentas, eficacia y calidad

Esta orientación tiene un acusado componente político-administrativo. Su discurso va a ser recurrente desde los años setenta hasta la actualidad, y aún cuando sus argumentos tienen matices específicos, en distintos momentos o fases, su finalidad es inequívoca: valorar la productividad y rentabilidad de las instituciones educativas. Lo que empezó siendo una forma de comprobar la eficacia de los grandes programas norteamericanos

de educación compensatoria en los inicios de los setenta, fue tomando cuerpo, en los años ochenta, con el movimiento de rendición de cuentas auspiciado por las primeras experiencias de las políticas neoliberales anglosajonas, y se generalizó al contexto internacional con las opciones y decisiones que la globalización económica ha venido concretando durante los últimos años.

Como han señalado repetidamente distintos autores, el surgimiento de la *accountability*, primero en EE.UU. y más tarde en el Reino Unido, está asociado a un momento histórico caracterizado por un clima de desconfianza hacia la educación. Cronológicamente se inicia en los primeros años de la década de los 70, coincidiendo con la crisis energética y económica de los países desarrollados y con un período de recesión, inflación y aumento del paro. Esto va a dar lugar a un replanteamiento de los valores sociales y educativos, con fuertes recortes presupuestarios en todo lo relacionado con los programas sociales y con la generalización de un cierto escepticismo con respecto a las instituciones educativas y los profesores.

En el Reino Unido, el detonante que abrió el llamado «Gran Debate» fue el tantas veces citado discurso del primer ministro Callaghan en Ruskin College, Oxford, en octubre 1976. Aunque allí no se usó el hoy corriente concepto de «*accountability*», sí se hicieron explícitas las principales premisas del mismo, esto es: *a)* la escuela debe dar cuenta de su acción, *b)* los no profesionales de la educación tienen que tener la oportunidad de intervenir en la política educativa local e institucional (Becher and Maclure, 1978). Estos planteamientos fueron inmediatamente reconocidos como una señal del cambio de la actitud oficial en el sentido de que se cuestionaba el control exclusivo de los docentes sobre el currículum. Como es lógico, tanto los sindicatos docentes como muchas autoridades locales vieron amenazada la larga tradición de autonomía y de alguna forma u otra tomaron partido en el debate. En 1977, el gobierno británico publica el Green Paper, donde el término «*accountability*» se utiliza por primera vez en un informe gubernamental. Desde entonces, el concepto fue potenciado y ampliado por la administración conservadora de Margaret Thatcher con un creciente énfasis en los aspectos de control y centralización. En la práctica esto significó la neutralización creciente de las Autoridades Locales de Educación (LEA) así como la aprobación de un currículum básico obligatorio a nivel nacional (*Core Curriculum*).

El movimiento de la «accountability», sin embargo, no suscitó exclusivamente resistencias frente al control. A diferencia de los programas de tests de rendimiento y evaluación institucional aplicados en muchos estados federales de los Estados Unidos, en el Reino Unido la rendición de cuentas adoptó un gran espectro de formas, hecho que sin duda contribuyó a la creciente aceptación de su necesidad. Según McCormick y James (1996: 33) se pueden distinguir, al menos, tres formas:

- «1. *Capacidad de responder ante los propios clientes (answerability)*, o sea, los alumnos y sus padres (rendición de cuentas moral).
2. *Responsabilidad (responsability) ante uno mismo y sus compañeros* (rendición de cuentas profesional).
3. *Rendición de cuentas (accountability)*, en sentido estricto, ante los propios empleadores o dirigentes políticos (rendición de cuentas contractual)».

La rendición de cuentas contractual es la forma que, en sentido estricto, adopta la *accountability* en las relaciones establecidas entre la administración educativa y los centros docentes. Se trata esencialmente de una modalidad formal-legalista, prescriptiva, con una única dirección, de los centros a los responsables políticos o empleadores, y con un fuerte componente de control e implicaciones de sanción (reducción de presupuestos y formas de apoyo). Su aplicación, supuso que a partir de 1977 se reforzaran tres tendencias muy claras: *a)* el incremento del personal destinado a la inspección, *b)* la aplicación de programas de tests y pruebas estandarizadas para medir el rendimiento y *c)* el aumento de la participación y responsabilidad de los padres en los órganos colegiados (Open University, 1984).

Las instituciones escolares, por tanto, se van a ver obligadas a asumir el imperativo de la rendición de cuentas. Legitimándose dicha orientación, unas veces, con argumentos sobre la opacidad de los centros en cuanto al conocimiento público de lo que en ellos se hace, y, otras, planteando la necesidad de medir la rentabilidad de las inversiones realizadas mediante criterios de costo-beneficios. Pero paralelamente las formas de responder (ante padres y alumnos) y de responsabilidad (ante uno mismo y los compañeros), es decir, las formas de rendición de cuentas moral y profesional, son iniciativas que emergen desde los propios centros como alternativas a la de relación contractual.

La rendición de cuentas contractual adoptó también muy diversos formatos. En muy pocos casos se desarrolló lo que propiamente puede carac-

terizarse como una evaluación democrática. Más bien el punto de partida estaba en las iniciativas que tomaron las autoridades locales, pero de una forma en ocasiones mucho más indicativa que coercitiva. Un ejemplo representativo de esta postura es el folleto editado por las autoridades educativas de Londres (ILEA, 1977) donde se insiste que el listado de temas para la evaluación de cada centro escolar no debe ser tomado taxativamente, sino como una base de reflexión acerca del rendimiento de cada institución. Las respuestas individuales de los profesores tenían carácter voluntario y podían ser orales. En otros casos la evaluación de los centros estuvo impregnada por una filosofía productivista, donde la valoración de la escuela se medía en términos empresariales de eficacia y eficiencia en relación a los resultados, siempre discutibles, de un rendimiento cuantitativo.

En este sentido, las líneas de investigación y evaluación desarrolladas por el movimiento de las «Escuelas Eficaces» es representativa de las preocupaciones del momento por identificar y potenciar los elementos institucionales que caracterizan a los centros educativos con un mejor rendimiento escolar. En efecto, sobre todo durante su primera fase, el movimiento de la eficacia escolar (Davis y Thomas, 1992) se centró especialmente en relacionar los resultados con las características de los centros, tratando de indagar cuáles eran los aspectos diferenciales que definían a las instituciones con mayores niveles de logro; pasando después (Reynolds y otros, 1997), a plantearse las formas en que podrían generarse en los centros las condiciones para que pudiesen obtener dichos niveles. El retrato robot de este tipo de escuelas, en las que no puede perderse de vista su carácter predominante de centros privados y el hecho de pertenecer, generalmente, a sistemas educativos descentralizados, es sobre todo el de instituciones con un fuerte liderazgo del director, que maximizan los tiempos de aprendizaje centrados en el currículum de las materias escolares, con una participación amplia de los padres, una cultura colegiada del trabajo docente, etc.

El problema, no obstante, de este movimiento reside en la concepción instrumental de la escuela; el concepto de eficacia, centrado exclusivamente en los resultados, se sustenta sobre una calidad extrínseca (Carr, 1998), de modo que, aun cuando identifica esos factores de eficacia, su orientación está más dirigida a la rendición de cuentas que al desarrollo de estrategias de aplicación y acción, dado que no se profundiza en la naturaleza del proceso que genera y conduce a dicha eficacia. Como acertadamente señala Bolívar (1999: 24),

«la limitación más fuerte en la investigación sobre escuelas eficaces, al menos en su primera generación, es haberle importado —como único indicador de eficacia— los resultados cuantificables medios obtenidos por un centro escolar, sin haber entrado en los procesos de innovación que puedan promover la mejora de los centros ineficaces (o ‘fracasados’).»

En efecto, lo distintivo del momento actual son las condiciones externas desde las que se aborda la evaluación. El afianzamiento de las políticas neoliberales introduce cambios de gran calado en el discurso educativo. La educación pasa a ser más un bien de consumo que un derecho, en el juego de la oferta y la demanda del mercado, y a los actores sociales, más desde su papel de consumidores que de ciudadanos, se les invita a juzgar la «calidad» de la misma mediante parámetros e indicadores que resaltan su valor en términos económicos. Como señala Gimeno (1999: 19),

«los valores de justicia, equidad, dignidad humana, solidaridad y distribución de la riqueza y del capital se van sustituyendo por la preocupación por la eficacia, por la competitividad, la “excelencia”, la búsqueda de resultados tangibles, el ajuste a las necesidades del mercado de trabajo y de la economía, la lucha por disponer de mejores condiciones de salida del sistema educativo ante un mundo laboral escaso, la formación en destrezas básicas, la necesidad de incorporar las tecnologías de la información, etc.»

Y, de acuerdo con Escudero (1999: 14),

«[el] formalismo de los procesos de mejora permanente oculta la sustancia, pertinencia y relevancia de los contenidos a mejorar, y esconde, de ese modo, los criterios normativos que hayan de validarse y presidir la fiebre innovadora. En el discurso y práctica de la calidad total es improbable encontrar el currículum y la pedagogía por algún sitio. Mucho menos, aquellos conflictos e interrogantes que sobre el mismo viene formulando la sociología crítica, el análisis cultural y antropológico, el multiculturalismo, o las perspectivas políticas y micropolíticas en torno a las organizaciones».

Lo prioritario en estos casos no va a ser el desarrollo integral de la personalidad de los alumnos ni la formación de ciudadanos, sino la adquisición de habilidades, destrezas y conocimientos que se adecuen a las demandas de un mercado del trabajo que cada vez más acentúa su orientación pragmatista de la formación y que, a su vez, impone las directrices de ésta en función de sus necesidades concretas de cualificación. La trasposición

de la lógica del mercado a la escuela (Rodríguez Moreno, 1998), enfatizada tanto por los discursos de la excelencia como por los de la reestructuración, desde la perspectiva tecnológica, retoman la necesidad de elevar los niveles de logro de los alumnos, sin plantearse las condiciones desiguales de su acceso, y acentuando la diversidad de las ofertas de formación de acuerdo con los rendimientos previos alcanzados. Pero en cualquier caso no debería olvidarse, como muy acertadamente precisa Pérez Gómez (1998: 260), que la escuela

«solamente desarrollará una tarea propiamente educativa cuando sea capaz de promover y facilitar la emergencia del pensamiento autónomo, cuando facilite la reflexión, la reconstrucción consciente y autónoma del pensamiento y de la conducta que cada individuo ha desarrollado a través de sus intercambios espontáneos con su entorno cultural».

Y ello, sin duda, supone un planteamiento muy diferente a cuando se prioriza exclusivamente la dimensión del desarrollo profesional de los estudiantes.

Por otro lado, junto a la perspectiva económica que toma cuerpo en los aspectos de eficacia y eficiencia en el logro de los resultados, también se van poniendo en evidencia los rasgos políticos de las tendencias neoconservadoras que cuestionan y se replantean la ausencia en los currículos de determinados contenidos de la enseñanza de ciertas áreas y/o asignaturas, tal y como pusimos de manifiesto en el Tema 1. La mayor parte de estos planteamientos se vuelven a vivir en nuestro país, en este mismo año de 2013, con algunas de las propuestas de la polémica y retrograda Ley Wert, como el hecho de que la religión sea evaluable y se proceda a la ampliación del currículum común compartido por todas las Comunidades Autónomas, al objeto de «españolizar», según afirmaba expresamente el ministro de Educación a los alumnos catalanes.

Tampoco parece muy claro que las políticas de descentralización que delegan responsabilidades a las instituciones y profesores no logren lo contrario de lo que pretenden si tenemos en cuenta que, en definitiva, se siguen manteniendo y reforzando la centralización de los mecanismos de control. Gimeno (1994: 19) identifica, precisamente, al mercado laboral y a las prácticas de evaluación como dos mecanismos recentralizadores en las políticas de desregulación curricular, precisando que «la evaluación puede hacerse con muy diversos propósitos y centrarse en el diagnóstico de infinidad de problemas,

pero establecida como método de control curricular sólo sirve para difundir una idea pobre y poco matizada del rendimiento escolar».

Es verdad que el control, ahora, se hace de una manera más sutil mediante el establecimiento de los propósitos generales, la definición de los objetivos mínimos y la elaboración de los criterios de calidad. Pero estos aspectos son lo suficientemente importantes para determinar las políticas y orientaciones que el sistema educativo ha de asumir, tanto en lo que significa la presencia de los centros en el mercado de la educación, como en la definición de los currículos en cuanto a áreas, asignaturas y contenidos que enfatizan visiones y valores nacionales de carácter ideológico tradicional.

La autonomía de centros y profesores, por decreto (Contreras, 1999) o decretada (Bolívar, 1999), no deja de ser una manera limitada de plantear el protagonismo de las instituciones y de sus actores, ya que, por una parte, su margen de maniobra es mínimo en la capacidad de decidir sus trayectorias, pues sus cometidos son de carácter metodológico-instrumental pero no entran en la deliberación de los objetivos o propósitos, y, por otra, no se facilitan las condiciones para la creación de una cultura institucional en la que se primen la formas de participación y colaboración que desde proyectos pedagógicos concretos conduzcan al desarrollo profesional e institucional.

Como puede verse, estas nuevas formas de concebir e instrumentar la evaluación suponen un reforzamiento de las primeras políticas de la *accountability* de los años 70 y 80, que van ir adquiriendo a lo largo de la década de los noventa, eso sí, una gran proyección internacional. Su discurso va a ser recurrente desde los años setenta hasta la actualidad, y aún cuando sus argumentos tienen matices específicos, en distintos momentos o fases, su finalidad es inequívoca, la de valorar la productividad y rentabilidad de las instituciones educativas. A este respecto, parece evidente que el fomento de los primeros proyectos de evaluación, que empezó siendo una forma de comprobar la eficacia de los grandes programas norteamericanos de educación compensatoria en los inicios de los setenta, fue tomando cuerpo, en los años ochenta, con el movimiento de rendición de cuentas auspiciado por las primeras experiencias de las políticas neoliberales anglosajonas, y se generalizó al contexto internacional con las opciones y decisiones que la globalización económica ha venido concretando durante los últimos años.

1.2. La evaluación de centros desde la perspectiva de los procesos de mejora y desarrollo institucional

El problema de la accountability no reside exclusivamente en los mecanismos de control que normalmente conlleva, sino en los procedimientos burocráticos con los que se acomete y en las formas tecnocráticas que utiliza en sus evaluaciones. Desde hace tiempo se acepta que la educación considerada como un bien de naturaleza social debe someterse al escrutinio y debate público y, por tanto, se asume democráticamente la necesidad de la rendición de cuentas como base de la deliberación y participación en la misma de los distintos sectores y actores sociales. El desafío en este caso consiste, por un lado, en superar una visión reduccionista y sesgada de lo que significa la valoración de un servicio público como la educación, centrada frecuentemente en los productos y resultados, y, por otro, en aceptar que cada una de las instancias presentes en éste son corresponsables de su desarrollo por lo que todas ellas deberían ser objeto de evaluación mediante procesos que retroalimenten la toma de decisiones y que, además, analicen con rigor las consecuencias y efectos de las medidas adoptadas en cada momento.

Los centros educativos son el eje y el motor de los procesos de cambio. Pero no deberíamos perder de vista que la comprensión e interpretación de lo que ocurre en la vida institucional significa, antes que nada, aceptar que los centros son organizaciones complejas, construidas socialmente, con elementos y culturas comunes en tanto que comparten objetivos, procedimientos y normativas derivados de su compromiso educativo, pero, a la vez, con características singulares, idiosincráticas, fruto de las experiencias concretas en que se encuentra inmersa, de los procesos de interacción que se llevan a cabo tanto en el interior de las mismas como en el contexto en que actúan, de los procesos simbólicos y micropolíticos que impregnan su quehacer cotidiano, etc. Todos estos aspectos no se suman entre sí, sino que están directamente vinculados e interrelacionados de forma que constituyen una unidad o totalidad por lo que difícilmente pueden aislarse los unos de los otros.

Esta forma de entender los centros se opone a la consideración de los mismos como organizaciones racionales, instituciones instrumentalizadas para el logro de unas metas definidas desde instancias externas, planificadas de arriba abajo, y que lineal y pasivamente ejecutan sin interpretar y dotar de significado las directrices recibidas. Díaz de Rada (1996: XVIII) plantea, a este respecto, que aunque

«la visión instrumental de la enseñanza tiende a reclamar para sí la validez de una realidad objetiva, el punto de vista antropológico llama la atención sobre la existencia de una constante *mediación sociocultural*, intersubjetiva, de los sentidos, los valores y los fines de *todas* las formas de racionalidad [...] La racionalidad instrumental, que enfatiza la dimensión objetivadora y universalista de la experiencia, se ve cotidianamente interpretada por una racionalidad de diverso orden, a la que denominaré *convencional*, que pone el acento sobre la dimensión intersubjetiva y local de las interacciones concretas».

Es en ese espacio de juego, de encuentro entre la racionalidad instrumental y la racionalidad convencional, en el que los actores y las instituciones escolares reconstruyen los valores y significados que atribuyen a sus acciones y prácticas concretas desde las experiencias singulares en que están inmersas. Las trayectorias y caminos construidos por los centros docentes están condicionados, por tanto, por las prescripciones curriculares, planes, dimensiones organizativas, etc., pero se expresan también por las mediaciones que las culturas idiosincráticas de cada institución hacen en sus procesos de interpretación y acción.

Desde esta perspectiva, es posible pensar en la escuela como una comunidad de aprendizaje, como organizaciones capaces de aprender (Bolívar, 2000; Santos Guerra, 2000). Los centros docentes necesitan reforzar la dimensión institucional de sus quehaceres educativos, al tiempo que crean espacios compartidos de aprendizaje y desarrollo institucional. Es a través del conocimiento y toma de conciencia de sus avances pero también de sus deficiencias, contradicciones e incoherencias como las instituciones escolares pueden definir y reforzar cursos de acción alternativos en los que se vaya logrando una simbiosis entre sus quehaceres o realizaciones y sus proyectos o aspiraciones. El problema, no obstante, es en qué aspectos se ha de centrar el aprendizaje institucional.

El movimiento para la «Mejora de la Escuela», en contraposición al de las Escuelas Eficaces, representa en esta línea un intento por fortalecer la capacidad de aprendizaje institucional en los problemas que los centros educativos han de afrontar. En su proyecto más consolidado, que fue auspiciado por la OCDE (Proyecto Internacional de Mejora de la Escuela, ISIP), impulsó enfoques de autoevaluación escolar. Su visión más amplia de lo que es un servicio educativo, le lleva a considerar que la calidad no es sólo un problema de rendimientos o logros de los alumnos, sino que también se

identifica con las condiciones internas que propician el desarrollo de los centros. Según Hopkins (1996: 381) los estudios realizados han encontrado

«que existe una capacidad para el cambio en los centros que mejoran y que dicha capacidad les capacita para adquirir las características internas de eficacia, al mismo tiempo que se incrementa en ellos el rendimiento y resultados de los alumnos».

Sin embargo, como Hopkins et. al. (en Bolívar, 1999: 27-28), ponen de manifiesto las limitaciones más importantes de los procesos de capacitación de los centros para el cambio han sido: el de haberse centrado preferentemente en las relaciones y trabajo conjunto de los profesores, olvidando que las mismas deben sustantivarse en los problemas curriculares para que se dé una mejora en el aprendizaje de los alumnos, y que el propio modelo de proceso se desvirtuó en un enfoque técnico-racional de cambio.

Fullan (1994: 147) ha analizado con detalle algunos de los motivos del fracaso de la gestión basada en el centro como consecuencia de no haber abordado los «<<grandes temas>>», esto es, en ir más allá de la etiqueta de gestión basada en el centro hacia los cambios más básicos que son imprescindibles para que una reforma verdaderamente marque la diferencia». Y dichos cambios, según él, están relacionados con la cultura de la escuela y con la redefinición de los roles desempeñados por el profesorado y de sus procesos de formación. Con respecto a la cultura de la escuela parece claro la necesidad de trascender el individualismo docente, estableciendo modos de proceder colaborativos que permitan afrontar colectivamente la experimentación de proyectos y la construcción de un conocimiento profesional compartido. En ellos deben evitarse algunas de las dificultades y problemáticas, también señaladas por Hargreaves (1996), de la cultura de la colaboración como la colegialidad artificial y la balcanización. Y, por parte del profesorado, no son menos importantes los cambios que han de acometerse en sus conocimientos y competencias que, de ningún modo, pueden limitarse ya a los aspectos de enseñanza y aprendizaje, sino que han de comprender temas y problemas como la colaboración, el conocimiento del contexto de actuación, actividades continuas de autoformación, los procesos de cambio y los componentes morales de su ejercicio profesional.

La evaluación de los centros desde las perspectivas de los procesos de mejora y desarrollo institucional han fundamentado sus propuestas en el protagonismo que asumen las instituciones y los profesores en la teoría del cam-

bio. Helen Simons (1995: 221-222) nos recuerda algunas de las lecciones extraídas de la evaluación externa de la reforma curricular centralizada de cara a apoyar el desarrollo de las escuelas democráticas. Nuestra autora destaca, entre otros elementos, que la escuela es la unidad básica del cambio. Todos los aportes que lleguen a ella desde el exterior son importantes, pero es necesario prestar una atención especial a la institución en cuanto lugar en el que se produce el cambio y desde el que puede abordarse las prácticas educativas realizadas en las aulas. Siguiendo a Stenhouse, resalta la necesidad de aunar el desarrollo curricular con el desarrollo del profesorado. La concepción del «*profesor como investigador*» hunde sus raíces en el principio de que el cambio curricular no proviene de las prescripciones del currículum sino del análisis y estudio que los profesores hacen de sus prácticas profesionales.

La interrelación entre el cambio de la cultura institucional y del profesorado es, también para nuestra autora, una condición *sine qua non* para la creación de nuevos espacios y formas de trabajo en el que pueda darse un desarrollo profesional del profesorado. Estos son, de alguna manera, los que han de elaborar teorías sobre el contexto escolar en que trabajan. La escuela debe constituirse en una comunidad de aprendizaje en la que el profesorado, de manera colectiva y participativamente, sea capaz de analizar críticamente sus quehaceres profesionales, de contar con marcos teóricos y conceptuales elaborados por ellos mismos, y de fundamentar en sus teorías las experiencias que llevan a cabo. El profesorado, y no sólo los teóricos de la evaluación y del currículum, puede de esta manera producir teoría sobre el cambio y desarrollo de la escuela. El reto pasa porque las escuelas se adueñen de los procesos de cambio.

En suma, nos encontramos, como precisa Helen Simons (1995: 222-223)

«con un proceso que es, a la vez, *cultural*: el conocimiento del mundo de la escuela puede significar que tengamos que cambiar ese mundo; *psicológico*: sólo si nosotros, como individuos y como individuos en colectividad, queremos cambiar, lo haremos; y *político*: si queremos cambiar, necesitamos una estrategia política que haga participar a quienes pueden respaldar el cambio mediante la provisión de condiciones y apoyo adecuados. En cualquier evaluación o intento de reforma han de integrarse estos tres procesos».

Contreras (1999) recoge el énfasis que se ha hecho en el desarrollo individual e institucional como elementos de la capacidad de cambio (Fullan, 1998), y en el desarrollo de la escuela como totalidad (Rudduck, 1994), en la evolución

de las distintas épocas en las tendencias de los estudios sobre el cambio y las formas de innovación para acometer las reformas curriculares. Por ello, resalta

«la consideración del profesorado y de las instituciones escolares como mediadores activos en los procesos de cambio, de los centros como unidades y como culturas totales, del cambio como un proceso continuo de aprendizaje, y de los enseñantes como fuerzas individuales que pueden dirigir sus energías en cuanto que agentes activos, se han convertido en argumentos de peso para el apoyo de reformas en las que se reconozca la autonomía del profesorado y de los centros escolares».

Llegados a este punto, no cabe duda que la evaluación de las instituciones educativas tienen que ser, por congruencia con las características de estas organizaciones, propuestas que consideren de manera holística la complejidad de sus distintas realidades. A esto es, precisamente, a lo que se refiere Santos Guerra (1993: 14) cuando habla

«[aunque] la evaluación puede focalizarse en un efecto zoom de contextualización y focalización, es preciso tener en cuenta que el centro es una unidad y que todos los elementos, relaciones, actividades, etc., pueden comprenderse en el análisis complejo del entramado general de propósitos, acciones agentes, actitudes y medios».

La filosofía que subyace en este planteamiento no es otra que se evalúa para conocer, y se conoce para transformar y/o mejorar (Sancho Gil, 1992). El problema es como se da el paso y se vincula el conocimiento con la acción. Las estrategias externas, de carácter administrativo coercitivo, lo hacen a través de prescripciones o normativas que las instituciones y sus actores deben observar en el desarrollo de sus actuaciones. En cambio, de acuerdo con McCormik y James (1996: 55),

«[la evaluación en el contexto] del desarrollo profesional y de la mejora de la enseñanza hace referencia a la revisión de la práctica con el fin de diagnosticar problemas y desarrollar, implementar y evaluar soluciones o, en caso contrario, asegurarse de que el proceso se desarrolla de manera correcta».

2. ESTRATEGIAS PARA LA EVALUACIÓN DE INSTITUCIONES EDUCATIVAS

Vistas las perspectivas analizadas hasta aquí, parece claro que los propósitos con los que puede llevarse a cabo la evaluación de las instituciones

educativas son muy distintos entre sí. Ahora bien, cabría interrogarse si estas dos orientaciones son las dos únicas soluciones posibles y si no existen entre ellas y dentro de cada una otro tipo de aspectos diferenciales que nos permitan clarificar su coherencia y contradicciones. Un autor como Nevo (1998: 89-91), por ejemplo, considera que la evaluación centrada en la escuela debe tener presente los cinco siguientes principios:

1. Los alumnos y sus resultados no deberían ser el único objeto de evaluación escolar;
2. Los resultados o impactos no deberían ser los únicos factores observados cuando se evalúa un programa, un proyecto o cualquier otro objeto de evaluación en el interior de la escuela;
3. La evaluación escolar tiene que cumplir las funciones formativas y sumativas de la evaluación y debe proporcionar información para la planificación y el perfeccionamiento, así como para la selección, la certificación y la responsabilización;
4. Las necesidades de evaluación interna de una escuela estarán mejor atendidas por un equipo de profesores y otros educadores para quienes la evaluación es sólo una parte de sus atribuciones, apoyados por una formación y una asistencia técnica externa adecuadas;
5. La evaluación centrada en la escuela debería elaborarse a partir de una combinación de las evaluaciones interna y externa, pero se requiere una evaluación interna antes de que una evaluación externa sea eficaz.

Los principios señalados por David Nevo trascienden la visión reduccionista de los centros escolares que mantienen los enfoques de evaluación de productos o resultados. Su propuesta de integración de las funciones sumativa y formativa, así como la necesidad de que se acometan las modalidades de evaluación interna y externa, suponen una visión más holística de la evaluación de centros y una apuesta por el protagonismo del profesorado al defender que ésta es un componente más de su desempeño docente. Sin embargo, convendría precisar más algunos extremos que no siempre se hacen explícitos cuando se contraponen enfoques en el siempre problemático campo de la evaluación de centros.

Uno de ellos es la identificación entre rendición de cuentas y evaluación sumativa. Según la distinción hecha por Scriven (1967) y Stake (1981), y dejando a un lado la validez de estas distinciones, la evaluación sumativa es

aquella que se realiza una vez que se ha concluido el programa o acción que se estuviera llevando a cabo en un ámbito concreto. Implica, por tanto, un balance final de la experiencia vivida. Pero eso no conlleva que la descripción, el retrato o la valoración de esa situación tenga que limitarse a un solo aspecto o fenómeno de dicha experiencia. Y es en eso justamente en lo que cae la rendición de cuentas al seleccionar como único criterio de valoración de las instituciones el rendimiento de los alumnos. Cosa que no ocurre, por ejemplo, con aquel otro de tipo de propuestas como la evaluación diagnóstica que, aunque pueda calificarse como sumativa, abre el foco de la evaluación a los distintos elementos sobre los que se articula la vida institucional, esto es, al aprendizaje, la enseñanza y la organización.

Si la evaluación de productos o resultados tiene estas limitaciones, la alternativa pasa por complementar la misma con propuestas de evaluación de proceso dado su carácter formativo. Esta es al menos la solución que discursivamente dan no pocos autores al justificar sus enfoques de evaluación. Sin embargo, también en este aspecto es necesario analizar cómo se sustancian dichas propuestas con respecto a la naturaleza del objeto de evaluación. Porque puede ocurrir, y de hecho así se sucede, que se parta de una concepción de las instituciones educativas como organizaciones racionales orientadas a la consecución de fines explícitos, definidos normalmente desde instancias externas a los propios centros, y cuyo funcionamiento puede preverse y planificarse a priori. Bajo estas premisas, se asume que los centros escolares presentan una serie de regularidades, elementos o aspectos constitutivos de su organización unitaria, que los hace ser más semejantes que diferentes. Bastará, pues, con seleccionar, Beltrán y San Martín (1993: 18), «como elementos constitutivos del objeto a evaluar aquellos que, repitiéndose en todos o la mayoría de los casos, resultan dominantes, esto es, aquellos que, de entre el universo de los posibles, imponen al resto su voz unívoca (su univocidad)».

Esta visión mecanicista de la organización escolar que implícitamente defienden o está presente en algunos enfoques de evaluación de ¿procesos o productos? desconsidera la capacidad de los centros y sus actores para recrear e interpretar, a la luz de sus experiencias y autobiografías, tanto el espacio institucional en que se desarrolla la enseñanza como los desempeños profesionales de los profesores. Desde estas propuestas de evaluación lo que se intenta es valorar la eficacia de los centros en la consecución de lo logros que previamente han sido regulados y prescritos en todo lo relativo a sus planteamientos institucionales, organización y funcionamiento,

objetivos mínimos, etc. La concepción instrumental de la escuela y la estandarización de los procedimientos de evaluación conducen a un conocimiento formal de lo que ocurre en el interior de los centros, en cuanto a si se han cumplido o no con los distintos cometidos encomendados, pero se ignora casi todo de los procesos, problemas, contradicciones, conflictos, avances, etc. que se generan en los desafíos cotidianos y que, en definitiva, son los únicos aspectos que nos permitirían conocer cómo se construyen y asumen, individual y colectivamente, los proyectos pedagógicos de una institución.

La comprensión y conocimiento que se tienen de los centros, desde estos enfoques de evaluación que se basan en una imagen estática de las instituciones y que se fijan en elementos atomizados de su organización, son muy limitados. Y a ello contribuye, en gran medida, no sólo los propósitos y la focalización del objeto de evaluación sino también el que los criterios y valores sean casi siempre determinados por sus patrocinadores o responsables (Simons, 1999), y que se dejen al margen las necesidades e intereses de las demás audiencias concernidas (Guba y Lincoln, 1989) desde una pretendida neutralidad valorativa que se sustenta en el uso de una epistemología objetivista, un encuadre ético subjetivo-utilitarista y una posición política engarzada en la democracia liberal. Los patrocinadores, a la vez que destinatarios, de estas evaluaciones son obviamente las administraciones educativas.

Frente a estas posiciones, que han sido perfectamente documentadas por House (1994) en función de las premisas epistemológica, ética y política en que se basan, los enfoques transaccionales (Angulo, 1987, 1988, 1990) suponen una alternativa a las limitaciones que venimos analizando. La conceptualización de la evaluación como una actividad política (MacDonald, 1976) conduce al cuestionamiento de la neutralidad valorativa, y genera no pocos interrogantes acerca de qué valores son los que guían nuestra construcción del conocimiento. En una sociedad democrática, y entendiendo la evaluación como un servicio público (MacDonald, 1995), todas las personas, grupos e instancias interesadas en la evaluación deben tener la oportunidad de expresar sus necesidades, experiencias y opiniones, House y Howe (2000). Es más, la evaluación tiene necesariamente que abordar las vivencias, interacciones, encuentros, conflictos y negociaciones entre los sujetos y grupos que interpretan y recrean el quehacer institucional y personal en las transacciones que establecen entre ellos mismos y con los demás miembros de la comunidad. La voz de todos ellos es imprescindible antes, durante y después de llevar a cabo la evaluación.

Antes, porque desde el mismo momento en que se va a afrontar el estudio se están ya prefigurando las líneas y directrices que van a definir el proceso de evaluación. Tanto en sus propósitos como en la concepción y delimitación del objeto de evaluación, las soluciones serán muy diferentes si se parte de la aceptación de la pluralidad de valores y del reconocimiento de las instituciones escolares como espacios en los que concurren actores y sectores sociales con opciones ideológicas, intereses y posiciones normalmente enfrentadas y, por tanto, en conflicto. En estos casos, la labor del evaluador tiene que comenzar por conocer cuáles son las necesidades e intereses de las distintas audiencias con respecto al estudio, porque es a partir de ellas como la planificación y diseño de la evaluación va a ir tomando cuerpo al aflorar la mayor parte de los problemas y preguntas relevantes a las que se ha de responder (Stake, 1983).

Estos planteamientos y modos de proceder son los que singularizan a las perspectivas transaccionales de las experimentales en su acercamiento al objeto de evaluación. Sobre todo porque estas últimas siguen enfatizando en sus enfoques y análisis los aspectos metodológicos sobre los conceptuales, proponiendo en ciertos casos, como hace De Miguel (1997: 161), que para comprobar hasta qué punto las instituciones cumplen con el cometido de proporcionar a los alumnos la educación, que se prescribe en las normas y disposiciones legales, **«necesitamos especificar en objetivos medibles los criterios que determinan la calidad de la educación en sus distintas etapas y niveles»**. Lo cual no significa otra cosa que, desde una posición de consenso o neutralidad valorativa, los que pueden, la administración, y los que saben, los expertos, determinen y midan los resultados conseguidos por los centros.

La voz de todos los implicados debe ser escuchada, asimismo, durante el proceso de evaluación. Es verdad que tanto los diseños preordenados como los respondientes o de réplica planifican a priori los escenarios y sujetos de los que obtener información. Pero tampoco es menos cierto que mientras los segundos están abiertos a las circunstancias, dinámicas y situaciones que pueden emerger en un momento dado, los primeros, es decir, lo preordenados, por las características mismas de sus propuestas metodológicas obvian todas estas situaciones, ya que las mismas contaminan el control y la relación entre las variables que se desean estudiar.

A este respecto, la lógica del proceso evaluador es de nuevo sustantivamente distinta en las dos perspectivas. Para la experimental o no existe o

ignora todo aquello que previamente no ha sido identificado y preespecificado como variables del estudio, dado que los que se busca es encontrar el grado de asociación o la relación causa-efectos entre las mismas. Y, en cambio, para la perspectiva transaccional son precisamente este tipo de fenómenos no previstos, las consecuencias no anticipadas, las caras y facetas no visibles de la organizaciones, las que nos facilitan la comprensión del carácter vivo de las instituciones, la singularidad e idiosincrasia de cada centro que, lejos de encajar en un nítido y perfilado retrato robot, presenta zonas claroscuras y líneas de diverso trazado y grosor. Por ello, la evaluación para esta perspectiva no deja de ser un desafío, un proceso eminentemente emergente y difícil de planificar hasta sus últimos extremos.

3. LAS MODALIDADES DE EVALUACIÓN

Desde la perspectiva de los responsables de llevarla a cabo, puede hablarse de evaluación externa y evaluación interna. La primera corresponde a los tipos de estudios emprendidos por personas que no son miembros de las instituciones o de las comunidades de formadores que se evalúan, mientras que la segunda es aquella que se realiza por los mismos miembros o instituciones que son objeto de evaluación. Esta distinción, sin embargo, es tan formal que dice muy poco de la naturaleza e implicaciones de estos dos tipos de evaluaciones.

En principio, habría que aclarar que una gran parte de las evaluaciones externas son llevadas a cabo por personas que, aunque no sean miembros de los centros o comunidades docentes, sí mantienen, por el contrario, una relación directa con el sistema educativo del que dependen éstos. Es el caso, por ejemplo, de las evaluaciones conducidas por los servicios de formación o de las instituciones específicas de las administraciones públicas. Por ello, estas evaluaciones son formalmente externas con respecto a las instituciones y formadores, pero no en relación con los responsables de determinar las políticas sobre las que se organiza y desarrolla la formación. Y lo que está en juego, en este tipo de estudios, es la capacidad e independencia de las personas u organismos para que los propósitos que las guían trasciendan los intereses concretos o coyunturales de esas políticas.

Esto último es el motivo de que se hable de la evaluación externa-independiente como una modalidad alternativa. A diferencia de la anterior, son

estudios emprendidos por personas cualificadas que o bien no tienen conexión directa con los distintos ámbitos de la administración de que se trate, o bien disponen de una cierta autonomía funcional, que les posibilita en uno y otro caso actuar con criterios informados por sus conocimientos profesionales. Tal y como han recalcado Angulo, Contreras y Santos (1991, 1994), la evaluación externa-independiente se caracteriza, entre otras notas, por ser un servicio a la sociedad, realizada por evaluadores que no están directamente implicados en las circunstancias y desarrollos de la realidad evaluada, que responde a las exigencias sociales y demandas democráticas de información pública, cuyo propósito es asegurar el funcionamiento democrático de las instituciones y sus actores, así como mostrar la calidad del servicio desde perspectivas no tecnocráticas.

En la evaluación interna, los programas, instituciones o formadores asumen, a la vez, ser objeto y sujeto de su propio proceso de evaluación. Hoy suele hablarse más de autoevaluación que de evaluación interna, identificándose ambas como una misma modalidad de evaluación. Sin embargo, en este caso concreto, habría que diferenciar entre la evaluación interna que surge desde instancias ajenas a los propios programas y profesionales, y la autoevaluación que emerge desde el protagonismo, el compromiso y la madurez de las instituciones y formadores.

La evaluación interna auspiciada, asesorada y hasta dirigida por instancias ajenas a los diferentes responsables y participantes de la formación ha sido con frecuencia una forma de rendición de cuentas encubierta para complementar u obtener la información que no era posible lograr a través de la evaluación externa. Ante la resistencia opuesta por las instituciones y actores a ser evaluados de esta forma, cada vez más se ha ido fomentando una modalidad de autoevaluación en la que la capacidad de decisión de aquéllos es ciertamente limitada, ya que las propuestas de evaluación que han de llevar a cabo vienen perfectamente definidas y totalmente delimitadas hasta en sus más mínimos detalles, limitándose los actores e instituciones a cumplimentar los protocolos de recogida de información que les han sido remitidos.

Esta forma de proceder ha suscitado, por una parte, el cuestionamiento de este tipo de autoevaluación dado que cabe preguntarse ¿a qué y quién se refiere el «auto» de la autoevaluación? Y, por otra, que aumente las sospechas de si la autoevaluación de las instituciones, concebida de esta manera, no es sino una forma de rendición de cuentas con un rostro más humano.

La autoevaluación, por el contrario, que surge desde el protagonismo, el compromiso y la madurez de las instituciones y formadores es antes que nada un proceso de valoración crítica del quehacer de las instituciones realizado **en** y **desde** la propia institución. Se trata esencialmente de que los centros inicien y desarrollen un análisis autocrítico de su funcionamiento y trabajo habituales en donde se ponga en cuestionamiento las distintas formas de proceder en sus actuaciones y programas de formación.

Desde estas posiciones, es posible pensar en las instituciones como comunidades de aprendizaje, como organizaciones capaces de aprender. Los centros necesitan reforzar la dimensión institucional de sus quehaceres formativos, al tiempo que crean espacios compartidos de aprendizaje y desarrollo institucional. Es a través del conocimiento y toma de conciencia de sus avances pero también de sus deficiencias, contradicciones e incoherencias como las instituciones de formación pueden definir y reforzar cursos de acción alternativos en los que se vaya logrando una simbiosis entre sus quehaceres o realizaciones y sus proyectos o aspiraciones.

4. LA ESPECIFICIDAD DE LOS DISTINTOS ÁMBITOS DE EVALUACIÓN

La concepción del proceso de evaluación implica situar en primer plano el objeto que se va a evaluar. Y en principio no es lo mismo referirse a un *ámbito individual*, es decir a la evaluación de los actores de la formación (participantes, formadores, directores, supervisores), que a un *ámbito institucional* o singular (evaluación de instituciones, de planes o proyectos de formación o materiales didácticos), o que a un *ámbito sistémico* (evaluación de políticas de formación o de sistemas de formación). Cada uno de estos espacios tiene su especificidad, sus propios elementos de identidad, y son difícilmente intercambiables entre sí y, cuando se hace, puede dar lugar a entremezclar conceptos que aunque relacionados son diferentes.

Ante cada uno de estos ámbitos surgen distintos problemas que exigen respuestas concretas para su evaluación. Entre otros, no pueden dejarse de tener en cuenta las distintas consideraciones y diferentes concepciones que pueden manifestarse sobre aquello que queremos evaluar; las diferencias más que apreciables que existen entre la evaluación del personal (formadores, profesores, participantes) y la de las instituciones, actuaciones o planes

de formación; la especificidad de los criterios que se utilizan para valorar cada uno de ellos; las metodologías que han de ponerse en juego para poder conocerlos, etc.

Por lo tanto, la referencia al objeto de evaluación significa fundamentalmente la selección de los elementos o aspectos sustantivos, en cualquiera de los ámbitos anteriores, que se necesitan o es deseable conocer. Esto supone, de alguna manera, una opción entre las diferentes formas de entender lo que son los procesos de formación, el desempeño profesional de los formadores o las características de las instituciones como organizaciones de formación. Es decir, que, además de por el tipo de objeto, la concepción del propio proceso de evaluación está influido y hasta determinado por los modos de entender, ver o seleccionar la perspectiva desde la que se aborda dicho objeto. Y, por tanto, nos lleva a reafirmar que la evaluación no es única ni exclusivamente un acto técnico, pese a la ansiedad de muchos por disponer de instrumentos de evaluación y a la abundancia de textos que se limitan a recoger repertorios de instrumentos y técnicas evaluativas. Si algo está meridianamente claro es que la evaluación no es una actividad atórica, ya que hasta, y sobre todo, en los planteamientos más pragmáticos y empiristas de evaluación subyacen concepciones del objeto evaluado.

Por ejemplo, tal y como hemos analizado en el Tema 2, en el contexto anglosajón a la hora de clarificar los posibles significados del concepto de evaluación se hace referencia a los términos *assessment* y *evaluation* para diferenciar dos formas distintas de entender la misma. Si se opta por el primero es evidente que, con respecto al objeto de evaluación, se entiende que la calidad de un centro puede evaluarse a través del conocimiento de su impacto en los usuarios, es decir, teniendo en cuenta sólo los resultados escolares o aprendizajes de los alumnos. En cambio, si nos identificamos con el segundo implicaría que la evaluación se concibe como el conocimiento de la calidad del servicio educativo prestado y, consecuentemente, ha de abarcar todos los componentes que influyen en los procesos de enseñanza y aprendizaje, además de los resultados escolares.

5. METODOLOGÍA DE EVALUACIÓN

En los apartados 1.1, 1.2 y siguientes del tema 2 se han planteado los distintos enfoques, modalidades, concepciones de las instituciones y significaciones de la evaluación. Todos ellos, junto con las diferentes posiciones

epistemológicas y metodológicas que fueron analizadas en el mismo tema 2 están presentes a la hora de la elaboración de un proyecto de evaluación institucional.

El evaluador se sitúa ante el objeto de evaluación tratando de dilucidar cómo puede llegar a conocer su naturaleza y/o valor. Es obvio que la forma en que se aborda dicho objeto es ya diferente, desde el principio, si el proyecto se acomete como una propuesta de autoevaluación o de evaluación externa. En el primer caso, son los miembros de la institución quienes deciden los asuntos sobre qué, cómo, cuándo y para qué evaluar a partir de las necesidades, preocupaciones o preguntas que se hagan o tengan sobre la trayectoria institucional —con anterioridad hemos visto cómo puede surgir o generarse la evaluación como consecuencia de la problematización de ciertos cometidos o desempeños profesionales e institucionales—. Por el contrario, en un proyecto de evaluación externa serán los responsables de llevar a cabo el mismo los que, en función de las demandas y necesidades de las diversas audiencias, determinarán los elementos sustantivos del mismo. En ambos casos el proyecto irá perfilándose a medida que vayan tomándose todo un conjunto de decisiones.

Convendría recordar, para un mejor entendimiento de nuestra propuesta, que cuando hablábamos de metodología precisábamos que nos referíamos al contexto y lógica de uso de los métodos. Pues bien, una de las decisiones que ha de tomarse es el tipo de diseño por el que se va a optar en el proyecto de evaluación institucional. Esta decisión exige considerar a la vez las características del objeto y lo que se desea o quiere conocerse del mismo ya que será lo que determine la metodología de evaluación. El uso de un diseño preordenado implica, entre otros requisitos, que es el evaluador el que, desde su conocimiento experto, determina los hechos que necesitan conocerse; que los hechos son los elementos centrales de la evaluación al margen de las percepciones y valores de los participantes; que el diseño del estudio, una vez establecido, no debe modificarse para que los resultados no resulten contaminados por fenómenos o circunstancias no contempladas en la relación entre las variables iniciales que han sido seleccionadas o controladas; que las variables han de ser susceptibles de cuantificación y medición; que los resultados se expresan generalmente a través de datos numéricos, etc.

En cambio, si en una propuesta de evaluación institucional como la que hacemos aquí, en la que se pretende conocer y valorar tanto los procesos

como los resultados, se opta por un diseño emergente en vez de por uno preordenado parece claro que, entre otros supuestos, el evaluador debe tener en cuenta a las audiencias a la hora de determinar lo que se necesita conocer; que no existe una separación entre los hechos y valores de los participantes; que el diseño del estudio es flexible y está abierto para ir incorporando las dimensiones personales e institucionales, no planificadas previamente, pero que van surgiendo a lo largo del estudio y que son importantes para complementar y enriquecer el conocimiento del objeto de evaluación; que las variables no tienen que ser necesariamente mensurables; que los resultados pueden expresarse no sólo a través de datos numéricos sino también mediante el discurso y la voz de los miembros de la institución, etc.

La lógica y contexto de uso de uno u otro diseño supone que el evaluador se posiciona ante el objeto de evaluación desde planteamientos ontológicos, epistemológicos, metodológicos y ético-políticos distintos. La metodología, en este caso, no se puede desvincular del resto de los elementos considerados ni actúa por separado, ya que en sus distintos modos de proceder para llegar al conocimiento están subsumidas la concepción de la realidad, la forma de asegurar que conocemos lo que afirmamos conocer, así como desde los valores e instancias que se construye el conocimiento. Piénsese, por ejemplo, las diferencias que se ponen de manifiesto cuando se utilizan las metodologías denominadas, de un modo un tanto genérico, experimental y transaccional. Otra cosa muy distinta son las técnicas e instrumentos para la recogida de información en las que, pese a su adscripción a unas u otras metodologías, cada día es más frecuente optar por el uso combinado de las mismas en correspondencia con la magnitud del objeto de evaluación.

Cualquier proyecto de evaluación ha de emprender, en su elaboración, una labor sistemática de análisis y selección de los elementos más relevantes de su objeto de estudio. En este proceso se juega una parte considerable de la calidad del proyecto, ya que se ha de clarificar y discernir lo que realmente es significativo para el conocimiento y valoración de la institución.

6. CRITERIOS E INDICADORES DE EVALUACIÓN

Existe un cierto consenso en que la evaluación conduce a un juicio razonado, ya sea de los logros obtenidos, del mérito o valía de algo, o de sus

consecuencias o efectos. Y, a este respecto, no deberíamos olvidar que la evaluación como valoración es esencialmente una actividad comparativa. La formulación de un juicio de valor exige que se disponga de una serie de referentes ideales, normas o criterios con los que contrastar lo que se está evaluando a los efectos de estimar o apreciar hasta qué punto o de qué manera dicho objeto posee o reúne las características deseadas. Existen, tal y como recoge Tiana (1998), tres posibles marcos de comparación, a saber: la que se lleva a cabo con respecto a un criterio o norma construidos a partir de las condiciones o requisitos que debe cumplir el objeto evaluado; la que se realiza sincrónicamente comparando al objeto con otros objetos de características semejantes; y aquella otra en que la comparación del objeto se hace consigo mismo en distintos momentos temporales, es decir con una perspectiva *diacrónica o longitudinal*.

Con relación al primer marco de comparación, los criterios que pueden utilizarse pueden provenir a través de tres fuentes: de su mérito, de su valía y de las expectativas de las distintas audiencias. El mérito normalmente se asocia con el valor intrínseco mientras que la valía se vincula a su valor contextual, de uso o extrínseco. De esta manera, se puede hablar del mérito de un plan, o del desempeño de un profesional, por la calidad o excelencia de los mismos en relación con otras propuestas de la misma naturaleza, y de la valía por su valor como la solución más adecuada a las características o problemas a los que intenta dar respuesta. Es decir, que en el primer caso la valoración se hace teniendo en cuenta exclusivamente sus atributos al margen de la realidad en que se va o esté aplicando, mientras que en el segundo su valor se pondera por su viabilidad y adaptabilidad a las características del contexto en el que se ha de desarrollar o esté ya implementándose.

Las expectativas de las distintas audiencias se refieren a que los criterios de evaluación deben recoger los intereses y necesidades de todos aquellos que están concernidos por el estudio. Robert Stake (1967) promovió uno de los más importantes cambios en la teoría y práctica de la evaluación al plantear que los estudios no podían estar exclusivamente al servicio de los patrocinadores sino que debían responder a las demandas y exigencias de la comunidad. En su propuesta de evaluación de réplica, Stake (1983: 96) supuso el paso de una tradición de evaluación gerencialista, definida por y para los patrocinadores, a una evaluación democrática, en la que se considera las exigencias de las audiencias y las distintas perspectivas de valor de las personas involucradas.

Los criterios de los marcos segundo y tercero de comparación no tienen por qué ser distintos a los del primero, ya que en uno de los casos lo que se compara entre sí son a dos objetos de parecidas características (actuaciones, planes o instituciones de formación) y, en el otro, lo que se va a valorar es la trayectoria de un mismo objeto en distintos momentos temporales.

Por otra parte, convendría tener en cuenta que para que se pueda llevar a cabo la valoración de cualquier fenómeno o elemento de la formación, además de tener claros cuáles son las condiciones y requisitos que han de poseer (criterios), es preciso dilucidar de dónde y de qué obtendremos información para conocer las características de los mismos. Pues será a través de la disponibilidad de datos e informaciones como los evaluadores van logrando indicios y evidencias de cuáles son las peculiaridades, o el modo en que se manifiesta, o las consecuencias o efectos de una determinada acción.

Uno de los primeros retos, por tanto, al que han de enfrentarse los responsables del estudio toma cuerpo cuando se interrogan acerca ¿de qué características, elementos o circunstancias deben y pueden obtener evidencias que les permitan conocer la totalidad o parte del ámbito que desean evaluar? Ante esta pregunta van surgiendo, poco a poco, respuestas que ayudan a ir concretando el objeto de evaluación.

El trabajo de delimitación y concreción del objeto de evaluación se lleva a efecto mediante la selección de lo que se denomina, de manera casi indistinta, como dimensiones, subdimensiones e indicadores de evaluación. Decimos que las denominaciones son «casi» equivalentes porque es verdad que, de un modo un tanto genérico, unas y otras se refieren a los elementos o componentes de una actuación, de un plan o de una institución de formación en los que se va a centrar la recogida de información y datos. Pero, no es menos cierto también que el uso predominante que suele darse actualmente al término indicador es equivalente al de subdimensión, en tanto que son más concretos y precisos, que el que tradicionalmente ha tenido y sigue teniendo el de dimensiones de evaluación.

Las dimensiones, por lo general, se refieren a los elementos o componentes estructurales de los distintos ámbitos de evaluación, mientras que las subdimensiones e indicadores son aquel conjunto de aspectos a través de los que podemos caracterizar cada una de las dimensiones que conforman un ámbito de evaluación cualquiera. Aunque tanto los indicadores

como las subdimensiones pueden ser expresiones de los procesos o de los resultados de la formación, aquéllos utilizan usualmente datos de tipo cuantitativo mientras que éstas combina información cualitativa y cuantitativa.

En los últimos años y, sobre todo, a raíz de la proliferación de estudios internacionales se ha ido extendiendo el uso de los indicadores de evaluación con la finalidad de lograr una información más objetiva y fiable en la comparación de la situación de los servicios públicos en distintos contextos nacionales. Normalmente los indicadores hacen referencia a series temporales de datos que tienen un carácter unívoco y homogéneo en diferentes situaciones y contextos y que, por ello, ponen de manifiesto los cambios que se producen en las dimensiones de evaluación. El IPC (índice de precios al consumo) es el ejemplo más claro de lo que es un indicador. A través de él los ciudadanos disponemos de una serie de datos a lo largo del año para valorar el encarecimiento o decrecimiento del coste de la vida en una comunidad o país, así como para comparar la misma entre diferentes comunidades o países.

En el Tema 2 hemos dedicado el apartado 4 al análisis de las características, tipos y modelos de indicadores. Asimismo se han aclarado las relaciones entre los modelos de indicadores y los modelos de escolaridad, y entre los indicadores y la toma de decisiones. Remitimos, por tanto, a los lectores al Tema 2 para refrescar y completar todas estas cuestiones.

7. A MODO DE CONCLUSIONES

La evaluación de centros se puede afrontar básicamente desde distintas perspectivas, orientaciones y estrategias. Las perspectivas de la rendición de cuenta y la de los procesos de mejora y desarrollo institucional se presentan como distintas e incompatibles, no tanto porque una y otra orientación no puedan ser complementarias, sino porque con cierta frecuencia a la primera solo le interesan los resultados y la eficacia institucional, mientras que la segunda se limita a estudiar y evaluar los procesos. Pero como se tendrá oportunidad de ver en el tema 4, la auténtica evaluación de centros es aquella que considera tanto los procesos y resultados. Sobre todo porque siendo el aprendizaje del alumnado un elemento muy importante para cualquier institución, no cabe duda que pierde toda su relevancia si se

desconocen los factores y actuaciones que, en definitiva, facilitan o impiden el aprender. Sólo desde el conocimiento de qué es lo que incide positivamente en el rendimiento escolar se puede afrontar la mejora, innovación y desarrollo de las instituciones y de sus principales actores, alumnos y profesores.

Estos aspectos son realmente pertinentes para que, ante un proyecto de evaluación, se tengan elementos de juicio con los que valorar su mérito y valía, así como discriminar la coherencia y contradicciones que pueden aparecer en el mismo. Es necesario tenerlos también muy en cuenta a la hora de analizar los proyectos y realizaciones que se llevan desde las diferentes modalidades de evaluación, así como desde las distintas instancias promotoras.

BIBLIOGRAFÍA

- BECHER, T. y MACLURE, S. (1978). *Accountability in education*. Introducción. NFER. London, Publishing Company.
- BELTRÁN, F. y SAN MARTÍN, A.(1993): «La organización escolar: ¿evaluación o devaluación?», *Cuadernos de Pedagogía*, n. 219, pp.16-21.
- BOLÍVAR BOTÍA, A. (1999). *Cómo mejorar los centros educativos*. Barcelona, Síntesis.
- (2000). «¿Pueden los centros educativos ser organizaciones que aprenden?», *Organización y Gestión Educativa*, n. 3, pp. 35-39.
- CARR, W. (1998). «¿Qué se entiende por “calidad” en la enseñanza?», en W. Carr, (Ed.): *Calidad de la enseñanza e investigación-acción*. Sevilla, Díada Editorial, 3ª edición revisada, pp. 5-24.
- CONTRERAS DOMINGO, J. (1999): «¿Autonomía por...» *opus. cit.*
- DE MIGUEL DÍAZ, M. DE (1997): «La evaluación de los centros educativos: una aproximación a un enfoque sistémico», *Revista de investigación educativa*, 15, n. 2, p. 145-178, (las negritas son del autor).
- DÍAZ DE RADA, A. (1996): *Los primeros de la clase y los últimos románticos. Una etnografía para la crítica de la visión instrumental de la enseñanza*. Madrid, Siglo XXI.

- ESCUADERO MUÑOZ, J.M. (1999). “Calidad de la educación: entre la seducción y las sospechas”, *HEREUSIS* (<http://www2.uca.es/HEURESIS>), 2, n. 5.
- FULLAN, M. G. (1994): «La gestión basada en el centro; el olvido de lo fundamental», *Revista de Educación*, 304, pp.147-161.
- (1998): «The meaning of Educational Change: A Quarter of a Century of Learning», en Hergreaves, A. et. al. (eds.): *International Handbook of Educational Learning*. Kluwer, Dordrecht, pp. 214-228.
- GIMENO SACRISTÁN, J. (1999). «Políticas y prácticas culturales en las escuelas: los abismos de la etapa posmoderna», *HEURESIS* (<http://www2.uca.es/HEURESIS/>), 2, n. 1.
- HARGREAVES, A. (1996): *Profesorado, cultura y posmodernidad (Cambian los tiempos, cambia el profesorado)*. Madrid, Morata.
- HOPKINS, D. (1996): «Estrategias para el desarrollo de los centros educativos», en ICE Universidad de Deusto (Ed.): *Dirección Participativa y Evaluación de Centros. II Congreso Internacional sobre Dirección de Centros Docentes*. Deusto, ICE de la Universidad de Deusto, pp. 377-402.
- HOPKINS, D.; AINSCOW, M.; FIELDING, M.; SOUTHWORTH, G. y WEST, M. (1996): Papers presented at a *Symposium improvement and school effectiveness: Towards a new synthesis*. Annual Meeting of the American Educational Research Association. New York
- HOUSE, E. R. (1992): «Tendencias en evaluación», *Revista de educación*, n. 299, p. 43-55
- (1998): “Acuerdos institucionales para la evaluación”, *Perspectivas*, vol. XXVIII, n. 1, marzo, pp.123-131.
- INNER LONDON EDUCATION AUTHORITY (1977): *Keeping the school under review*. London, I.L.E.A
- MCCORNICK, R. y JAMES, M. (1996). *Evaluación del currículum en los centros escolares*. Madrid, Ediciones Morata.
- NEAVE, G. (1988). «On the cultivation of quality, efficiency and enterprise: an overview of recent trends in higher education in Western Europe, 1968-1988», *European Journal of Education*, 23(1/2), pp. 7-23.

- (2001): *Educación superior: historia y política. Estudios comparativos sobre la universidad contemporánea*. Barcelona, Edit. Gedisa,
- NEVO, D. (1998): «La evaluación mediante el diálogo: una contribución posible al perfeccionamiento de la escuela», *Perspectivas*, XXVIII, n. 1, marzo, pp. 87-100.
- OPEN UNIVERSITY (1984). *Accountability and evaluation*. Milton Keynes, The Open University Press, pp. 11-14.
- RODRÍGUEZ MORENO, M. MAR (1998): «El cambio educativo y las comunidades discursivas: representando el cambio en tiempos de postmodernidad», *Revista de educación*, n. 137, pp. 157-184.
- RUDDUCK, J. (1994): «Reflexiones sobre el problema del cambio en las escuelas», en J. F. Angulo y N. Blanco, (Comp.): *Teoría y Desarrollo del Currículum*. Málaga, Ediciones Aljibe, pp. 385-393.
- SANCHO GIL, J. M. (1992): «Evaluar, conocer, transformar, mejorar: La evaluación del centro y el desarrollo del criterio profesional», *Aula de innovación educativa*, 6, septiembre, pp. 47-51.
- SANTOS GUERRA, M. A. (1993): «Los tres vértices del triángulo: evaluación de centros, perfeccionamiento profesional y mejora de la práctica», *Aula de innovación educativa*, 20, noviembre, pp.14-19.
- (2000): *La escuela que aprende*. Madrid, Morata.
- SIMONS, H. (1995a): «La autoevaluación escolar como proceso de desarrollo del profesorado: en apoyo a las escuelas democráticas», en AA.VV.: *Volver a pensar la educación*. Madrid, Edic. Morata, Vol II, pp. 220-242.
- STAKE, R. (2006). *Evaluación comprensiva y evaluación basada en estándares*. Barcelona, Graó.
- G. WHITTY,; S. POWER y D. HALPIN, (1999): *La escuela, el estado y el mercado*. Madrid, Morata, pp. 55-56.

Tema 4

La evaluación de centros promovida por las administraciones educativas

Introducción

1. Evaluación y administración
2. La evaluación externa de centros desde las iniciativas de las administraciones
 - 2.1. La evaluación de productos o rendimientos de los alumnos en el contexto de la evaluación de centros educativos
 - 2.2. La evaluación de procesos y productos o evaluación de centros escolares
 - A) Objeto de evaluación
 - B) Criterios de evaluación
 - C) Procedimientos y utilización de los resultados de la evaluación
 - 2.3. Propuestas y realizaciones españolas
3. Estrategias de evaluación interna y combinada de evaluaciones interna y externa
4. A modo de conclusiones

Bibliografía

TEMA 4

LA EVALUACIÓN DE CENTROS PROMOVIDA POR LAS ADMINISTRACIONES EDUCATIVAS

Eustaquio Martín Rodríguez.

UNED

INTRODUCCIÓN

Las instituciones educativas han sido desde siempre, y especialmente desde la institucionalización de los sistemas educativos en el siglo XIX, objeto de evaluación. En la mayoría de los casos se trataba de un tipo de evaluación, en cuanto a sus propósitos y formato, muy diferente de la que se va a ir estableciendo y conformando a partir de la década de los años sesenta del siglo XX. Hasta entonces, la evaluación siempre había estado asociada al control e inspección de los centros sobre el cumplimiento de la ley, ya fuera para la acreditación de las instituciones o para comprobar si éstas observaban, en sus dimensiones de aprendizaje, enseñanza y organización, las prescripciones y normativas emanadas de la administración educativa de cada momento.

Este tipo de evaluación con un acusado componente de control, que fue y sigue siendo una de las funciones de los servicios de supervisión e inspección de las administraciones educativas nacionales, se va a complementar con nuevas propuestas de carácter más técnico de cara a valorar sistemáticamente los rendimientos de los centros educativos. En ellas no se pierde de vista tampoco la dimensión de control, la evaluación siempre se asocia quiérase o no al control, pero ahora se toma conciencia de que para que las reformas e innovaciones curriculares y organizativas trasciendan el plano prescriptivo o normativo con el que se presentan y puedan hacerse realidad, es decir, puedan implementarse, es necesario el compromiso y trabajo de las instituciones y sus actores.

La evaluación de centros propiamente dicha es la que considera y aborda a la vez los procesos y resultados. En el contexto europeo es un tipo de evaluación que se ha ido desarrollando poco a poco y, en ocasiones, como ha puesto de manifiesto la investigación de Buisson-Fenet, H. et Pons, X. (2011), la parte doctrinal es la que ha prevalecido sobre la experimentación y puesta en práctica de los proyectos y propuestas nacionales. Las dificult-

tades para su desarrollo provienen tanto de la resistencia de las instituciones al cambio de los modelos como de las derivadas del tiempo y dedicación necesarias para su realización. Los costes económicos y la falta de personal cualificado son también factores que no pueden dejarse de considerar.

Tal vez como consecuencia de estas dificultades, la evaluación de productos, es decir, la de los aprendizajes de los alumnos en las áreas y/o asignaturas nucleares del currículum ha ido ocupando el protagonismo estelar. De hecho, cada vez más la evaluación externa de los centros aborda a través de pruebas de nivel, de los resultados de los exámenes nacionales y de las calificaciones de los estudiantes, la valoración del rendimiento escolar del alumnado, proyectándolo como una dimensión relevante de la evaluación de centros. Es más, teniendo como base la evaluación de productos, de inequívoca tradición psicométrica, han surgido nuevas propuestas como el modelo de evaluación de valor añadido que en el contexto anglosajón, especialmente en Estados Unidos y en menor medida en Reino Unido, tiene cierta importancia

Por contra, el conocimiento de los procesos se ha ido dejando de lado, o se lleva a cabo a través de la modalidad de evaluación interna o autoevaluación, aunque cada vez con menos capacidad y posibilidades de los actores e instituciones para decidir por sí mismos los elementos básicos de los proyectos.

En cualquier caso, la documentación de Euridyce sobre la evaluación de los centros escolares europeos de educación obligatoria es sumamente interesante por la variedad de propuestas y de experiencias que se recogen. Animamos a los lectores a que manejen y lean los informes de los años 2002 y 2004.

1. EVALUACIÓN Y ADMINISTRACIÓN

Las administraciones educativas justifican sus iniciativas de evaluación de centros en la necesidad de contar con información relevante desde la que fundamentar sus políticas sobre los sistemas de enseñanza. Prácticamente en todos los países desarrollados, y en la mayor parte de los países en vías de desarrollo, a partir de la década de los 70, durante los años 80, y princi-

pios de los 90, se van a ir creando instituciones y elaborando propuestas y planes para la evaluación de las instituciones educativas. En la mayoría de los casos se presentan como evaluaciones externas por ser realizadas por personal ajeno a los centros escolares, aunque las personas o servicios que las llevan a cabo están vinculados normalmente a las administraciones educativas como sucede con los inspectores, servicios de supervisión y centros o institutos de evaluación. La tipología de estos estudios suelen coincidir con las evaluaciones de producto y de proceso o con formas combinadas en que se contemplan ambos.

Con respecto a las modalidades hay que señalar que, a raíz de las experiencias anglosajonas del movimiento de la *accountability*, la evaluación interna o autoevaluación se ha ido generalizando como una propuesta menos invasiva y coercitiva de valorar la trayectoria y logros de las instituciones escolares. Como ya se ha señalado y, más tarde volveremos a retomar aquí y en otros temas, la evaluación interna o autoevaluación fue y sigue siendo, en determinadas experiencias y realizaciones, una forma encubierta de rendición de cuentas, dadas las limitaciones presupuestarias y de personal con las que las administraciones se enfrentan para realizar evaluaciones externas de todos los centros del sistema educativo. Pero, también, es verdad que la autoevaluación que surge a iniciativa de las instituciones y de sus actores es una herramienta para abordar el desarrollo profesional e institucional, para que la comunidad educativa conozca sus logros y deficiencias y se aborden propuestas de mejora de sus actuaciones, así como un elemento básico del diagnóstico intrainstitucional para acometer tanto el desarrollo y las adaptaciones del currículo como la optimización de la organización y funcionamiento del centro.

Teniendo en cuenta las dificultades y limitaciones que presentan por separado las evaluaciones interna y externa, cada vez han ido tomando más impulso las iniciativas y proyectos que tratan de generalizar estrategias que combinan ambos tipos de modalidades. Consideradas individualmente, sobre la evaluación interna siempre recae la sospecha de su falta de objetividad y rigor, de sus tendencias autojustificadoras y endogámicas; de la evaluación externa, por el contrario, se recela de sus actuaciones vinculadas al control de las instituciones, de intervenciones demasiado instrumentalizadas y ajenas a las características singulares de cada centro educativo. Por ello, mediante el uso combinado de ambas modalidades se pretende superar dichas sospechas y recelos tratando de que haya una mayor aceptación

de las evaluaciones externas y de que las evaluaciones internas, a su vez, contribuyan a enriquecer las aportaciones de las externas.

Los responsables de llevar a cabo las evaluaciones externas en la mayoría de los países europeos, según se recoge en el estudio monográfico realizado por la Red Europea de Información sobre la Educación en Europa (EURYDICE, 2004), son generalmente los inspectores en todo lo que se refiere a los aspectos educativos y de una parte de la gestión. Junto a los cuerpos de inspectores, y en lo que concierne al control de otros aspectos concretos como la gestión financiera de los centros o de otras áreas más precisas como la seguridad o la documentación (archivos), también participan personal de distintas categorías dependientes de distintos organismos administrativos a escala central o regional. [El texto completo de la versión de este informe ha sido traducido al castellano, Secretaría General Técnica del MEC (2004), y puede encontrarse en <http://www.eurydice.org>, o en <http://www.mec.es/cide/>].

Como ya se ha señalado en el Tema 1, en España es en la Ley General de Educación de 1970 donde se introduce, por primera vez, el concepto de evaluación, proyectándose su uso no solamente para el aprendizaje de los alumnos sino también, de una manera más ambiciosa, referida al sistema educativo. Se establece, asimismo, tanto en la propia Ley como en el Decreto 664/1973, de 22 de marzo, que una de las funciones de la inspección será la de «Evaluar el rendimiento educativo de los centros docentes y profesores de su zona respectiva o de la especialidad a su cargo en colaboración con los Institutos de Ciencias de la Educación».

No es, sin embargo, hasta finales de los años 80 cuando, con la promulgación del Real Decreto 1524/1989, de 15 de diciembre, se vuelve a retomar y definir las funciones del Servicio de Inspección Técnica de Educación (SITE). Concretamente en el artículo 2, apartado tres, de este texto se recoge como una de sus funciones:

«[la de] Evaluar el rendimiento educativo del sistema, a través del análisis de la organización, funcionamiento y resultados de los Centros docentes y servicios, así como de la ejecución y desarrollo de los programas y actividades de carácter educativo, promovidos o autorizados por el Ministerio de Educación y Ciencia».

Cometidos, estos últimos, que también aparecen reflejados, en la Orden de 27 de septiembre de 1990, al fijar las labores del servicio Central de

Inspección en lo relativo a su «colaboración con los Servicios Provinciales de Inspección en el desarrollo y aplicación de programas o actividades de Evaluación de Centros y servicios».

La Ley 1/1990, de 3 de Octubre, de *Ordenación General del Sistema Educativo* (LOGSE) puso las cosas en su sitio, estableciendo ahora de una manera expresa, en su artículo 61 —apartado 2, subapartado b—, que una de las funciones de la inspección es la de «participar en la evaluación del sistema educativo». Y para disipar cualquier posible duda recogía, en el artículo 62 —apartado 3—, que la «evaluación del sistema educativo se realizará por el Instituto Nacional de Calidad y Evaluación».

Doctrina, esta última, que será refrendada en los distintos documentos y disposiciones legales que irán apareciendo en los años sucesivos. En el 94, con motivo de la publicación del documento *Centros Educativos y Calidad de la Enseñanza* se recuerda que para garantizar el cumplimiento de las leyes y la mejora de la calidad del sistema educativo, la inspección tiene encomendada «la función de participar en la evaluación del sistema educativo y de colaborar en la mejora de la práctica docente y del funcionamiento de los centros», (MEC, 1994: 30). En la Ley Orgánica 9/1995, de 20 de noviembre, de *la participación, la evaluación y el gobierno de los centros docentes* (LOPEG), en su artículo 36 dedicado a las funciones de la inspección educativa, el apartado c recoge la función de «Participar en la evaluación del sistema educativo, especialmente en la que corresponde a los centros escolares, a la función directiva y a la función docente, a través del análisis de la organización, funcionamiento y resultados de los mismos». Cometidos que vuelven a recogerse en los mismos términos, y sin ninguna alteración textual, en la Orden de 29 de febrero de 1996 por la que se regula *la organización y funcionamiento de la Inspección de Educación*, en el capítulo I : *Ámbito territorial, funciones y atribuciones*, apartado Primero dedicado a Fines y funciones.

En la Ley Orgánica 10/2002, de 23 de diciembre, de *Calidad de la Educación* (LOCE) en el Título VI, dedicado a la evaluación del sistema educativo, se refrenda la responsabilidad del Instituto Nacional de Calidad de la Educación (INCE) que, a partir de este momento, pasará a denominarse Instituto Nacional de Evaluación y Calidad del Sistema Educativo (INECSE), para acometer esta tarea, tal y como originariamente le asignó la LOGSE en 1990, sin perjuicio de la evaluación que puedan realizar las Comunidades Autónomas en sus ámbitos respectivos.

A tal efecto, se contemplan las siguientes actuaciones: las evaluaciones generales de diagnóstico sobre áreas y asignaturas en la Educación Primaria y la Educación Secundaria Obligatoria, que versarán sobre competencias básicas del currículo; la participación de las administraciones educativas en las evaluaciones generales de diagnóstico y en las evaluaciones internacionales en las que participe el Estado español; y la elaboración a iniciativa del Ministerio de Educación, a través del INECSE, de un Sistema Estatal de Indicadores de Educación y de un Plan General de Evaluación del Sistema. Asimismo, en otros planes de evaluación, se sigue apostando por la de centros, en sus modalidades externa e interna, y se afirma que la inspección participará en la evaluación externa de los centros docentes sostenidos con fondos públicos.

En la Ley Orgánica de Educación (LOE), de 3 de mayo de 2006 se introducen muy pocas novedades en relación con lo ya establecido en la LOCE sobre la evaluación del sistema educativo y otros programas de evaluación. Cabe señalar, por ejemplo, que vuelve a cambiarse el nombre del Instituto Nacional de Calidad de la Educación (INCE), creado por la LOGSE, que, en la LOCE, pasaba a denominarse Instituto Nacional de Evaluación y Calidad del Sistema Educativo (INECSE) y que, con la LOE, se llamará Instituto de Evaluación (IE).

En la LOE se recoge que el Instituto de Evaluación (IE) elaborará planes plurianuales de la evaluación general del sistema educativo, en colaboración con las administraciones educativas. Y se responsabiliza, asimismo, al IE de la coordinación de las administraciones educativa tanto para la participación del Estado español en las evaluaciones internacionales, como para la elaboración del Sistema Estatal de Indicadores de Educación. Por otro lado, se siguen manteniendo las evaluaciones generales de diagnóstico sobre las competencias básicas del currículo en la enseñanza primaria y secundaria para obtener datos representativos del alumnado y de los centros del Estado y las Comunidades Autónomas.

Por último, en el artículo 145 se contempla la evaluación de centros en sus modalidades externa y de autoevaluación. Y específicamente, en el artículo 151 dedicado a las funciones de la Inspección, se regula que la misma participará en la evaluación del sistema educativo y de los elementos que lo integran.

Esta es, a grandes rasgos, la historia de cómo se va configurando normativamente la evaluación desde los años setenta hasta la actualidad en

España. Los cambios producidos han sido considerables ya que al ámbito de la evaluación de alumnos se va a añadir el de centros en 1970, y a continuación, desde principios de los noventa con la LOGSE se contempla, uno nuevo, el del sistema educativo. En el devenir de la evaluación, a la Inspección se le ha asignado casi siempre la responsabilidad de participar, primero, en la evaluación de los centros y, después, que colaborara también en la evaluación del sistema educativo. En las diferentes etapas, sin embargo, la responsabilidad de la evaluación de los ámbitos de los centros escolares y del sistema educativo no ha sido asumida exclusivamente por la Inspección, puesto que de los primeros, en la Ley General de Educación de 1970, ha de hacerlo en colaboración de los, creados entonces, Institutos de Ciencias de la Educación (ICEs), y, del segundo, ya desde la LOGSE se encarga el Instituto Nacional de Calidad de la Educación.

2. LA EVALUACIÓN EXTERNA DE CENTROS DESDE LAS INICIATIVAS DE LAS ADMINISTRACIONES EDUCATIVAS

La evaluación externa de los centros escolares promovidas por las administraciones educativas es, en estos momentos, una práctica generalizada en la mayor parte de los países europeos y en todos aquellos que han alcanzado un cierto nivel de desarrollo en el contexto internacional. No obstante, es preciso que aclaremos que existen distintas variantes para evaluar los centros educativos. En el panorama de los estudios realizados es posible encontrarse con propuestas que toman a las instituciones como su objeto central y específico de evaluación, presentando en sus informes un análisis y valoración de la enseñanza, aprendizaje y organización de los mismos, así como con otros mucho más limitados, en cuanto a la amplitud del foco con el que captan la imagen de los centros, ya que a través de los resultados de los alumnos en las pruebas nacionales o de las realizadas en las evaluaciones generales de diagnóstico o en evaluaciones internacionales infieren o proyectan una valoración de las instituciones y de sus actores. El primer tipo de estudios son evaluaciones de proceso y producto, y en sentido estricto serían los únicos que podrían considerarse como evaluaciones de centros, mientras que los segundos al limitarse a los productos o rendimiento escolar no pueden catalogarse como tal de centros y sí de evaluaciones del rendimiento de los alumnos.

2.1. La evaluación de productos o rendimientos de los alumnos en el contexto de la evaluación de centros

Las evaluaciones de productos se centran especialmente en los resultados, siendo la relación entre los objetivos y el rendimiento de los alumnos lo que determina el nivel de logro y la eficacia de los centros educativos. Los programas de pruebas de rendimiento y los resultados de exámenes nacionales suelen ser las dos formas más frecuentes que adopta este tipo de evaluaciones. Las puntuaciones obtenidas por los alumnos en las pruebas estandarizadas de rendimiento han sido utilizadas en EE.UU., House (1978), para medir la eficacia de los programas educativos a nivel federal, así como para valorar el logro de los objetivos por parte de los profesores en el ámbito estatal, dentro del movimiento de rendición de cuentas. Las calificaciones escolares en los exámenes nacionales son también contemplados en países como Francia, Inglaterra, Holanda, etc., y, en menor medida, en España como indicadores de la evaluación de los centros educativos.

En Francia, por ejemplo, anualmente se publican los resultados obtenidos por cada liceo en los exámenes del Baccalauréat (BAC). Desde 1995, la Direction de l'Évaluation et de la Prospective (Demailly, Deubel, Gadrey, Verdière, 1998: 83-84) viene utilizando una batería de indicadores para la evaluación cuantitativa y administrativa de los liceos (IPES, *indicateurs de pilotage des établissements secondaires*). Dicha batería incluye cuatro tipos indicadores: resultados de los alumnos (tasas de resultados en el BAC y de acceso a niveles superiores, proporción de bachilleres entre los que salen del liceo, y destino de los alumnos con diploma y sin diploma a su salida del liceo), características de la población de alumnos (sexo, edad media, % de repetidores y repetidores, número de alumnos en Première y en Terminale), recursos y medios (horarios, características de los profesores y número de alumnos por opción), funcionamiento y entorno (servicios asociados a la enseñanza y a la vida escolar, movilidad de los profesores, relaciones con el entorno económico, estatus del liceo, etc.).

Todos estos indicadores, (Thélot, 1995, *Le Monde de l'Éducation*, 1996) son considerados globalmente a la hora de valorar cada liceo. Quizá, la tasa de resultados de cada liceo en el BAC es la más conocida ya que se difunde por los medios de comunicación (*Le Monde de l'Éducation* lanza todos los años un número monográfico). La información que se ofrece permite contextualizar las tasas brutas de cada liceo con las medias de la academia y la nacional, y conocer el resto de los indicadores señalados de cara a que la

información proporcionada se valore en relación con las demás características de cada centro.

En el Reino Unido también, desde 1980, se vienen publicando por el Departamento de Educación y Empleo (DEE) los resultados de los alumnos en los distintos exámenes nacionales, como por ejemplo en el CGSE (Certificado General de Educación Secundaria). A partir de los años noventa, y ante las críticas que suscitaba este tipo de datos brutos al margen de las características de cada escuela, se ha ido imponiendo el sistema de «medidas de valor añadido de la eficacia de la escuela» (MVA). Por valor añadido, (Thomas, 1998: 104) se entiende

«la medida del progreso medio de los alumnos a lo largo de un periodo de tiempo dado (que en las escuelas de enseñanza secundaria se extiende, por regla general, desde el momento de ingreso en el centro docente hasta los exámenes oficiales, y en las escuelas primarias abarca algunos años específicos) en una escuela determinada, en comparación con otras escuelas de la misma muestra».

El procedimiento que se sigue para establecer estas medidas consiste en determinar el nivel de conocimiento inicial y final de los alumnos, a través de una serie de pruebas normalizadas (pretest-postest), y comparar el grado de adquisición de conocimientos adquiridos tanto en el interior de cada escuela como en relación con alumnos de características similares de otras escuelas. Para ello se utilizan técnicas estadísticas de modelización de múltiples niveles con las que, controlando variables continuas y discretas del alumnado (capacidades, edad, sexo, etnia, etc.), se trata de medir los efectos de la escolarización.

Saunders (1999: 22), investigadora de la National Foundation for Educational Research (NFER), institución promotora de este tipo de estudios en el Reino Unido, alerta sobre las precauciones que deben tomarse sobre las medidas de valor añadido:

- «son tan buenas como los datos en que se basan;
- tratan de correlaciones, no de causas;
- están basadas en una normativa y un modelo retrospectivo que no pueden decirnos nada sobre los niveles deseables de actuación futura;
- son sólo un instrumento de evaluación;
- deben, por tanto, ser utilizadas como un recurso de investigación, ya que no proporcionan soluciones rápidas o respuestas correctas a los problemas de la mejora de las escuelas».

En España, la Inspección de Educación, el CIDE y el INCE tienen, asimismo, una cierta tradición en los estudios de evaluación de resultados. Los primeros, es decir, los de la Inspección, son los menos conocidos dado que normalmente no se han difundido públicamente, pero lo cierto es que desde el curso 1986-87 se han venido publicando informes anuales sobre los resultados de las evaluaciones de los alumnos de EGB y enseñanzas medias, así como estudios sobre las calificaciones por áreas y asignaturas (desde el curso 1989-90) y algunas monografías en aspectos como la relación entre el rendimiento y abandono escolar. Sobre este último tema, y como ejemplo de la naturaleza de este tipo de estudios, el SITE (1988) abordaba la relación de la deserción de los estudiantes de enseñanzas medias (FP y BUP-COU) con las siguientes variables: condición de alumno repetidor, evaluación negativa (no promoción) al final del curso académico, entorno del centro, tipo de población atendida por el centro, tamaño del centro, y estudios con los que se accede (sólo en el caso de la FP). El análisis de los datos establece las correlaciones entre las variables consideradas, resaltando la significación de las dos primeras variables (alumno repetidor y evaluación negativa) en el abandono de los alumnos.

Todos estos estudios, por lo general, se limitan a tratar estadísticamente las calificaciones de los alumnos de los distintos niveles del sistema educativo, recogidas de las actas y libros de escolaridad oficiales, relacionándolas con distintas variables como tipo y tamaño de los centros, sexo y edad de los alumnos, etc. La información que proporcionan se centra en aspectos de promoción, fracaso escolar, abandono, etc., pero no ofrecen datos ni conocimientos sobre los procesos y dinámicas de los centros educativos.

Primero el Centro de Investigación y Documentación Educativa (CIDE) y después, a raíz de su creación, el Instituto Nacional de Calidad de la Educación (INCE) han venido realizando desde finales de los años 80 toda una serie de estudios que pueden catalogarse como evaluaciones de producto o resultados. Las evaluaciones externas de la reforma experimental de las enseñanzas medias, —CIDE (1988)—, es el prototipo de estudios que con diseños cuasi experimentales tratan de comparar los resultados de los alumnos, de los grupos experimental y de control, en rendimiento, aptitudes, expectativas, actitudes, reacciones emocionales ante el estudio y variables didácticas. Los datos y resultados, dada la naturaleza de estos estudios, tienen la limitación de centrarse en uno solo de los aspectos de la vida escolar, y aún en el tema de los resultados sus aportaciones son escasamente

relevantes dado que no suponen un avance con respecto a lo ya conocido sobre el rendimiento escolar en las investigaciones realizadas a finales de los años 70 y principios de los 80.

Por su parte, el INCE ha emprendido más recientemente toda una serie de estudios de cara a la evaluación del sistema educativo. Su metodología no difiere sustantivamente de la que ha venido empleando el CIDE, se trata básicamente de diseños cuasi experimentales y *ex post facto* al objeto de comparar los niveles de rendimiento de los alumnos en una secuencia temporal y de establecer las relaciones de asociación o concomitancia con respecto a cierto tipo de variables discretas. Así, por ejemplo en la Evaluación de la Educación Primaria, centrada en los alumnos de 12 años (INCE, 1997, 2000), se ponen a prueba los conocimientos de los alumnos en determinadas áreas/asignaturas (Lengua castellana y Literatura, Matemáticas y Conocimiento del medio), a través de las puntuaciones obtenidas en pruebas estandarizadas, comparando los resultados según el sexo de los alumnos, el tipo de centros, el tamaño de los centros y el nivel cultural de las familias.

En el Diagnóstico de la Enseñanza Secundaria Obligatoria (INCE, 1998), además de los resultados escolares de los alumnos de 14 y 16 años (Comprensión lectora, Gramática y Literatura, Matemáticas, Geografía e Historia, Ciencias de la Naturaleza, Ortografía, Expresión escrita), distribuidos por sexo y Comunidades Autónomas, se ha recogido también información sobre planes de estudios y métodos de enseñanza, funcionamiento de los centros, función docente y contexto socioeducativo: escuela y familia. Pero todos estos aspectos, al menos en el «Avance de Resultados» que nosotros hemos manejado, son procesados y analizados independientemente de los resultados escolares por lo que el enriquecimiento de la información tampoco, en este caso, nos permite avanzar en el conocimiento de la relación entre todos estos factores.

Más recientemente, las propias Comunidades Autónomas han empezado a realizar las evaluaciones generales de diagnóstico sobre las competencias básicas de los alumnos en determinadas áreas del currículo o asignaturas. Quizá, una de las primeras experiencias, con más repercusión pública, fue la evaluación llevada a cabo, en el año 2005, por la Comunidad de Madrid de los alumnos de 6º curso de Educación Primaria. La polémica se suscitó antes de realizarse el estudio, ya que una parte del profesorado, las asociaciones de padres y madres de alumnos y los sindicatos de enseñanza

cuestionaron que, como tal, el estudio pudiera considerarse una evaluación diagnóstica al haberse llevado a cabo con el alumnado del último curso de primaria, y, después de haberlo concluido, al filtrarse a los medios de comunicación los resultados de cada uno de los centros participantes y establecerse una clasificación de los mismos sobre la base de las puntuaciones obtenidas (*El País*, 10 de octubre de 2005).

2.2. La evaluación de procesos y productos o evaluación de centros escolares

La evaluación de centros que combina el análisis y valoración de procesos y productos en la Unión Europea dispone de una buena base documental, ya que desde el año 2000 la red Eurydice inicio una serie de estudios monográficos sobre las características de este ámbito de evaluación en muchos de sus países (Eurydice, 2002) y posteriormente (Eurydice, 2004), ha publicado el informe de un estudio comparado, en el que se contemplan globalmente todos los países que formaban parte de pleno derecho de la UE en esas fechas, es decir, los 15, más los tres países de la Asociación Europea de libre comercio que son miembros del Espacio Económico Europeo, —Islandia, Liechtenstein y Noruega—, y los países candidatos en el momento de la redacción del documento —otros 12 más—. En todos los casos, tanto en los estudios individuales por países como en el global, en el que se comparan entre sí al conjunto de los países, se aborda la evaluación de centros de enseñanza obligatoria.

Una vez hecha esta aclaración, trataremos de presentar los aspectos sustantivos que definen la evaluación de centros europeos. Con el fin de ofrecer una visión lo más completa posible, nos detendremos en los elementos que concretan el objeto de evaluación, o lo que es lo mismo qué es lo que se evalúa, en la definición de los criterios y los procedimientos seguidos para la evaluación, es decir, cómo y con qué fines se realiza la evaluación, y si se utilizan y para qué los resultados de las evaluaciones. Conviene tener en cuenta, no obstante, que la casuística de los 30 países participantes hace muy complicado dar una visión general, así como precisar las particularidades que se dan en muchos de ellos, por lo que para tener un conocimiento más completo de ciertos países es necesario consultar las monografías de los mismos (disponibles en http://www.eurydice.org/Documents/Evaluation/en/frameset_eval.html).

A) Objeto de evaluación

El objeto de evaluación son los centros como entidades, ocupándose de las actividades de su personal en conjunto sin atribuir la responsabilidad de dichas actividades a uno o más de sus miembros concretos, excepto en el caso del director que en algunos casos también son evaluados individualmente como parte de la evaluación del centro. En ciertos países, también, la evaluación de un centro considera simultáneamente la valoración individual del trabajo de sus profesores junto con la evaluación de su funcionamiento como entidad. No obstante, estos dos últimos casos no suelen ser frecuentes entre los países europeos.

A los efectos de evitar equívocos, hay que tener presente que los estudios de Eurydice (2002 y 2004) se han ocupado exclusivamente de los centros que reúnen las siguientes características: de enseñanza obligatoria del sector público, es decir, los financiados y gestionados por los poderes públicos; de enseñanza obligatoria a tiempo completo; y de los que imparten educación general por lo que no se consideran los de formación profesional ni, tampoco, los de educación especial, por no ser en el caso de estos últimos de educación general ordinaria. Por último, hay que tener en cuenta que las fechas de referencias de los estudios (monografía y análisis comparado) son los años 2000/01.

Las actividades del centro que pueden ser objeto de evaluación se clasifican en seis grandes tipos de funciones, diferenciándose a la vez las que pueden caracterizarse como educativas de aquellas otras que son fundamentalmente administrativas.

Cuadro 4. 1. Funciones educativas y administrativas que son objeto de evaluación.

Funciones educativas
<ul style="list-style-type: none"> • (a) la enseñanza/aprendizaje de aptitudes y conocimientos, y • (b) la enseñanza/aprendizaje de los modelos apropiados de conducta social y desarrollo personal, incluida la orientación profesional.
Funciones administrativas
<ul style="list-style-type: none"> • (c) la gestión de los recursos humanos; • (d) la gestión de los recursos operativos; • (e) la gestión de los recursos de capital, y • (f) las actividades relacionadas con la información, las relaciones externas y la colaboración con otras asociaciones.

Fuente: Eurydice (2004:11).

B) Criterios de evaluación

En relación con los criterios de evaluación utilizados, el panorama es más diverso del que presentan las funciones que delimitan y concretan el objeto de evaluación. En este sentido, podemos encontrarnos con las siguientes variantes: países que tienen listas predeterminadas de criterios que, normalmente, son establecidos por las autoridades de las administraciones central, regional o local; países en los que son los evaluadores quienes establecen los criterios basándose en la legislación educativa nacional, independientemente de su ámbito territorial de actuación —central, regional o local—, aunque en algunos de ellos se tienen en cuenta las categorías de evaluación formuladas en el ámbito central, los objetivos de la política educativa local o regional, o los objetivos concretos de los centros; y países en donde los evaluadores tienen autonomía total para seleccionar los elementos sobre los que establecer sus criterios (Eurydice, 2004: 56-60).

Sólo diez países de los que utilizan listas predeterminadas de criterios, disponen de una información similar y comparable en este aspecto concreto, entre ellos se encuentran España, Países Bajos, Portugal, el Reino Unido, etc. En el Cuadro 4.2, se presentan las 14 categorías en que se estructuran los criterios de proceso.

Cuadro 4. 2. Criterios de evaluación de procesos de los países que disponen de listas preestablecidas.

Categorías que estructuran los criterios de proceso
1. Enseñanza/aprendizaje en el aula
2. Orientación y apoyo a los alumnos
3. Funcionamiento de los órganos/organización del centro
4. Política educativa/general del centro
5. Relaciones entre el centro y la comunidad local/relaciones externas
6. Gestión de los recursos humanos
7. Gestión del tiempo escolar
8. Actividades extraescolares
9. Evaluación interna
10. Liderazgo
11. El clima del centro
12. Gestión de las instalaciones
13. Gestión de los recursos financieros y materiales
14. Procedimientos administrativos

Fuente: Elaboración propia a partir de los datos de Eurydice (2004:66-68).

Cada una de estas categorías se puede dividir en varios parámetros. En el Cuadro 4.3 puede verse, a modo de ejemplo, los parámetros incluidos en la categoría de orientación y apoyo a los alumnos. En el Anexo de este tema recogemos el listado completo de categorías y parámetros de los criterios de evaluación de proceso.

Cuadro 4.3: Parámetros que integran la categoría de Orientación y apoyo a los alumnos.

Orientación y apoyo a los alumnos	
Orientación escolar	Actividades de recuperación
Actividad tutorial	Desarrollo personal de los alumnos
Desarrollo de las habilidades sociales de los alumnos	Medidas para mejorar el comportamiento de los alumnos
Tratamiento de los alumnos con necesidades especiales	Lucha contra el absentismo

Fuente: Elaboración propia a partir de los datos de Eurydice (2004:67).

En el Cuadro 4.4, se presentan los parámetros que se utilizan en los criterios de evaluación de producto o resultados, en la evaluación externa de los centros europeos que disponen de listas preestablecidas por los responsables de las administraciones.

Cuadro 4.4. Parámetros de los criterios de evaluación de producto o resultados.

Datos cuantitativos
Índice de éxito/fracaso en la promoción de un año al siguiente o de una etapa a otra, porcentaje de alumnos admitidos en Educación Especial, tasa de abandono
Resultados en tests o en exámenes (que pueden tener relación con habilidades sociales o capacidades cognitivas)
Absentismo escolar
Trayectoria tras la enseñanza obligatoria
Datos cualitativos
Habilidades sociales de los alumnos
Capacidades cognitivas de los alumnos

Fuente: Elaboración propia a partir de los datos de Eurydice (2004:69).

A tenor de lo expuesto hasta aquí en torno a los criterios de evaluación, convendría que hiciéramos algunas precisiones sobre el uso de la terminología empleada ya que el empleo simultáneo de los vocablos criterio, categoría y parámetro puede inducir a ciertas confusiones o errores. Tal y como planteábamos en el Tema 3, los **criterios** son los referentes, normas intelectuales o características deseables con los que se compara la información recogida de cualquiera de los objetos o ámbitos de evaluación. La utilización de criterios de evaluación es consecuente con la necesidad de enjuiciar, de emitir en suma un juicio de valor, ya que la evaluación como proceso o actividad comparativa requiere comparar las características que un objeto cualquiera (instituciones, docentes, currículo) tiene en un momento dado, en diferentes fases temporales o en relación con otros objetos de la misma naturaleza con las que son deseables que posea.

Estos son el significado y uso que se da al término criterio, en el campo de la evaluación, y que coinciden con una de las acepciones que se contemplan en el *Diccionario del Uso del Español* (Manuel Seco *et al.*): 'Norma intelectual para juzgar o decidir'. En esta definición, sin embargo, hay un matiz que no deberíamos pasar por alto, el hecho de que además de para 'juzgar' puede utilizarse para 'decidir'. Y es, en este sentido, como se emplea también en este estudio europeo cuando se asocia a los términos categorías y parámetros para ir determinando y concretando los distintos elementos del objeto de evaluación de los que se recogerá información, utilizándose estos dos últimos como sinónimos de dimensiones y subdimensiones de evaluación. De hecho, en la terminología general sobre los modelos de evaluación de los centros (disponible en http://www.eurydice.org/Documents/Evaluation/terminology_fr.html), se precisa que «un criterio se compone de dos elementos: los parámetros (u objeto de evaluación) y el nivel de exigencia (norma, regla o nivel de competencia)...»

Los criterios de evaluación para la formulación de un juicio de valor de cada una de las categorías (dimensiones) y parámetros (subdimensiones), se definen a través de escalas de puntuaciones con las que se comparan los rendimientos y/o resultados obtenidos. Las variantes que pueden encontrarse en los 10 países con listas preestablecidas (Eurydice, 2004: 71), se centran en los cuatro siguientes grupos:

- a. Para cada parámetro analizado se propone una escala de puntuación especial con descripciones precisas de los diferentes niveles de rendimiento posibles.
- b. Se emplea la misma escala de evaluación, con distintos grados, para todos los parámetros evaluados.
- c. Únicamente se utiliza una escala de puntuación para reflejar la actividad global de los centros.
- d. La autoridad central o superior no establece ninguna escala para la evaluación de los centros.

C) Procedimientos y utilización de los resultados de la evaluación

En general, la secuencia a través de la que se afronta la evaluación de los centros, en diferentes países, pasa por una serie de momentos o fases, a saber:

- Un momento inicial o fase preliminar en la que los evaluadores, con anterioridad a su entrada a los centros educativos, disponen de y consultan documentación diversa.
- Una fase en la que, una vez recogida y procesada la información, se vuelve nuevamente a contactar con los responsables de los centros antes de la redacción definitiva del informe de evaluación.
- La fase de seguimiento de los centros para supervisar si los mismos han asumido y puesto en práctica las recomendaciones e instrucciones que se les daba en el informe de evaluación.

En la gran mayoría de los países y en los más grandes e importantes, a excepción de Francia, en la evaluación de los «Collèges», y en Alemania, los evaluadores antes de empezar el estudio propiamente dicho consultan una documentación que, aunque es diversa en cada país, está relacionada con: Informes administrativos o descriptivos (autoridad superior u órgano nacional), resultados basados en pruebas o en el rendimiento (autoridad superior u órgano nacional), evaluaciones anteriores (evaluador), cuestionario para los padres, personal, etc. (evaluador), documentos elaborados/proporcionados por el centro.

En todos los países participantes, los documentos elaborados por el centro educativo que se ponen a disposición de los evaluadores se clasifican (Eurydice, 2004: 86), en cuatro posibles tipos, siendo los dos primeros los más frecuentemente utilizados:

1. Planes de políticas educativas y organizativas del centro (proyecto educativo, proyecto curricular, proyecto para el desarrollo de las tecnologías de la información y la comunicación, etc).
2. Datos objetivos relativos al centro escolar (resultados de las pruebas o exámenes internos, datos sobre las expulsiones de alumnos o sobre la composición de las clases, informe sobre la gestión financiera, etc.).
3. Documentos producidos por el centro para uso de personas externas a éste (folletos, prospectos, etc.).
4. Evaluaciones o revisiones de cuentas llevadas a cabo por el centro.

Una vez que se ha recogido la información prevista en el plan de evaluación y se ha procesado convenientemente la misma, los evaluadores antes de elaborar su informe definitivo dan a conocer a los responsables de los centros un informe provisional de evaluación. Normalmente es al director, equipo directivo y consejo escolar, dependiendo de los países, a los que se les proporciona información, en ocasiones de forma oral, de los resultados de la evaluación. En algunos países —Países Bajos, Portugal, Inglaterra, República Checa y Rumania—, los responsables de los centros disponen de un plazo de tiempo prefijado para enviar sus alegaciones. En cualquier caso, parece claro que la devolución de la información a los centros y la posibilidad de que estos puedan analizarla, haciendo comentarios y observaciones, ponen de manifiesto que la evaluación trasciende los propósitos de los procesos habituales de supervisión y refuerzan, por el contrario, las actuaciones y propuestas de mejora de la calidad de la educación

La fase de seguimiento de los centros para supervisar la puesta en práctica de las medidas y/o recomendaciones que se les hizo en el informe definitivo de evaluación, no es algo que se haya generalizado en todos los países. La casuística va desde aquellos en los que el seguimiento se hace de forma sistemática y habitual después de cada evaluación —España (Andalucía y Canarias), Reino Unido (Escocia), etc.—, pasando por otros en los que solamente se lleva a cabo en ciertos casos, como cuando el centro

presenta deficiencias reseñables, —Reino Unido (Inglaterra, Gales e Irlanda del Norte), Países Bajos, Irlanda, etc.—, hasta llegar a países en que no se hace seguimiento de la evaluación, -Alemania, Francia, España (País Vasco).

En relación con las consecuencias de la evaluación externa de los centros, el estudio comparado de Eurydice (2004: 90) señala las tres siguientes:

1. La evaluación da lugar a recomendaciones, o incluso a instrucciones por parte de los evaluadores o autoridades educativas. Estas recomendaciones/instrucciones se consideran el punto de partida para iniciar el proceso de mejora de la calidad del centro.
2. La evaluación insta a los centros escolares a elaborar un plan de mejora en el que se estructuren y organicen los objetivos que deban alcanzarse.
3. La evaluación da lugar a sanciones que son establecidas por los propios evaluadores o por las autoridades educativas. La sanción puede afectar al centro escolar o a sus responsables. Puede tratarse de una disminución de los recursos concedidos, de la retirada de algunas prerrogativas (como la expedición de certificados) del centro, de multas dirigidas a algunas personas o incluso de su despido.

Por último, se ha de señalar que la publicación de los resultados e informes de la evaluación externa es una práctica habitual en algunos países como Reino Unido (Gales, Inglaterra, Irlanda del Norte), Países Bajos, Suecia, Portugal (página web de la Inspección), etc., difundiéndose dichos resultados a través de los medios de comunicación por lo que no es infrecuente que se establezcan clasificaciones de los mismos. En cambio, en otros como Alemania, España (Andalucía y País Vasco), Francia, Irlanda, Polonia, etc., no se publican. Ambas posturas tienen sus pros y contras. Los partidarios de hacer públicos los resultados de la evaluación afirman que favorecen la elección consciente de los ciudadanos de los centros escolares, en tanto que los detractores se niegan a que la educación sea una mercancía más sometida a las reglas del mercado. La polémica tiene difícil solución, dados los valores y presupuestos que subyacen en cada una de las posiciones, aunque lo que es difícil de entender son los impedimentos a los que han de hacer frente docentes, instituciones e investigadores para conocer qué se está haciendo y cuál es el balance de las experiencias de evaluación institucional.

2.3. Propuestas y realizaciones españolas

A comienzos de los años 90, la administración educativa elabora un plan de evaluación de centros, en un trabajo de colaboración entre el CIDE y la Inspección Central de Educación, denominado posteriormente Plan de Evaluación de Centros (Plan EVA), que se desarrollará principalmente, (Cerdán, Sahuquillo, Luján y Puente, 1992; Menéndez i Pablo, 1993), como un propuesta de evaluación formativa externa para centros docentes no universitarios y en el que inicialmente se resalta su carácter cualitativo así como su propósito de fomentar la evaluación interna de las instituciones².

El Plan Eva (MEC, 1992: 25-26) articula su propuesta de evaluación en torno a los aspectos materiales, personales y formales del centro escolar, así como en los resultados obtenidos por el mismo. Establece siete dimensiones con sus consiguientes subdimensiones de evaluación:

1. **Contexto:** tipo de centro, hábitat, tamaño y características del centro
2. **Recursos humanos y materiales.** Personal docente: características personales, experiencia docente, formación permanente, grado de satisfacción con su trabajo en el centro. Personal no docente: características personales, mobiliario, material didáctico y psicopedagógico
3. **Apoyos externos:** servicio de inspección, equipos psicopedagógicos, centros de profesores, centros de recursos, otros.
4. **Alumnado:** características personales, características sociofamiliares, antecedentes escolares, aspiraciones y expectativas, autoconcepto académico, motivación
5. **Organización y funcionamiento.** Programación general anual, gestión económica, gestión administrativa, elección y constitución del equipo directivo y del Consejo Escolar. Organización pedagógica: equipos docentes, agrupación de alumnos, reglamento de régimen interior, distribución de horarios y espacios, funcionamiento del equipo directivo, funcionamiento conjunto del equipo directivo, estilo directivo, director, jefes de estudios, secretario, funcionamiento de

² Decimos inicialmente porque, más tarde, en Luján, J. y Puente, J. (1996: 114) se dice textualmente que el “enfoque metodológico del Plan EVA se identifica con una postura ecléctica”. Y tampoco sabemos, por otra parte, que durante su vigencia generará procesos de atoevaluación.

los órganos colegiados (Consejo Escolar, Claustro), funcionamiento de los equipos docentes. Participación de la comunidad educativa: padres y madres, alumnos, relaciones con la comunidad.

6. **Procesos didácticos:** organización del aula, metodología, relación didáctica, evaluación, actividades de orientación y tutoría, actividades complementarias.
7. **Rendimiento educativo:** resultados académicos, actitudes, tasas e índices.

El Plan, como tal, se estructuraba en cinco módulos: elementos contextuales y personales, proyectos, organización y funcionamiento, procesos didácticos y resultados. Cada módulo estaba diseñado de manera autosuficiente para su aplicación, permitiendo una evaluación tanto sectorial como global en función de que dichos módulos se apliquen por separado (uno o más módulos) o conjuntamente (todos los módulos). Se precisa, no obstante, que todos los módulos están interrelacionados y «constituyen una unidad estructural».

El proceso de evaluación se realizaba a través de una serie de fases: visita inicial (reuniones con el Consejo Escolar y el Claustro para informar sobre el Plan), visita de evaluación (recogida de información), análisis de la información y elaboración de conclusiones (cumplimentación de las fichas de evaluación y estudio y análisis de las mismas), visita final (contrastar las conclusiones con la comunidad educativa) e informe final (redacción del informe conforme a un esquema propuesto).

En sus cinco años de vigencia, desde el curso 1991-92 (proyecto piloto), al 1995-96, el número total de centros que vivieron la experiencia del Plan EVA fue de 1076 (cifra que hemos obtenido de los datos proporcionados para cada curso por Luján y Castro, 1996: 332-337).

Con la llegada al gobierno del Partido Popular en 1996, la administración educativa va a emprender una nueva etapa en el desarrollo de la evaluación de centros. Por un lado, a partir de ese momento, el Plan EVA va a ser sustituido por el Modelo de Gestión de Calidad como instrumento de autoevaluación (Resolución de la Dirección General de Centros de 27 de mayo de 1998), iniciándose la experimentación del mismo durante el curso 1997-98 en 6 colegios de primaria y 10 institutos de educación secundaria.

Y, por otro, se crea el Plan Anual de Mejora de centros educativos que con carácter experimental se ensaya en los cursos 1996-97 y 1997-98 y se generaliza y establece su aplicación permanente desde el curso 1998-99 (Orden de 9 de junio de 1998).

Más adelante, en el apartado dedicado a la evaluación interna, analizaremos las implicaciones y presupuestos del Modelo de Gestión de la Calidad, así como las características del Plan Anual de Mejora en tanto que instrumento que trasluce los valores y orientaciones de las políticas educativas de la actual administración.

La mayor parte de las propuestas de evaluación promovidas por el Ministerio de Educación perdieron su ímpetu inicial o quedaron paralizadas desde el mismo momento que finalizó el proceso de transferencias en materia de educación no universitaria del Estado a las Comunidades Autónomas, llevado a cabo entre 1998 y 2000 en aquellas Comunidades que no habían asumido hasta entonces estas competencias, y su continuidad dependería, a partir de entonces, de que éstas adoptaran y siguieran trabajando con dichas propuestas y modelos.

La situación creada a partir de entonces se pone de manifiesto en la monografía que el CIDE (2002) elaboró sobre la evaluación de centros en España para Eurydice y que, a diferencia de otros países, lleva el título de *Aproximaciones a la evaluación de centros. España (2000/01)*. El uso del término aproximaciones se entiende perfectamente ya que en el momento de redactar el informe se dice:

«[en] España coexisten dos modalidades de evaluación de centros docentes: la evaluación externa y la evaluación interna. Tanto la una como la otra, debe desarrollarse en todos los centros del territorio español, aunque hasta el momento sólo algunas Comunidades Autónomas han elaborado planes sistemáticos de evaluación. Concretamente, Andalucía, Canarias y Cataluña han elaborado y puesto en marcha sus propio plan de evaluación de centros, encontrándose el resto de las Comunidades Autónomas en la fase de elaboración» (CIDE, 2002: 4).

Cuadro 4.5. Tareas y evaluadores de los distintos tipos de evaluación en España.

	Tareas escolares evaluadas	Evaluadores		
		Externa	Interna	Mixta
Tareas educativas	1. Enseñanza-Aprendizaje de capacidades y conocimientos	Inspección de Educación	Consejo Escolar Claustro Profesores	
	2. Enseñanza -Aprendizaje sobre vida adulta/ orientación profesional	Inspección de Educación	Consejo Escolar Claustro Profesores	
	3. Gestión del personal	Inspección de Educación	Consejo Escolar Claustro Profesores	
	4. Gestión de recursos operativos	Inspección de Educación	Consejo Escolar Claustro Profesores	
Tareas administrativas	5. Gestión de recursos de capital	Inspección de Educación	Consejo Escolar Claustro Profesores	
	6. Información/recursos informativos/relaciones/ asociaciones	Inspección de Educación	Consejo Escolar Claustro Profesores	

Fuente: CIDE (2002:5)

En el Cuadro 4.5 se presentan los distintos elementos del objeto de evaluación, así como los responsables de llevarlas a cabo en las diferentes modalidades de evaluación de centros en España, de acuerdo con el esquema general utilizado por Eurydice (2002 y 2004) para realizar los estudios monográficos de cada país y el estudio comparado entre todos los países.

En el estudio español, CIDE (2002), se contempla el análisis de las propuestas de evaluación externa de las tres Comunidades Autónomas citadas (Andalucía, Canarias y Cataluña) y el programa de evaluación de la Educación Secundaria Obligatoria del País Vasco. En la modalidad de eva-

luación interna, además de Andalucía, Canarias (en donde se distinguen dos variedades: autoevaluación y evaluación interna) y Cataluña, se incorpora la de Castilla y León. No se recoge, sin embargo, por la fecha en que se concluyó el estudio, el Plan de evaluación de centros docentes de la Comunidad de Castilla-La Mancha, que se aprobó por la Orden de 6 de marzo de 2003, (DOCM 17-03-03), y la Resolución de 30 de mayo de 2003, (DOCM. 11/06/03), y en la que se desarrollan las modalidades de evaluación externa e interna.

En general, en lo que respecta a la evaluación externa de centros docentes, en todas las Comunidades Autónomas existen actuaciones aunque no hayan desarrollado planes sistemáticos. En estos casos, la evaluación se lleva a cabo a través de las actuaciones de la Inspección con la recogida de datos e informaciones en sus visitas a los centros, aun cuando el propósito de dichas visitas sea la supervisión de la organización y funcionamiento de los centros y no específicamente su evaluación. Tanto en los planes ya existentes como en los que se han ido desarrollando con posterioridad, las orientaciones de las propuestas de evaluación son también diferentes. Así, por ejemplo, los de la Comunidad Valenciana y Castilla y León han adoptado el Modelo Europeo de Gestión de Calidad (EFQM) y los Planes Anuales de Mejora (PAM), promovidos por el Ministerio de Educación durante los gobiernos del partido Popular, mientras que en Andalucía, Canarias, Cataluña, Castilla la Mancha, Navarra, País Vasco, La Rioja, etc., la situación es muy diversa. Navarra y La Rioja, según el estudio del CIDE (2002: 33), están aprovechando las evaluaciones generales del sistema educativo para iniciar una aproximación a la evaluación de centros, mientras que el País Vasco también está desarrollando un programa de evaluación de centros.

Las Comunidades con una mayor continuidad en sus planes son Andalucía, Canarias y Cataluña. Las tres abordan la evaluación externa de los centros con planteamientos diversos y singulares. Andalucía, por ejemplo, que adoptó desde sus comienzos el Plan de Evaluación de Centros (Plan EVA), promovido por el Ministerio de Educación del gobierno socialista, ha ido transformando el mismo en la dirección de disponer de indicadores de resultados. Canarias, también ha ido cambiando su plan inicial, aunque lo sigue estructurando en torno a las dimensiones de contexto, procesos y resultados en los ámbitos de enseñanza, aprendizaje y organización de los centros. Cataluña, por su parte, desde el año 1997, viene realizando su evaluación en los ámbitos de enseñanza-aprendizaje y organización. Para un conocimiento pormenorizado de todos

estos planes, consúltese el estudio completo del CIDE (1992), en la siguiente dirección URL: http://www.eurydice.org/Documents/Evaluation/en/frameset_eval.html.

3. ESTRATEGIAS DE EVALUACIÓN INTERNA Y COMBINADA DE EVALUACIONES INTERNA Y EXTERNA

La evaluación interna o la combinación de las modalidades externa e interna son actuaciones defendidas tanto por las administraciones educativas como por algunos autores que vienen trabajando en el campo de la evaluación. Claude Thélot, director de la Direction de l' Evaluation et de la Prospective (DEP) del Ministerio de Educación francés, planteaba en una entrevista de *Le Monde de l' Éducation* en 1996 que no se podía defender a la vez la idea de proyecto institucional y la imposición de una evaluación exclusivamente estándar, señalando que (Guibert et Reverchon, 1996: 37) «la evaluación de establecimientos escolares y universitarios debe ser en parte estándar; es decir, común a todos los establecimientos del mismo tipo, y en parte específica para cada establecimiento».

Para Thélot (1995: 87), una evaluación puramente interna y no pública, incluso si toma la forma de autoevaluación, corre el riesgo de perder pertinencia y utilidad, tanto para facilitar información al exterior como para alimentar la propia reflexión interna:

«[mientras que] una evaluación puramente externa, destinada a crear una competencia excesiva entre los centros y desconectada de los intereses del director y del conjunto de los actores, conducirá a establecer clasificaciones de centros, corriendo el riesgo de que no pueda llevarse a cabo por la falta de información fiable sobre cada centro».

En Inglaterra, el Informe Blanco *Excellence in Schools* (GB. HoC. Parliament, 1997) reconocía que para abordar la mejora del rendimiento escolar era necesario la implicación de los centros en procesos de autoevaluación. Por ello, el Ministerio de Educación y Trabajo británico propone un ciclo de cinco fases para acometer la autoevaluación y la mejora anual que incluye:

- La revisión de la calidad del rendimiento, la pedagogía y la gestión.
- La comparación diagnóstica con respecto a otras escuelas similares.
- La fijación de objetivos específicos para cada escuela y centrados en conseguir un aumento del rendimiento de los alumnos.

- La planificación de desarrollo para poner en práctica los objetivos.
- La evaluación de la acción llevada a cabo y de su impacto en los alumnos.

La OFSTED (1998) — <http://www.ofsted.gov.uk/>, defiende que la autoevaluación complementa la evaluación externa realizada por la inspección y, en este sentido, han propuesto la necesidad de que las evaluaciones de las escuelas, de cara a incrementar su eficacia y complementariedad, utilicen los mismos criterios e indicadores y hasta técnicas similares a las empleadas por la inspección.

Por otra parte, como ya hemos tenido oportunidad de ver la mayoría de las iniciativas de evaluación externas son partidarias, asimismo, de impulsar propuestas de evaluación interna en la medida que se es consciente de la necesidad de implicación de los centros y la comunidad educativa para llevar adelante los planes de mejora. No obstante, uno de los problemas que se plantea en estos casos se refiere concretamente a si la autoevaluación ha de ser o no una actuación complementaria a las evaluaciones externas llevadas a cabo por la Inspección.

Lo más frecuente es que la evaluación interna se subordine a las exigencias, requisitos, objetos e indicadores de la evaluación externa. Muchas de las propuestas de autoevaluación en la actualidad pervierten su sentido y significado educativos, produciéndose una instrumentalización de la misma ya que es utilizada para llegar y obtener de los centros información que, de otro modo, sería imposible ante la evidencia de la falta de recursos y personal para generalizar administrativamente proyectos de evaluación externa (recuérdese, por ejemplo, que durante los cinco años de vigencia del Plan Eva sólo se evaluaron 1076 centros, número realmente insignificante con relación al total de los existentes).

En España, como ya se ha señalado con anterioridad, la administración popular va a emprender una nueva etapa en el desarrollo de la evaluación de centros con el desarrollo de propuestas como el Modelo de Gestión de Calidad y los Planes Anuales de Mejora.

Más adelante, analizaremos las implicaciones y presupuestos del Modelo de Gestión de la Calidad. Ahora, señalaremos algunas de las características del Plan Anual de Mejora en tanto que instrumento que trasluce los valores y orientaciones de las políticas educativas de la administración de esos años. De una manera muy sintética, cabría resaltar el papel relevante que se le asigna en todos estos proyectos al «liderazgo efectivo» de los equipos directivos, la adver-

tencia de que el diagnóstico explícito de la situación de partida se base en hechos o resultados y no en opiniones subjetivas, la conjugación de la evaluación interna (realizada por el equipo directivo y el equipo de mejora) y la evaluación externa (competencia de la Inspección), la asignación de incentivos económicos (de hasta un millón de las antiguas pesetas) a los centros que se hayan distinguido por la calidad de su plan, la difusión de los mejores planes para extender las buenas prácticas, etc.

Cuando se leen detenidamente los documentos sobre los planes de mejora siempre surge la duda de si realmente estamos ante propuestas que intentan aunar voluntades en el trabajo cotidiano de los docentes, o realmente se trata de instrumentos que van a ir preparando a los centros para la implantación de nuevas modalidades de evaluación, dada la importancia que se da a la misma en dichos planes. Ya es curioso, por ejemplo, que en la introducción de la publicación *Prácticas de buena gestión en centros educativos públicos* (en la que se difunden los mejores planes anuales de mejora) se hagan afirmaciones como: «El <Plan Anual de Mejora> se integra dentro del <Plan para una Gestión de Calidad en Educación> (cuando el Modelo de Gestión de Calidad se define como un instrumento de autoevaluación); y de un modo más claro cuando se habla, en la misma introducción,

«[los] conceptos de evaluación y mejora, en tanto que elementos fundamentales de esta nueva orientación, están conectados entre sí de un modo circular: la mejora conduce a la evaluación que conduce a la mejora [...]. La ordenación del <Plan Anual de Mejora> ha hecho de la evaluación de las actuaciones previstas en el Plan y de sus resultados un elemento fundamental, pero ha procurado deliberadamente, poner el acento en la mejora que es lo que otorga a la evaluación su dimensión más profunda, su sentido más elevado».

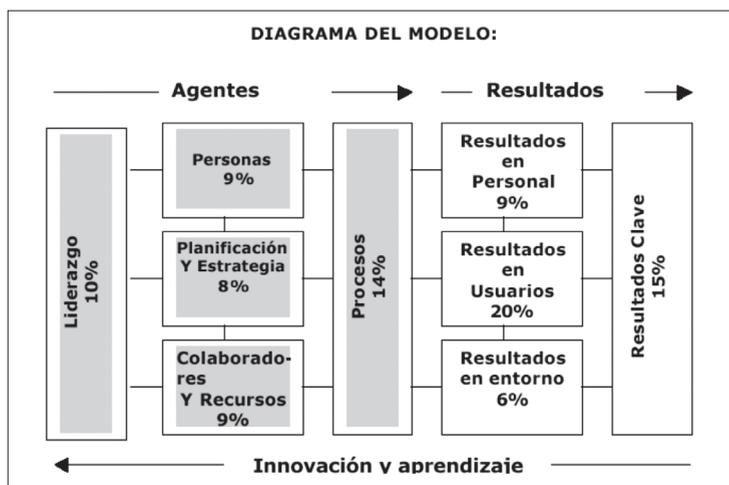
La propuesta auspiciada por el Ministerio de Educación y Cultura (1997) adopta la definición de autoevaluación, dada por la Fundación Europea de Gestión de la Calidad, en el sentido de que consiste en «un examen global, sistemático y regular de las actividades y resultados de una organización comparados con un modelo de excelencia». Dicho modelo no es otro que el que se conoce como Modelo Europeo de Gestión de la Calidad. Si se leen con detalle los documentos elaborados para tal fin, parece claro que el «modelo» y la autoevaluación son los dos instrumentos con los que se pretende introducir una nueva forma de gestión en los servicios públicos de la educación. Siendo la autoevaluación el «método», que permite descubrir las fortalezas y debilidades de cada centro con respecto a un concepto de gestión

«[que] se aleja de [una] visión restringida, propia de la orientación burocrático-administrativa, para adquirir un significado más global, comprensivo de la gestión de recursos, de las personas, de los procesos y de los resultados, elementos que, en interacción recíproca, se dan cita en toda organización» (palabras de López Rupérez en la *Introducción*)³.

El proceso mismo de autoevaluación se determina y concreta en relación con la esencia misma del modelo de calidad que, según se recoge en estos documentos, se resume en

«[que la] satisfacción de los usuarios del servicio público de la educación, de los profesores y del personal no docente, y el impacto en la sociedad se consigue mediante un liderazgo que impulse la planificación y la estrategia del centro educativo, la gestión de su personal, de sus recursos y sus procesos hacia la consecución de la mejora permanente de los resultados».

El modelo se articula en torno a nueve criterios, que agrupados en dos grandes categorías, —los criterios *agentes* (liderazgo, gestión del personal, planificación y estrategia, recursos, procesos) y los criterios resultados (satisfacción del personal, satisfacción del cliente, impacto en la sociedad, resultados del centro educativo)—, van a ser los elementos sustantivos en que se centre el proceso de evaluación.



³ En la misma Introducción, se habla de que los colegios e institutos «son organizaciones de vocación humanista que trabajan con personas, entre personas y para personas, de modo que un modelo orientado hacia las personas podría ser asumido por esta modalidad de organizaciones, por lo general, reacias a aceptar referentes vinculados al mundo empresarial», para un poco más adelante señalar que los ganadores del Premio Europeo a la Calidad han sido organizaciones como: Ran Xerox, Texas Instruments, Ericsson S.A., etc.

La evaluación puede llevarse a cabo a través de «dos enfoques sencillos», por medio del «enfoque del formulario» y del «enfoque del cuestionario» y en ambos casos «se realiza por comparación con los nueve criterios que componen el Modelo». El formulario consiste en una matriz o tabla sobre cada uno de los criterios y subcriterios en que se subdividen los primeros, que deben ser puntuados, en una escala de porcentajes con cinco intervalos (0%, 25%, 50%, 75%, 100%), en los apartados de enfoque (manera de afrontar o solucionar un problema o alcanzar un objetivo) y despliegue (extensión en que se aplica el enfoque relacionándola con el potencial total). Mediante el cuestionario, en cambio, se puntúa en una escala del 1 al 4 las 50 preguntas elaboradas a partir de los nueve criterios. Las puntuaciones obtenidas, en uno y otro caso, en cada uno de los criterios en relación con las asignadas inicialmente en el modelo permiten detectar los puentes fuertes y las áreas de mejora de cada centro.

Las experiencias de autoevaluación desarrolladas por Canarias y Cataluña son sustancialmente distintas a la auspiciada por el MEC. En principio, por ejemplo, en la convocatoria sobre Proyectos de Evaluación Interna y Desarrollo de Centros, de la Consejería de Educación del Gobierno canario (BOC de 27-XI-96), no se parte de ningún modelo predeterminado, dado que se deja a la iniciativa de los centros que fijen la concreción del objeto de evaluación, elijan las técnicas e instrumentos que vayan a utilizar y decidan los indicadores de evaluación. Para ello, AA.VV. (1999: 58), el Instituto Canario de Evaluación y Calidad Educativa (ICEC) «se comprometía a hacer un seguimiento del proceso, a prestar los apoyos técnicos necesarios, y a facilitar la formación requerida para llevarlo adelante». Por su parte, la propuesta de autoevaluación de Cataluña, según Alcalde y Rué (1999: 50), en cuanto a sus procedimientos y metodologías desarrollados se basa en las «aportaciones realizadas desde el campo de la investigación-acción, así como de aquellas registradas en el ámbito de la evaluación democrática».

Desde estos presupuestos de partida, las experiencias canaria y catalana no pretenden comprobar si los centros se ajustan a un determinado modelo de gestión, sino que los centros inicien la autoevaluación desde los problemas o preguntas que son relevantes para conocer y comprender la dinámica institucional al objeto de fundamentar sus cursos de acción y plantearse propuestas de mejora. Así, por ejemplo, en la evaluación interna de un instituto de educación secundaria canario se destaca que, a pesar de las dificultades e inconvenientes surgidos,

«se valora positivamente la información que en estos momentos se posee acerca del funcionamiento del centro, ya que la variedad de aspectos evaluados permitirá reutilizarlos en momentos posteriores con objetividad, como punto de partida para otro tipo de proyectos, por ejemplo, el Proyecto Educativo de Centro» (AA.VV., 1999: 62).

Opinión que también es compartida por los participantes en la autoevaluación de una escuela pública catalana, cuando se resalta (Alcalde, 1999: 57) que el proceso

«ha propiciado que el equipo de maestros tomase conciencia del valor de evaluar el propio trabajo y la repercusión directa que este análisis tiene en la mejora educativa. Los objetos de evaluación escogidos han constituido el motor de un trabajo de reflexión pedagógica que hacía ya tiempo que la mayor parte del claustro deseaba realizar».

A diferencia del modelo del MEC, los proyectos de autoevaluación canario y catalán se plantean desde la autonomía de los centros para decidir la mayor parte de los elementos de la evaluación, contando en el primer caso con el asesoramiento de instituciones *ad hoc* (ICEC y CEP) y disponiendo, en el segundo, de un conjunto de materiales que pueden utilizarse con total flexibilidad y que «no se muestran en términos de tablas o de pautas <<listas para ser usadas>>».

En Andalucía, CIDE (2002: 25-26), todos los centros deben evaluar su propia organización y funcionamiento, los programas y las actuaciones educativas que llevan a cabo y los resultados que obtienen.

Las dimensiones sobre las que se desarrolla la autoevaluación son:

- Diseño y elaboración de los elementos para la planificación del centro (Proyecto de Centro, *Proyecto Curricular* y Plan Anual).
- Proceso de aplicación y desarrollo de los mismos.
- Resultados obtenidos.
- Nivel de satisfacción, en los distintos sectores de la comunidad, por los resultados obtenidos en relación a los objetivos propuestos.

Los responsables de la Dirección General de Evaluación Educativa y Formación del Profesorado del Gobierno andaluz ponen a disposición de los centros modelos de indicadores de calidad, instrumentos y orientaciones para su aplicación. No obstante, el Equipo Técnico de Coordinación Pedagógica del Centro tiene autonomía para determinar los instrumentos y procedimientos a utilizar.

En cuanto a la distribución de cometidos, el Consejo Escolar se encarga de la evaluación del proyecto de centro, del plan anual, de la evolución de los resultados académicos y de la eficacia en la gestión del centro, sin entrar en los aspectos docentes de los mismos. Por su parte, el Claustro de Profesores asume la evaluación del proyecto curricular de etapa y ciclo, de la programación de los procesos de enseñanza-aprendizaje y de los resultados de los alumnos; analizando y valorando, asimismo, los aspectos docentes del proyecto de centro y del plan anual.

El programa de Castilla y León y la propuesta de la Comunidad Valenciana se plantearon desde los presupuestos y principios de los Planes Anuales de Mejora y del Modelo de Gestión de la Calidad que ya hemos descrito con anterioridad.

Estos son los programas de autoevaluación que, en estos momentos, conocemos lo cual no quiere decir que sean los únicos que se estén desarrollando, ya que la descentralización y transferencia de competencias a las Comunidades Autónomas no ha facilitado la disponibilidad de información actualizada sobre estas y otras experiencias educativas. Sin duda, el Ministerio de Educación sigue teniendo el reto de articular un sistema de redes de información y datos que, en conexión con todas las Comunidades Autónomas, permita conocer con fiabilidad y transparencia qué está ocurriendo en el sistema educativo, como ocurre normalmente en cualquier Estado democrático y desarrollado en el siglo XXI.

4. A MODO DE CONCLUSIONES

La evaluación de centros por las administraciones educativas sigue siendo un reto pendiente en la mayoría de los países europeos. Como ya apuntábamos en la Introducción una parte importante de los proyectos que se recogen en las publicaciones de Euridyce no habían sido experimentados en la realidad de los sistemas educativos.

Con respecto a la evaluación de productos hay que señalar que cada vez es más frecuente que los resultados de los alumnos en pruebas oficiales, ya sean de diagnóstico, de exámenes nacionales o de proyectos internacionales, se asuman por los medios de comunicación y por la sociedad en general como un indicador infalible de la calidad de los centros educativos y del profesorado.

BIBLIOGRAFÍA

- AA.VV. (1999): «Un proceso de mejora. El modelo canario de autoevaluación», *Cuadernos de Pedagogía*, n. 283, pp. 58-62.
- ALCALDE, I. Y RÚE, J. (1999): «La autoevaluación. Procedimientos y metodología desarrollados en Cataluña», *Cuadernos de Pedagogía*, n. 283, pp. 50-53.
- ALCALDE I SASTRE, I. (1999): «Llevar el modelo a la práctica. La experiencia del CEIP <<Granullarius>> de Granollers», *Cuadernos de Pedagogía*, n. 283, pp. 53-57
- BELTRÁN, F. Y SAN MARTÍN, A.(1993): «La organización escolar: ¿evaluación o devaluación?», *Cuadernos de Pedagogía*, n. 219, pp. 16-21.
- BOYD, S. (2002): *Approaches to the evaluation of schools which provide compulsory education: The United Kingdom – England*. Eurydice
- BUISSON-FENET, H. et PONS, X. (2011). *Les pratiques d'évaluation externe des établissements scolaires en France, au Royaume-Uni et en Suisse*. Créteil, Université Paris-Est-Créteil.
- CERDÁN, J.; SAHUQUILLO, L.; LUJÁN, J. y PUENTE, J. (1992): «La evaluación externa de centros», *Aula*, n. 6, pp. 65-68.
- CIDE (2002). *Aproximaciones a la evaluación de Centros. España (2000/01)*. Madrid, Centro de Investigación y documentación Educativa (CIDE), MECD, Unidad Española de Eurydice.
- EURYDICE (2004). *Evaluación de los centros de enseñanza obligatoria en Europa*. Madrid, Ministerio de Educación y Ciencia, disponible en: <http://bookshop.europa.eu/en/evaluation-of-schools-providing-compulsory-education-in-europe-pbEC3112831/> , (consultado el 6 de mayo de 2013).
- GREAT BRITAIN. PARLIAMENT. HOUSE OF COMMONS (1997). *Excellence in Schools* (Cm. 3681). London: The Stationery Office.
- LUJÁN CASTRO, J. y otros (1996): *Evaluación de centros docentes. El Plan EVA*. Madrid, Ministerio de Educación y Ciencia, Centro de Publicaciones.
- GUIBERT, N. y REVERCHON, A. (1996): «A chacun de se saisir des outils de l'évaluation. Un entretien avec Claude Thelot, directeur de l'évaluation», *Le Monde de l'Éducation*, n. 240, septembre, pp. 34-38.

- MEC (1988): *Abandono escolar en Enseñanzas Medias*. Madrid, Dirección General de Coordinación y de la Alta Inspección, Servicio de Inspección Técnica de Educación (documento mecanografiado).
- (1992): *Plan de Evaluación de Centros docentes. Niveles no universitarios. Programa piloto para el curso 1991-1992*. Madrid, MEC.
- MEC (1994). *Centros Educativos y Calidad de la Enseñanza. Propuesta de actuación*. Madrid, MEC.
- (1998). *Modelo europeo de gestión de la calidad*. Madrid, Ministerio de Educación y Ciencia.
- MENÉNDEZ I PABLO, J. (1993). «Evaluar los centros», *Cuadernos de Pedagogía*, n. 217, pp. 68-75.
- MIGUEL DÍAZ, M. DE (1997). «La evaluación de los centros educativos: una aproximación a un enfoque sistémico», *Revista de investigación educativa*, 15, n. 2, p. 145-178.
- NEVO, D. (2001). “School Evaluation: Internal or External”, *Studies in Educational Evaluation*, 27, pp. 95-106.
- OFFICE FOR STANDARDS IN EDUCATION (1998). *School Evaluation Matters (Raising Standards Series)*. London, OFSTED.
- OUARDANI, A. (2002): *Modes d'évaluation des établissements scolaires dans l'enseignement obligatoire: France*. Eurydice
- Orden de 9 de junio de 1998 por la que se establece el Plan Anual de Mejora en los centros docentes públicos dependientes del Ministerio de Educación y Cultura y se dictan instrucciones para su desarrollo y aplicación (BOE n. 141, de 13 de junio de 1998).
- SAUNDERS, L. (1999): «Work in <<value-added>> analysis of schools' performance at the National Foundation for Educational Research (NFER)», *Infancia y Aprendizaje*, n. 85, pp. 19-32.
- THÉLOT, C. (1995): *L'évaluation du système éducatif*. Paris, Nathan.
- THOMAS, S. (1998): «Medir el valor añadido de la eficacia de las escuelas en el Reino Unido», *Perspectivas*, vol. XXVIII, n. 1, marzo, pp. 101-122.

ANEXO

Parámetros que integran las 14 categorías de los criterios de evaluación de proceso para la evaluación externa de los centros de enseñanza obligatoria europeos, en los países que disponen de listados preestablecidos.

1. Enseñanza/aprendizaje en el aula

- Currículo/materia
- Métodos de enseñanza
- Evaluación de los alumnos
- Participación de los alumnos
- Actitud de los alumnos hacia el aprendizaje
- Interacción profesor/alumno
- Utilización de los deberes para casa
- Utilización/enseñanza de las TIC
- Tratamiento de las necesidades diferenciadas de los alumnos

2. Orientación y apoyo a los alumnos

- Orientación escolar
- Actividades de recuperación
- Actividad tutorial
- Desarrollo personal de los alumnos
- Desarrollo de las habilidades sociales de los alumnos
- Medidas para mejorar el comportamiento de los alumnos
- Tratamiento de los alumnos con necesidades especiales
- Lucha contra el absentismo

3. Funcionamiento de los órganos/organización del centro

- Junta/Consejo escolar
- Claustro de profesores
- Consejo de clase
- Órganos consultivos
- Sindicatos de estudiantes

4. Política educativa/general del centro

Oferta de cursos
Definición de objetivos
Política de deberes para casa
Proyecto de centro
Reglamento interno del centro
Informe anual

5. Relaciones entre el centro y la comunidad local/relaciones externas

Relaciones con:

- los padres
- otros centros
- las instituciones locales
- los responsables de educación
- la comunidad local en general

Difusión de documentos sobre la oferta educativa, las actividades y/o el rendimiento de los alumnos
Participación en programas locales, actividades en el entorno local
Participación en programas internacionales

6. Gestión de los recursos humanos

Desarrollo profesional
Apoyo a los nuevos docentes
Criterios para la distribución de los alumnos en las clases y/o de las clases a los profesores
Distribución de las tareas no relacionadas con la docencia
Número de docentes y su titulación
Política de salarios y/o existencia de incentivos
Política en caso de que existan puestos vacantes

7. Gestión del tiempo escolar

Duración de las clases
Periodicidad de las vacaciones
Tiempo asignado a cada materia
Gestión del tiempo durante las clases

8. Actividades extraescolares

Organización de estas actividades

9. Evaluación interna

- de la enseñanza
- del centro
- de los profesores/del director

Medidas adoptadas tras la evaluación interna

Participación del personal en la evaluación interna

10. Liderazgo

Actividad del director/equipo directivo

Seguimiento y consecución de objetivos comunes

Cumplimiento del reglamento interno

Mejoras llevadas a cabo desde la última inspección

Coordinación, comunicación interna y/o participación en la toma de decisiones

Gestión de los conflictos

Valores comunes del centro

11. El clima del centro

Relaciones:

- entre los alumnos
- entre los profesores
- profesores/alumnos

12. Gestión de las instalaciones

Utilización del espacio y de la infraestructura de acuerdo con las necesidades educativas

Estado de las aulas/edificios

13. Gestión de los recursos financieros y materiales

Utilización de los recursos operativos (cualitativa)

Utilización/estado de los equipos científicos

Utilización/estado del equipo de las TIC

Gestión del presupuesto salarial
Gestión del presupuesto operativo 2-3
Gestión del presupuesto global (incluidos salarios y recursos operativos)
Gestión de los recursos adicionales y específicos

14. Procedimientos administrativos

Calidad de la gestión administrativa de:

- la inscripción de los alumnos
- el registro de faltas/asistencia
- los expedientes de los alumnos
- los boletines de notas de los alumnos (relativos a su rendimiento)
- los expedientes del personal
- las fichas de evaluación del personal

Tratamiento de las quejas

Cumplimiento de la normativa en relación con los certificados expedidos a los alumnos

Tema 5

La evaluación basada en el centro

Introducción

1. Formas y tipos de autoevaluación institucional
2. Orientaciones y estrategias de autoevaluación
3. Propuestas y experiencias de autoevaluación
 - 3.1. Análisis de algunas experiencias de autoevaluación promovidas por investigadores y proyectos internacionales
4. Presupuestos y condiciones para la realización de la autoevaluación
5. Balance de las experiencias y propuestas de autoevaluación institucional

Bibliografía

TEMA 5
LA EVALUACIÓN BASADA EN EL CENTRO:
LA AUTOEVALUACIÓN INSTITUCIONAL

Eustaquio Martín Rodríguez.

UNED

INTRODUCCIÓN

A diferencia de la evaluación externa que normalmente llevan a cabo los supervisores, la autoevaluación institucional o escolar —denominaciones que utilizaremos indistintamente— es una modalidad en donde el protagonismo de la evaluación es asumido por los centros y los distintos actores educativos. Ellos son los que acometen la responsabilidad de llevar a cabo la misma, ya sea siguiendo las orientaciones que les vienen dadas en los proyectos desarrollados por instancias ajenas a los propios centros, o bien generando por sí mismos estudios o experiencias en torno a los problemas que consideran relevantes de la vida institucional.

Las primeras propuestas de autoevaluación institucional surgieron en los últimos años de la década de los 70 como consecuencia de las políticas anglosajonas de la *accountability*, o rendición de cuentas, y sin solución de continuidad fueron extendiéndose hasta generalizarse con las políticas de descentralización educativa en los años 90. En ellas cabe distinguir, al menos, dos tipos o formas de autoevaluación. Por un lado, las experiencias que ya entonces fueron, y siguen siendo hoy, promovidas por las administraciones educativas con el propósito de que los centros y profesores informen públicamente de sus quehaceres y resultados, así como para verificar el ritmo de implantación y los cauces de desarrollo que van siguiendo las reformas escolares en curso. Y, por otro, las que emergieron vinculadas a los círculos académicos-universitarios en conexión con determinadas propuestas de formación permanente del profesorado —como la formación en centros— y con ciertos proyectos auspiciados por movimientos como la mejora de la escuela, la Investigación-acción o la revisión basada en la escuela, para apoyar y reforzar los estudios que, emprendidos por los propios actores, privilegian el desarrollo profesional e institucional.

Ambas formas o tipos de autoevaluación tienen plena actualidad en estos momentos. Ya desde el último tercio del siglo XX se han ido genera-

lizando las políticas gubernamentales orientadas al control y evaluación de la calidad de los sistemas educativos y, asimismo, los centros de enseñanza han concitado el interés de distintas líneas de estudio e investigación por la importancia que tienen los mismos en los procesos de cambio e innovación en la educación.

1. FORMAS Y TIPOS DE AUTOEVALUACIÓN INSTITUCIONAL

Como se ha señalado con anterioridad, los enfoques de autoevaluación surgen en un contexto de demandas y exigencias de que los centros escolares informen públicamente de sus logros y resultados. Las políticas de la *accountability* generaron distintas formas de rendición de cuentas y, consecuentemente, de evaluación. McCormick y James (1996:33) precisan que en el Reino Unido se podían distinguir, al menos, tres formas:

- «1. *Capacidad de responder ante los propios clientes (answerability)*, o sea, los alumnos y sus padres (rendición de cuentas moral).
2. *Responsabilidad (responsability) ante uno mismo y sus compañeros* (rendición de cuentas profesional).
3. *Rendición de cuentas (accountability)*, en sentido estricto, ante los propios empleadores o dirigentes políticos (rendición de cuentas contractual)».

La rendición de cuentas contractual es la forma que, en sentido estricto, adopta la *accountability* en las relaciones establecidas entre la administración educativa y los centros docentes. Se trata esencialmente de una modalidad formal-legalista, de carácter prescriptivo, con una única dirección, de los centros a los responsables políticos o empleadores, y con un fuerte componente de control. Alternativamente, las formas de responder ante los padres y alumnos y la de responsabilidad (ante uno mismo y los compañeros), es decir, las formas de rendición de cuentas moral y profesional, son iniciativas que emergen desde los propios centros como alternativas a la de relación contractual.

Helen Simons (1985: 2) afirmaba, a mediados de los 80, que era la autoevaluación

«... un movimiento en estado embrionario y, desgraciadamente, una de las características de este estado es la confusión y la contradicción mientras se

superponen múltiples etiquetas: autoevaluación escolar, evaluación internivelar, evaluación dentro del centro, evaluación basada en el centro, evaluación institucional, evaluación interna, autocrítica escolar, autorevisión, autoestudio, autoseguimiento... En el centro de toda esta confusión existe una incógnita política que tal vez podríamos expresar como pregunta dual: ¿Es la autoevaluación una contraoferta de las instituciones amenazadas a las demandas coactivas del gobierno para intervenir en el currículum, o es el resultado de un movimiento iniciado por el gobierno para abrir las escuelas a la opinión pública?».

El paso del tiempo ha dejado muy claro que el sentido de las propuestas de autoevaluación era, en la pregunta de Helen Simons, tanto una contraoferta de las instituciones como el resultado de las políticas gubernamentales de rendición de cuentas. Otro tanto ocurre con las modalidades o tipos, en la medida que la autoevaluación es susceptible de posiciones ideológicas dispares, y pueden observarse entre unas y otras formas no pocas diferencias y desencuentros.

James (1987: 186), considerando las dimensiones operativas y los constructos teóricos de la autoevaluación, distinguía al menos dos tipos con características muy diferentes entre sí, según se recoge en el siguiente Cuadro 5.1.

Cuadro 5.1. Tipos de autoevaluación institucional

TIPO A: AUTOEVALUACIÓN INSTITUCIONAL BASADA EN LA AUTORIDAD	
Dimensiones operativas	
Participación y control interno:	conducción jerárquica y centralizada
Propuestas:	accountability y revisión curricular
Organización:	gestión racional
Foco:	antecedentes y resultados (input-output)
Métodos:	reuniones y datos cuantitativos principalmente
Constructos teóricos	
Modo de accountability:	legal-formal (explícito)
Perspectiva sociopolítica:	positivista o teoría de sistemas
Tradición de evaluación:	análisis de sistemas, objetivista, modelos de gestión tecnológico

TIPO B: AUTOEVALUACIÓN INSTITUCIONAL BASADA EN LA RESPONSABILIDAD**Dimensiones operativas**

Participación y control interno:	sin distinción de estatus (no jerárquica)
Propuestas:	desarrollos profesional y curricular
Organización:	colegiada
Foco:	procesos
Métodos:	discusiones y datos cualitativos principalmente

Constructos teóricos

Modo de <i>accountability</i> :	moral (implícita)
Perspectiva sociopolítica:	interpretativa o humanística
Tradicción de evaluación:	evaluación de proceso, responsiva, democrática iluminativa

Fuente: James (1987: 186).

Como puede comprobarse, los supuestos, orientaciones y características de uno y otro tipo de autoevaluación son totalmente distintas. La basada en la autoridad se asocia a la *accountability* de relación contractual, es decir, a la potestad del Estado o los gobiernos de controlar lo que ocurre en las escuelas y, por tanto, a la obligación de éstas de rendir cuentas de sus logros y resultados. En el Reino Unido, de las tres orientaciones que influyeron en la autoevaluación escolar —el movimiento de reforma curricular, la formación permanente del profesorado y la rendición de cuentas— la más importante fue sin duda la de la rendición de cuentas auspiciada por las administraciones educativas. Tanto el entonces Departamento de Educación y Ciencia (DES), con sus iniciativas de revisión curricular, como las propias autoridades locales (LEA), con sus esquemas de autoevaluación escolar, crearon las condiciones que impulsaron la *accountability*.

Así, por ejemplo, las LEA desarrollaron durante los años setenta una serie de orientaciones formales para la autoevaluación de las escuelas. Consistían, por lo general, en listas de preguntas que, preparadas por los administradores y asesores externos, debían ser respondidas por los directores y profesores de las escuelas. Bajo la retórica del desarrollo escolar y del desarrollo del profesorado, estos esquemas y preguntas de autoevaluación tenían claramente la finalidad de obtener información de lo que ocurría en

las escuelas. Simons (1999: 228) afirmaba rotundamente que «apelar a esos enfoques de autoevaluación escolar es una burla de un concepto esencialmente educativo». Y, como hemos resaltado en otro momento, John Elliott (1986: 254) planteaba

«[las] listas cumplen una importante función de control dentro de la autoevaluación técnica. En primer lugar, actúan a modo de recordatorios de las reglas prescritas y, en segundo lugar, permiten al profesor comprobar rápidamente su actuación en comparación con tales reglas. Las listas son esencialmente un instrumento para incrementar la eficiencia técnica y el control en lugar de la capacidad para tomar decisiones profesionales reflexivas».

En la autoevaluación basada en la autoridad afloran distintos dilemas y controversias. Nutall (1985: 24) planteaba las tensiones que generaba la autoevaluación en función de que los enfoques o motivos de la misma fueran internos o externos, ya que, por un lado, la «evaluación implica la crítica y la búsqueda de cambios para mejorar la práctica y, por lo tanto, resulta hostil para las organizaciones y formas de trabajo establecidos», y, por otro, si «el motivo principal de la evaluación es externo a la organización, entonces la tensión más importante se dará entre los valores internos (de carácter profesional) y los valores externos (que tienden a ser utilitaristas)». Estas tensiones entre los propósitos externo e interno de la autoevaluación son los que, de alguna manera, van a estar presentes en las distintas experiencias que van a desarrollarse en esta modalidad de evaluación institucional.

Las directrices político-administrativas de la autoevaluación para la rendición de cuentas se encontraron entonces, en el Reino Unido, con un bagaje de experiencias y realizaciones provenientes tanto del legado del desarrollo curricular como del replanteamiento de la formación permanente del profesorado que configurarían un espacio y un clima muy diferentes para el debate sobre las consecuencias de este tipo de orientaciones. Las aportaciones de Stenhouse (1975) sobre la consideración del «profesor como investigador», las propuestas de “evaluación democrática” de MacDonald (1976) y de la «evaluación como iluminación» de Parlett y Hamilton (1976), entre otras, así como los intercambios de trabajos e ideas con autores norteamericanos como Ernest House y Robert Stake sobre las realizaciones de la accountability y los avances en el campo de la evaluación, propiciaron un contradiscurso, de carácter académico, desde el que se defendería la autoevaluación desde una perspectiva de desarrollo profesional de los docentes.

La motivación y razón de estos enfoques alternativos están vinculadas más con las necesidades internas de la escuela que con las presiones del exterior. Los promotores de la autoevaluación basada en la responsabilidad no eluden el tema del control, ni creen que la escuela deba estar exenta de informar y rendir cuentas a la sociedad de sus quehaceres, pero al mismo tiempo sostienen que este proceso debe ser conducido democráticamente en beneficio de la institución y de los profesores, para que pueda darse un compromiso de los mismos en la mejora de prácticas escolares y en el desarrollo institucional.

Diversas iniciativas académicas como los movimiento de «investigación-acción», la «mejora de la escuela» y la «revisión basada en la escuela» han venido desarrollando otras maneras de concebir y realizar la evaluación institucional. La consideración del centro educativo como unidad básica del cambio y de los profesores como los agentes primordiales del mismo, refuerzan la idea de la autoevaluación como una propuesta que, trascendiendo su carácter evaluativo, se incardina en los procesos de innovación y formación de las instituciones educativas. Esta es la forma y, a la vez, el gran reto al que se enfrenta la autoevaluación escolar, al igual que ocurrió antes con el movimiento de investigación-acción, para superar los problemas que se derivan de su percepción como iniciativas externas a la propia institución. Por ello, Simons (1999: 239) se reafirma en la idea de que la «autoevaluación escolar para el desarrollo institucional a largo plazo ha de basarse en el análisis de las metas, objetivos y valores de la misma institución».

La evaluación de los centros para los procesos de mejora y desarrollo institucional han fundamentado sus propuestas en el protagonismo que asumen las instituciones y los profesores en las teorías del cambio. Helen Simons (1995: 221-222) nos recuerda algunas de las lecciones extraídas de la evaluación externa de la reforma curricular centralizada de cara a apoyar el desarrollo de las escuelas democráticas. Todos los aportes que lleguen a ellas desde el exterior son importantes, pero es necesario prestar una atención especial a la institución en cuanto lugar en el que se produce el cambio y desde el que puede abordarse las prácticas educativas realizadas en las aulas. La necesidad de vincular el desarrollo curricular con el desarrollo del profesorado, hunde sus raíces en el principio de que el cambio curricular no proviene de las prescripciones del currículo sino del análisis y estudio que los profesores hacen de sus prácticas profesionales.

Algo similar plantea el movimiento de la Mejora de la escuela ya que, en tanto que es una propuesta asumida y desarrollada por los docentes, se ha preocupado preferentemente por crear las condiciones y poner en juego las estrategias que posibiliten el desarrollo institucional y profesional como componentes indisociables de la calidad de la educación. Su perspectiva cultural de la innovación, se basa fundamentalmente en la capacidad de los actores educativos (docentes, alumnos, administradores, etc.) para crear y recrear el mundo social, en la medida que los mismos no están predeterminados sino que interactúan con un relativo grado de autonomía y libertad. La interacción requiere la atribución de significados a la propia conducta y a la conducta de los demás porque en definitiva toda interacción es a la larga una negociación entre significados e interpretaciones subjetivas de la realidad social y escolar.

Como nosotros mismos hemos señalado en distintos momentos (Martín Rodríguez, 1988, 1989, 1996, 2002), la autoevaluación institucional es antes que nada un proceso de valoración crítica del quehacer de la institución educativa realizado **en** y **desde** la propia institución. Se trata esencialmente de que los centros educativos inicien y desarrollen un análisis autocrítico de su funcionamiento y trabajo habituales en donde se pongan en cuestionamiento las distintas formas de proceder docentes.

Desde estas posiciones, es posible pensar en la escuela como una comunidad de aprendizaje, como organizaciones capaces de aprender. Los centros docentes necesitan reforzar la dimensión institucional de sus quehaceres educativos, al tiempo que crean espacios compartidos de aprendizaje y desarrollo institucional. Es a través del conocimiento y toma de conciencia de sus avances pero también de sus deficiencias, contradicciones e incoherencias como las instituciones escolares pueden definir y reforzar cursos de acción alternativos en los que se vaya logrando una simbiosis entre sus quehaceres o realizaciones y sus proyectos o aspiraciones.

2. ORIENTACIONES Y ESTRATEGIAS DE AUTOEVALUACIÓN

La evaluación interna auspiciada, asesorada y hasta dirigida por instancias ajenas a los centros y profesores ha sido con frecuencia una forma de rendición de cuentas encubierta para complementar u obtener la

información que no era posible lograr a través de la evaluación externa. Así, al menos ocurrió en EE.UU. e Inglaterra durante los años 70 y 80, y en esta dirección cabe interpretar las iniciativas de autoevaluación que desde los años noventa se están impulsando en numerosos países, una vez que se ha comprobado que la evaluación de todos los centros educativos requiere una dedicación y la disponibilidad de equipos profesionales que, en la mayoría de los casos, no están al alcance de los organismos que las promueven. Por ello, tanto en los años setenta como hoy tienen plena vigencia las sospechas de si la autoevaluación de las escuelas no es sino una forma de rendición de cuentas con un rostro más humano, tal y como Nuttall adelantaba, ya en 1985, con el título de uno de sus libros: *School self-evaluation. Accountability with a human face?*

Ante estos problemas, un autor como David Nevo (1998: 91) defiende una propuesta que denomina evaluación centrada o basada en la escuela que no puede concebirse «ni como sinónimo de evaluación interna ni como antónimo de evaluación externa, sino como una combinación de ambas». Desde sus planteamientos considera imprescindible que exista una evaluación interna para que pueda darse una utilización constructiva de la evaluación externa. Es más, para él,

«una escuela que no disponga de un mecanismo de autoevaluación tendrá dificultades para desarrollar una actitud positiva hacia el proceso de evaluación y carecerá de la confianza necesaria para entablar un diálogo constructivo con el equipo de evaluación externa».

En las experiencias llevadas a cabo en Israel, Nevo (1997: 58) nos informa que una de las formas de introducir la evaluación externa de los centros ha sido impulsando previamente la evaluación interna mediante la formación de los directores y profesores, la constitución e institucionalización de equipos de evaluación en los centros con capacidad para decidir los objetos de evaluación, y una vez operativizada la evaluación interna, pero no antes, iniciar proyectos de evaluación externa.

Frente a esta posición, otros autores se cuestionan que la autoevaluación sea una actuación complementaria de las evaluaciones externas llevadas a cabo por los servicios de supervisión e inspección. Saunders (1999: 425) advierte «que si los mismos enfoques se utilizan como base tanto para la evaluación externa como para la autoevaluación, se corre el riesgo de caer en una monocultura [...] y de conseguir mejoras que lo son sólo en apariencia». Helen Simons (1999: 205-206) señalaba, asimismo, la necesidad de que la

autoevaluación se separe de las respuestas a las exigencias de rendición de cuentas porque, en tanto las escuelas no adquieran y consoliden una cultura de evaluación, este tipo de procesos puede seguir percibiéndose como amenazantes, los autoinformes pueden ser utilizados como auditorías, no se está seguro del uso que puede darse a los informes, etc.

La clonación de la autoevaluación de los proyectos de evaluación externa, y el hecho de que las propuestas tengan un formato cerrado, conduce a que los centros sean objeto pero no sujetos de evaluación. Un aspecto central que no pasa desapercibido es la negación de la capacidad e iniciativa de las instituciones para determinar sus propias propuestas y metodologías de evaluación. Todo les viene dado de antemano, limitándose su responsabilidad a cumplimentar los protocolos de recogida de información, cuestionarios o entrevistas. Este modo de proceder dirigista impide que los distintos actores institucionales decidan por sí mismos y acuerden en qué se va a centrar la evaluación, cuáles van a ser los límites y alcances del estudio, quiénes van a participar y de qué forma en el proyecto, cómo se va a gestionar y conducir todo el proceso, etc. Y, además, pone en evidencia las posibles respuestas que pueden darse a la pregunta de ¿a qué o quién se refiere el 'auto' de la modalidad de autoevaluación?

El problema de la generalización de este tipo de experiencias es que no facilitan o hacen más complicado el que las escuelas inicien autónomamente estudios sobre los problemas o temas que, en realidad, les preocupa o interesa en el acontecer cotidiano de la vida institucional. Una institución cualquiera puede estar preocupada o interesada, en un momento concreto, por los altos índices de fracaso escolar de los alumnos, o por los problemas de disciplina que se dan en determinados grupos y asignaturas, o por las dificultades y obstáculos que impiden la coordinación de los profesores de un mismo curso o de distintos seminarios. Todas estas dinámicas u hechos tienen la suficiente entidad e importancia para que emerjan como focos de un proceso de autoevaluación que, en la mayoría de los casos, son más relevantes y significativos en tanto que conectan con los problemas más usuales con los que se convive en el día a día, al tiempo que permiten conocer de manera más sistemática qué está ocurriendo en los centros.

Gather Thurler (2002), en un artículo en el que analiza las virtualidades de la autoevaluación como motor de los procesos de cambio, propone una clasificación exhaustiva de los tipos de autoevaluación institucional en función de los elementos a los que se da prioridad o delimitan el objeto de evaluación.

Cuadro 5.2. Dimensiones de las propuestas de autoevaluación institucional

Objeto de evaluación	Objetivos y procedimientos	Ventajas / Limitaciones
<i>Centrado sobre las finalidades del sistema educativo en su globalidad</i>	Verificación de la adecuación de las prácticas a los objetivos previamente determinados a través de informes.	Las escuelas deben aceptar los programas oficiales como único criterio de referencia. Este tipo de evaluación permite identificar las lagunas, sin que estén obligadas a poner en evidencia sus causas, condición necesaria para poder determinar las soluciones adecuadas
<i>Centrado sobre las prioridades propuestas por las escuelas y recogidas en sus proyectos.</i>	Control del cambio previsto, por ejemplo en lo que concierne a la modificación de estructuras de gestión internas. Introducción de una planificación a largo plazo en combinación con un plan anual, así como con una evaluación de proceso y de los efectos de la actividad desarrollada	La asociación entre las prioridades propuestas y su realización favorece la regulación continua. Su implementación individualizada en las escuelas no permite estudios comparativos, ni la obtención de datos para conducir el sistema educativo.
<i>Centrado sobre los criterios de eficacia establecidos por la investigación.</i>	Las autoridades escolares se basan en los resultados de la investigación para el desarrollo de programas de acción que faciliten la generación del cambio. La autoevaluación es conducida a través de listas de control y herramientas desarrolladas por los investigadores.	La confrontación entre los criterios teóricos y los criterios locales basados en las 'teorías prácticas' de los actores educativos pueden generar fuertes conflictos, en vez de facilitar la construcción del consenso y permitir el establecimiento de las prioridades del cambio

Fuente: Gather Thurler (2002)

En el Cuadro 5.2 presentamos una versión adaptada de la que ella hace. Nosotros los hemos reducido a tres tipos, dado que recogen con carácter general las formas más habituales de emprender esta modalidad de evaluación.

De los tres tipos que se recogen, el primero y el tercero encajan dentro de las propuestas en las que la capacidad de decisión de las instituciones es sumamente limitada ya que todo el proceso queda condicionado al ser instancias externas a ellas quienes delimitan el objeto y establecen los criterios de evaluación. En un caso, lo que se persigue es informar de cómo se están desa-

rollando las políticas educativas en la realidad de las aulas, mientras que, en el otro, se pretende comprobar las formas de organización y modos de funcionamiento de los centros que la investigación ha identificado por su influencia en el logro de altos niveles de rendimiento escolar. Aunque en la clasificación de Monica Gather el primer tipo de evaluación se refiere fundamentalmente a la valoración que, de modo más informal que sistemático, hacen los centros en sus Memorias anuales, cabría también integrar en el mismo todas aquellas propuestas de formato cerrado en las que los cometidos de las instituciones consisten en responder o cumplimentar los diferentes protocolos elaborados para recoger una información previamente determinada.

Las experiencias de evaluación externa e interna basadas en los resultados de ciertas líneas de investigación han adquirido, asimismo, una presencia destacable en los sistemas educativos. Desde finales de los años 80, a raíz de la difusión de las aportaciones y resultados de la investigación de la eficacia escolar, por organismos internacionales como la OCDE, se han ido elaborando propuestas de evaluación articuladas en torno a los elementos curriculares y organizativos que aparecen con regularidad en los centros escolares cuya eficacia se pone de manifiesto por los resultados escolares que alcanzan sus alumnos.

En el segundo tipo, en cambio, son los centros escolares los que deciden y proponen sus dimensiones de evaluación entre los elementos prioritarios que integran su proyecto institucional. La autoevaluación, en este caso, es una herramienta más incorporada en su planificación que va aportando información y datos para la autorregulación de sus actuaciones. De esta manera, la conducción de los procesos de cambio y mejora propuestos por los propios centros se sirven del conocimiento que van obteniendo para continuar o modificar el curso de las dinámicas y trabajos emprendidos.

3. PROPUESTAS Y EXPERIENCIAS DE AUTOEVALUACIÓN

Desde finales de los años 80 y durante la década de los 90 no han sido pocos los países que han abordado la modernización de sus sistemas educativos, incorporando la evaluación de centros en sus proyectos de reforma. Igualmente durante este tiempo se han ido desarrollando distintos proyectos, en el contexto internacional, con el fin de promover el intercambio de metodologías y experiencias de evaluación. Por lo general, la modalidad de

autoevaluación se ha ido introduciendo, junto con la evaluación externa, en la mayoría de los sistemas educativos de los países desarrollados y en vías de desarrollo.

Hoy es posible contar con una buena cantidad de experiencias de autoevaluación, tanto de las promovidas por las administraciones educativas como de las llevadas a cabo por instituciones, investigadores y organismos internacionales. La clasificación de las mismas es una tarea difícil, ya que no es posible su categorización en función de criterios homogéneos y podemos encontrarnos que dentro de un mismo tipo o modalidad se dan variaciones sustantivas en relación con el objeto, los propósitos, metodologías, etc., de la evaluación. Por ello hemos optado por presentar las mismas en dos grandes apartados dedicados, uno, a las promovidas por las administraciones educativas y, el otro, a las realizadas por investigadores y organismos internacionales. En uno y otro apartado se recogen experiencias diversas que pueden centrarse o bien en evaluaciones de proceso o de resultados, estar planteadas o bien como propuestas complementarias de evaluaciones externas o ser independientes de ellas, con propósitos de verificar la puesta en práctica de determinadas políticas o de acometer procesos de mejora y desarrollo institucional, etc.

Dentro del Programa Sócrates (MacBeath, Schratz, Meuret y Jakobsen, 2000), en el bienio 1997-98 se ha realizado el programa de Evaluación de la Calidad en la Enseñanza Escolar con el objetivo de desarrollar procesos de cambio y de mejora de la calidad de los centros. En esta experiencia en la que participaron 101 escuelas de educación secundaria de 18 países, de las que 10 eran españolas, se trató de que cada una de las escuelas explorara su nivel de calidad mediante un diagnóstico realizado a través de una guía de autoevaluación, para posteriormente, y en función de los avances de cada centro en los procesos de autoevaluación, elaborar un plan de actuación de cara a mejorar sus logros.

Las dos partes del programa, es decir, la realización del diagnóstico y la elaboración de planes de mejora se constituyeron en dos subproyectos, ya que no todos los centros tenían experiencia o estaban iniciados en los procesos de autoevaluación, por lo que los que participaron en el subproyecto 1 centraron su actividad en los métodos de evaluación, y los del subproyecto 2 se centraron en el aprovechamiento de los resultados de la autoevaluación para generar la mejora de su calidad. El perfil de

autoevaluación, igual para todos los centros, comprendía 12 áreas de la vida de la escuela agrupadas en cuatro grandes bloques: I. Resultados (1. rendimiento académico; 2. desarrollo personal y social; 3. trayectoria de los alumnos). II. Procesos en el aula (4. el tiempo como recurso para el aprendizaje; 5. calidad de la enseñanza y el aprendizaje; 6. Apoyos para las dificultades en el aprendizaje). III. Procesos en la escuela (7. la escuela como espacio de aprendizaje; 8. la escuela como espacio social; 9. la escuela como espacio profesional), y IV. Entorno (10. escuela y hogar; 11. escuela y comunidad; 12. escuela y trabajo).

De los diez centros españoles que participaron seis eran públicos y cuatro privados, y de las 12 áreas que recogía el perfil de evaluación se acordó que la autoevaluación se centraría en las siguientes: 1 (éxito escolar), 2 (desarrollo personal y social), 5 (calidad del aprendizaje y de la enseñanza), 7 (el centro como lugar de aprendizaje) y 10 (el centro escolar y el hogar), aunque los centros podían seleccionar otras de su interés.

El proyecto ofrecía una guía para la autoevaluación y dejaba libertad a cada país para poder utilizar otras estrategias de evaluación. En la experiencia española, se optó por utilizar el Modelo Europeo de Gestión de Calidad elaborado por el MEC, en vez de la guía del proyecto, porque, aunque son propuestas totalmente distintas, «los centros acogieron favorablemente la formación que se les podía proporcionar al iniciarles en el modelo europeo de gestión de calidad», (INCE, 1999: 33). Dos representantes de cada uno de los centros participaron en un curso de formación de 13 horas, proporcionándoseles además a los centros una serie de recursos e instrumentos para la evaluación. Así, por ejemplo, para el área 1 (éxito escolar), se empleó una parte del estudio estadístico de los *Resultados de la evaluación de los alumnos de Educación Primaria y Secundaria*, que forma parte de la evaluación externa del sistema educativo español; para el área 2, de desarrollo personal y social, se propuso como referencia la segunda parte del cuestionario para el alumnado que figura en el plan EVA; para el área 5 se utilizó el bloque de procesos didácticos que figura en la primera parte de ese mismo cuestionario, así como la guía de observación sistemática del aula que figura en el plan EVA, etc.

El informe de los procesos seguidos por los diez centros españoles participantes, dados su enfoque y la forma concreta como se llevó a cabo, no se parece en nada al relato que hacen MacBeath, Schratz, Meuret y Jakobsen

(2000) de la historia de una escuela vista a través de los ojos de una alumna, sus padres, un profesor, la directora de la escuela y una amiga crítica.

3.1. Análisis de algunas experiencias de autoevaluación promovidas por investigadores y proyectos internacionales

Empezaremos por la propuesta de un autor como Stufflebean (1996) que desde el marco de una dirección de centros participativa defiende la necesidad de la evaluación interna de cara a fundamentar la toma de decisiones en el interior de las instituciones, estableciendo los elementos sustantivos que deben considerarse para, primero, definir una perspectiva apropiada para la evaluación basada en las escuelas, y establecer, a continuación, los pasos que deben seguir las mismas para hacer realidad esta perspectiva. En su propuesta, es de destacar que la evaluación debe estar al servicio de todos los alumnos, independientemente de sus niveles de éxito; no basarse únicamente en las calificaciones de los alumnos en los exámenes o tests, sino considerar todos los elementos de la escuela; abordarse desde una perspectiva longitudinal y, por tanto, con una continuidad en la recogida de los datos; ser consistente desde una perspectiva teórica; que aúne la rendición de cuentas y la mejora; que clarifique los valores de la escuela para que desde ellos se establezcan los criterios de evaluación y los instrumentos y técnicas de evaluación; que promueva la excelencia y la igualdad de oportunidades; que busquen la credibilidad, imparcialidad y justicia, etc.

La propuesta de Stufflebean, a mitad de camino entre las basadas en la autoridad y la responsabilidad, recoge en gran medida los avances teóricos por los que ha caminado la evaluación profesional norteamericana. Sobre todo en la sensibilidad a los valores, en la subordinación de los aspectos metodológicos a los criterios y necesidades de los centros, en las dimensiones éticas y políticas de la evaluación, etc. Y además es todo una apuesta por una modalidad de evaluación que no ha sido precisamente de las más potenciadas en EE.UU.

David Nevo (1997. 153-163), por su parte, establece seis grupos de indicadores generales, con un número variable de indicadores específicos en cada uno de éstos, para evaluar la calidad de un centro y desde ellas comprender las maneras en que puede ser mejorado. En el Cuadro 5.3, recogemos todos los indicadores propuestos.

Cuadro 5.3. Indicadores para la evaluación de un centro educativo.

Indicadores generales de evaluación del centro	Indicadores específicos
1. Comunidad y alumnado a quienes presta servicio el centro docente.	1.1.tipo de comunidad a la que presta servicios, 1.2. implicación de la comunidad/padres, 1.3. procedencia socioeconómica del alumnado, 1.4. nivel de aptitud del alumnado, 1.5. grupos especiales de alumnos
2. Visión del centro	2.1. metas y objetivos, 2.2. perspectiva pedagógica, 2.3. criterios de admisión, 2.4. criterios de integración, 2.5.perspectivas sobre autoridad y responsabilidad, 2.6 políticas de evaluación, 2.7 estilos de dirección
3. Personal del centro	3.1. calidad del profesorado, 3.2.calidad de los directivos, 3.3. calidad del personal de apoyo, 3.4.ratio profesor / alumno, 3.5.estabilidad del personal, 3.6formación permanente y desarrollo profesional
4. Recursos materiales	4.1.tamaño del centro, 4.2.presupuesto, 4.3 tamaño de las clases, 4.4.horas de enseñanza por clase, 4.4. espacios dentro del aula, 4.5. laboratorios y bibliotecas, 4.6. equipo informático, 4.7. instalaciones deportivas y de recreo, 4.8. zonas ajardinadas
5. Actividades y programas educativos	5.1. principales programas educativos, 5.2. programas selectivos y de alta calidad, 5.3. programas de atención a la diversidad, 5.4. actividades extraescolares, 5.5. tutorías y programas de apoyo para alumnos con necesidades educativas especiales, 5.6. proyectos innovadores / experimentales, 5.7. participación en proyectos nacionales o regionales, 5.8 programas para la implicación de padres / comunidad, 5.9. actividades de autoevaluación y de programación
6. Logros del centro	6.1. tasas de abandono / retención de alumnos, 6.2. resultados del alumnado en tests nacionales o regionales, 6.3. resultados del alumnado en proyectos de investigación, 6.4. resultados en tests preparados por el centro, 6.5. resultados en competiciones nacionales, 6.6. resultados de los antiguos alumnos, 6.7. juicios críticos de inspectores y supervisores, 6.8.satisfacción de los padres, 6.9. reconocimiento nacional y premios concedidos al centro

Fuente: David Nevo (1997. 153-163)

Si observamos la relación de aspectos de la vida institucional que se recogen, no cabe duda que, como el mismo Nevo reconoce, se trata de una propuesta maximalista. A partir de ella, y dentro de su defensa de la evaluación basada en la escuela, se pueden llevar a cabo evaluaciones externa e interna sin que, por otra parte, cualquiera de los proyectos tengan necesariamente que abarcar todos los indicadores. Un centro, por ejemplo, puede iniciar un proceso de autoevaluación centrándose en los indicadores 1 y 2, a través de los que llevaría a cabo un estudio diagnóstico del contexto comunitario y características de su alumnado así como de algunos de los elementos de su proyecto educativo, dejando para otra ocasión el resto de los indicadores.

El desarrollo de los enfoques de autoevaluación también han sido impulsados en el contexto europeo de las políticas de evaluación de centros. La Unión Europea, con el Programa Sócrates, el Consorcio de Instituciones para la Investigación y el Desarrollo Educativo en Europa (CIDREE) y otros muchos organismos e instituciones nacionales han financiado en los últimos años toda una serie de programas destinados a fomentar la autoevaluación en las escuelas europeas. Las experiencias llevadas a cabo son muy diversas. Veamos algunas más en detalle.

John Macbeath (1999) por encargo de la National Union of Teachers ha venido trabajando en la elaboración de un modelo de autoevaluación para el contexto inglés / galés, tomando como base el desarrollado en Escocia en 1990. El trabajo es sumamente interesante, pues recoge las expectativas, necesidades y opiniones de todos los actores de la escuela (profesores, directores, alumnos, padres), trata de superar las limitaciones de las listas de indicadores de resultados y de los informes de los inspectores en la medida que ambos cuentan una historia parcial. En su propuesta de un marco para la autoevaluación, en el que se va describiendo cómo puede generarse un proceso de autoevaluación (creación del clima, acuerdos necesarios, establecimiento de criterios, desarrollo de herramientas, elementos sobre enseñanza y aprendizaje), defendiendo la presencia en el mismo de una amigo crítico, establece 10 indicadores en los que concreta, para cada uno de ellos, los elementos, evidencias cuantitativas, evidencias cualitativas y métodos / instrumentos con los que llevar a cabo su evaluación. Los diez indicadores que utiliza, Macbeath (1999: 131-149) son: clima de la escuela, relaciones, clima de la clase, apoyos para el aprendizaje, apoyos para la enseñanza, tiempos y recursos, organización y comunicación, equidad, reconocimiento de rendimiento, vínculos entre la familia y la escuela.

El proyecto Innovative Approaches in School Evaluation (INAP), financiado también por el programa Sócrates, recoge un conjunto de experiencias de evaluación de centros realizadas en Inglaterra, Escocia, Holanda y España. En su informe final (Tiana, 1999), se describen las modalidades de evaluación desarrolladas en los distintos centros participantes. Los informes de cada país, y de las tres Comunidades (Madrid, Cataluña y Canarias) que integran la experiencia española, plantean las características generales del enfoque de evaluación adoptado para pasar, después, a analizar la realización de la evaluación de un centro concreto. Aunque todos se presentan como propuestas de evaluación interna, habría que señalar que en los enfoques de medidas de valor añadido (casos de Inglaterra y del Equipo Redes de Madrid), los procesos de autoevaluación se pretenden generar a través del feedback establecido con los centros en la devolución de los datos cuantitativos y resultados, cosa que no ocurre en los enfoques de evaluación del resto de las experiencias que desde el principio se definen como de autoevaluación (casos de Escocia, Holanda, y de Cataluña y Canarias en España).

Los enfoques que utilizan el sistema de medidas de valor añadido de la eficacia de la escuela son aquellos que, como Thomas (1998: 104) aclara, se centran

«[en] la medida del progreso medio de los alumnos a lo largo de un periodo de tiempo dado (que en las escuelas de enseñanza secundaria se extiende, por regla general, desde el momento de ingreso en el centro docente hasta los exámenes oficiales, y en las escuelas primarias abarca algunos años específicos) en una escuela determinada, en comparación con otras escuelas de la misma muestra».

El procedimiento que se sigue para establecer estas medidas consiste en determinar el nivel de conocimiento inicial y final de los alumnos, a través de una serie de pruebas normalizadas (pretest-postest), y comparar el grado de adquisición de conocimientos adquiridos tanto en el interior de cada escuela como en relación con alumnos de características similares de otras escuelas. Para ello se utilizan técnicas estadísticas de modelización de múltiples niveles con las que, controlando variables continuas y discretas del alumnado (capacidades, edad, sexo, etnia, etc.), se trata de medir los efectos de la escolarización.

Cuadro 5.4. Principios Directrices para la Autoevaluación Escolar (CIDREE).

AUTOEVALUACIÓN ESCOLAR:

- Es un criterio de cualquier escuela “eficaz”
- **Implica valores y cuestiones éticas** al igual que *técnicas*
- Forma parte del *espectro de la evaluación de calidad del sistema educativo de un país*
- Mejora la relación entre las escuelas y los organismos externos

La autoevaluación escolar puede:

- **Utilizar enfoques creativos e imaginativos, ej. fotografía**, al igual que *técnicas de investigación, ej. cuestionarios, entrevistas y datos numéricos*
- **Comenzar partiendo de un espíritu de curiosidad** al igual que de *una intención de mejorar o desarrollar la escuela*
- **Ofrecer unas mejores perspectivas de cómo son realmente las cosas** al igual que *indicadores para la acción y cambio*

La autoevaluación escolar debe:

- **Ser “democrática”**, por ejemplo, promover las creencias y prácticas democráticas mediante la consulta y la negociación
- Ser “responsable”, por ejemplo, tener muy presentes el *aprendizaje*, el *éxito* y la **felicidad** de sus alumnos
- Involucrar a todas las partes interesadas, ej. profesores, alumnos, padres
- Permitir la participación de un “amigo crítico”
- Estar fundamentada en la confianza, el trabajo en equipo, un “sentimiento de pertenencia” —y **diversión!**— al igual que fundamentada en *objetivos y procedimientos claros*
- Tener como objetivo comprender las **experiencias reales** de la gente, de los intereses creados, etc., al igual que el *analizar los datos*
- Compartirse con los participantes, desde la fijación de objetivos hasta la comunicación de los resultados.

Fuente: Saunders (1999: 417)

Esta diversidad de iniciativas y experiencias vienen a demostrar la pluralidad de enfoques que se están dando en esta modalidad de evaluación. Lesley Saunders (1999) analiza, en un interesante artículo, las diferencias que existen entre el «modelo inglés» de autoevaluación y los principios elaborados, después de dos años de trabajo, por el proyecto CIDREE como marco orientador para llevar a cabo el mismo en ocho países europeos (entre ellos el Reino Unido). Sus reflexiones, a nuestro juicio, trascienden el contexto nacional en el que sitúa su análisis y lo que realmente reflejan

son las supuestos y concepciones que, implícita o explícitamente, se manejan en algunos de los enfoques actualmente existentes. En el Cuadro 5.4 se recogen los principios del CIDREE, los resaltados en *cursiva* se corresponden con el «modelo inglés» y los que aparecen en **negrita** están ausentes en dicho modelo.

A tenor de los principios ausentes (resaltados en **negrita**) en el «modelo inglés», Saunders (1999: 419) afirma que éste se caracteriza por mantener un enfoque instrumentalista, orientado a la acción, racionalista y directivo. Y en este sentido, destaca que los modos de pensar éticos, afectivos, no racionales y democráticos (ausentes en el modelo) son una parte integrante de las actividades de autoevaluación en cuanto son componentes fundamentales del proceso educativo. La dimensión intrínsecamente ética de la práctica educativa y del conocimiento pedagógico; la relevancia de los aspectos emocionales y afectivos para explicar el comportamiento personal y las dinámicas institucionales; la capacidad, por parte de los actores educativos, de interpretar y resignificar los procesos y objetivos burocráticamente establecidos; y la aceptación de la diversidad y pluralidad de intereses y necesidades como un indicador de la vida democrática de las instituciones son, entre otros, argumentos con suficiente peso a la hora de clarificar las posibilidades y limitaciones de los enfoques de autoevaluación.

4. PRESUPUESTOS Y CONDICIONES PARA LA REALIZACIÓN DE LA AUTOEVALUACIÓN

El desarrollo de cualquier institución educativa no puede abordarse sino es sobre la base de una revisión rigurosa y continua de su propia trayectoria. Esto supone plantear que el centro o el foco principal de la evaluación no puede ser otro que el cuestionamiento del quehacer institucional.

Generalmente, la puesta en marcha de un proceso de autoevaluación pone en evidencia las distintas perspectivas desde las que se han elaborado las categorías usualmente consideradas unívocas y que subyacen en la vida institucional. Como nosotros mismos hemos puesto de manifiesto en otro momento (Martín Rodríguez, 1988: 38), las definiciones presentes en el ámbito escolar acerca de temas tales como «lo que es <<un buen alumno>>», <<conocimientos útiles>>, <<disciplina>>, <<participación>> y muchas más, no resultan tan obvias ni son consensuadamente compartidas cuando

son confrontadas entre los distintos profesores. El contraste probablemente sería mucho mayor si se toma en cuenta el contenido que pueden adjudicarles los alumnos o los padres».

Todas y cada una de estas cuestiones, consideradas aislada o conjuntamente, conducen a la problematización del desempeño institucional y consecuentemente a la consideración de que es necesario recoger evidencias que permitan un debate crítico, participativo y plural sobre los mismos. Si ello fuera así estaríamos en los comienzos de lo que será la puesta en práctica de un proceso de evaluación interna. En su diseño y desarrollo convendría tener en cuenta, entre otros, los siguientes elementos: los aspectos organizativos, la conformación del equipo de evaluación, la concreción del objeto de evaluación, la selección de estrategias, técnicas e instrumentos y la metaevaluación.

Aspectos organizativos

La determinación de iniciar un proceso de autoevaluación debe ir acompañada de una planificación mínima de sus aspectos organizativos y metodológicos. El hecho de que la evaluación sea un proceso emergente y difícil de planificar hasta sus últimos extremos no debería plantearse como un obstáculo o contradicción insalvable a la hora de considerar y ponerse de acuerdo sobre ciertos extremos de los que, en última instancia, va a depender su realización.

En relación con la organización, es necesario que se prevean y establezcan, entre otros, aspectos tales como el tiempo, la delimitación de los contenidos y límites de la propia evaluación, la distribución de cometidos y el papel de los distintos participantes, los tipos de informes, etc.

La consideración del tiempo es, pese a lo que pueda parecer a primera vista, una de las primeras variables que necesitan acotarse. Entre otras razones porque hay que tener muy claro que es imposible abordar grandes proyectos en poco tiempo, así como que es contraproducente dilatar temporalmente los mismos. Teniendo en cuenta la secuencia del trabajo de los centros, el tiempo máximo dedicado a la autoevaluación debería ser el de un curso escolar, aunque el proceso pueda repetirse año a año, ya que ésta es la unidad temporal de coincidencia de un mismo equipo docente y sobre la que se puede integrar sus aportaciones en el diseño del siguiente plan anual de actuación. Paralelamente, se ha de decidir también los tiempos

que se van a dedicar a la evaluación, con el compromiso de todos los participantes de que van a respetar escrupulosamente los días y horas fijados para ello, por lo que como actividad prioritaria tiene que anteponerse siempre a cualquier otra urgencia cotidiana que pudiera aparecer.

Los contenidos y límites de la evaluación no supone ni más ni menos que la labor de precisar sus alcances. En este punto, al igual que ocurría con el cuidado de los tiempos, se necesita ser prudentes para no sobrepasar las posibilidades de todo orden (formación, recursos, esfuerzos...) que han de ponerse en juego para llevar adelante el proyecto. Nos referimos concretamente a que de la noche a la mañana no se debería emprender la evaluación de todos los ámbitos del quehacer institucional, si con anterioridad no ha habido ninguna otra experiencia en este sentido y, sobre todo, si no se ha delimitado qué es lo que se va a evaluar. En el fondo lo que se plantea es que es preferible una propuesta modesta que contribuya a la transformación de la institución, antes que el más completo y ambicioso de los proyectos, en donde para empezar la magnitud de la información y evidencias obtenidas pueden llegar a ser de por sí inmanejables.

Igualmente no debería improvisarse la distribución de los cometidos y las funciones de cada uno de los participantes. Desde el comienzo ha de aceptarse democráticamente que todas las sugerencias, ideas y propuestas hechas por cualquiera de los integrantes del equipo de evaluación se considerarán en un plano de igualdad y que las decisiones sólo se tomarán consensuadamente a través del debate y de la discusión grupal. La gestión, no obstante, de todo el proceso requiere de la constitución de un grupo o comisión que sea la que asuma la responsabilidad de conducir, prever y adoptar estrategias para superar los casi seguros altibajos que aparecen en cualquier dinámica colectiva. Lo sustantivo, en estos casos, de acuerdo con los planteamientos de la autoevaluación es que la presencia de un liderazgo compartido, así como el conocimiento del propósito y situación de la institución por la mayor parte de sus miembros, permita encarar los problemas con respuestas que vayan más allá de las incertidumbres y/o soluciones individuales.

Los papeles y funciones que pueden jugar y adoptar los distintos participantes no deben impedir que todos tengan una visión global y unitaria del proceso que se está siguiendo. La formación de diferentes grupos o comisiones, necesarias para avanzar en la evaluación, tiene que hacerse sobre la

base de que su trabajo ha de ser conocido y compartido por los demás miembros de la institución puesto que las decisiones que se tomen, además de negociadas, han de ser asumidas y proyectadas por cada uno en sus actuaciones cotidianas. No estaría de más, profundizando en esta misma línea, que se aprovecharan al máximo las cualificaciones y habilidades propias de cada participante, dado que son muchas las destrezas de gestión, personales y técnicas que han de concurrir para llevar a buen término este empeño. Los no miembros del centro o participantes externos pueden, por ejemplo, ser elementos decisivos para aflorar preguntas o cuestiones no contempladas por los integrantes de la institución, para validar una parte de la información en la medida en que pueden constituirse como uno de los polos de los procedimientos de triangulación y/o para verificar y contrastar las percepciones de los demás participantes.

Tampoco tiene ningún sentido implicarse en un proceso de esta naturaleza si previamente no se detecta que disponemos de mecanismos, recursos o voluntad para vincular la evaluación con las decisiones que al respecto deben asumirse en la mejora y transformación del quehacer institucional. La autoevaluación, tal y como la planteamos aquí, no es una actividad cosmética o que se quede en lo meramente epidérmico, por lo que puede ser enormemente frustrante y peligroso embarcarse en dicho menester sin que sea posible o no se esté decidido a sentar las bases de un modo de proceder alternativo y consecuente con las aportaciones de la evaluación.

La conformación del equipo de evaluación

Un Proyecto de autoevaluación, aun cuando se caracteriza primariamente por ser una propuesta de evaluación interna, puede y debe contemplar la presencia y participación de personas externas con distintos grados de vinculación al centro. En todos los casos es conveniente y tal vez hasta necesario que, además del equipo docente del centro, participen miembros del Consejo del centro, del propio alumnado y de la comunidad educativa, de los Equipos de apoyo a los centros, de algún Departamento universitario, de la Inspección, etc. La incorporación o no de miembros de la comunidad, asesores y/o especialistas en evaluación es una decisión que corresponde al equipo docente del centro. En caso afirmativo, una tarea que no debe demorarse mucho en el tiempo será la de determinar la composición del Equipo de Evaluación, especialmente de las personas externas al centro, ya que es necesario que sean partícipes, desde el primer momento, de todas las decisiones que puedan

tomarse. Conviene sopesar también el número ideal de personas incorporables al equipo para que este no sea mastodóntico e ingobernable.

Todos los componentes del equipo docente serán miembros del Equipo de Evaluación. Inicialmente parece conveniente el que se constituya una Comisión Coordinadora, de un número razonable de miembros, en la que estén representados los distintos grupos de participantes, que será la encargada de gestionar, coordinar y llevar a cabo todos los asuntos propuestos y aprobados por el Equipo de Evaluación. La dinámica del trabajo emprendido será la que informe de la necesidad de crear distintas Comisiones con encargos y tareas específicas.

El Equipo de Evaluación, compuesto por el equipo docente y las personas externas incorporadas al centro, será el encargado de dilucidar y decidir sobre todas las cuestiones generadas por el proceso de evaluación emprendido. Todas las personas integradas en el Equipo de Evaluación deben estar de acuerdo en emprender un Proyecto de Autoevaluación del centro, aun cuando el compromiso y responsabilidad de cada uno de los miembros estará en función del trabajo que razonablemente puedan asumir en cada momento. En cualquier caso, debe quedar muy claro que nadie estará obligado a participar en el proyecto en contra de su voluntad, aunque todos han de comprometerse a facilitar la información necesaria que se les requiera a través del Equipo de Evaluación.

El trabajo del Equipo de Evaluación estará presidido por los principios de funcionamiento democrático y colegiado, de tal forma que sólo le competen a él la toma de decisiones. Las Comisiones formadas por delegación se responsabilizarán de los cometidos para los que específicamente se creen. En todos los casos presentarán sus informes, dictámenes y sugerencias al Equipo de Evaluación, que será quien tenga capacidad de decisión.

La concreción del objeto de evaluación

Uno de los primeros pasos en la elaboración de una propuesta de evaluación es determinar qué es lo que se necesita conocer. La selección de las dimensiones de evaluación comporta el trabajo de ir centrando y acotando los aspectos que son relevantes para tener un conocimiento lo más completo y significativo de aquello que se desea evaluar. Una parte importante de la calidad y utilidad de la propuesta evaluadora se va a jugar precisamente en estos momentos, dado que éstas van a depender en gran medida del rigor y acierto con que se haya realizado dicha selección.

A este respecto convendría recordar que la evaluación no es un acto puramente técnico por ello hay que recalcar que tan o más importante que el **cómo** es el **qué** se evalúa. Y ello supone reconocer que la evaluación es una labor que requiere la clarificación de todo un conjunto de cuestiones sustantivas, relacionadas tanto con la problemática de la propia propuesta evaluadora como con la especificidad de los escenarios y actores implicados en dicho proyecto.

El conocimiento de la especificidad de lo que se quiere evaluar es una condición sine quanon para que se produzca una concreción y selección de los elementos más relevantes del objeto de evaluación, para que los criterios de evaluación sean congruentes con la caracterización y conceptualización de aquello que se va a evaluar y para que el aparato metodológico, como corresponde, esté determinado por el objeto de evaluación. Porque parece evidente que, tal y como afirma Norris (1990: 101) «diferentes tipos de objeto reclaman diferentes tipos de evaluación, y para expresar la relación más formalmente podemos decir que el objeto de evaluación se mantiene en una relación determinante a la metodología de la evaluación».

Aunque la autoevaluación suele concebirse como un planteamiento de evaluación holística, es decir, que abarca a todos los aspectos de la vida institucional: aprendizaje, enseñanza y organización, el proceso puede comenzar estudiando cualquier tema o problema que centre la preocupación de un grupo de profesores:

- El rendimiento de los alumnos
- El trabajo de los profesores
- Aspectos curriculares
- Dimensiones organizativas
- Comunicación y resolución de conflictos
- Las características o clima propio de la escuela
- El aprovechamiento más eficaz de los recursos
- El fomento de las relaciones con los padres y la comunidad, etc.

Selección de estrategias, técnicas e instrumentos

Una vez clarificado qué es lo que nos interesa conocer, estaremos en disposición de seleccionar los dispositivos con los que recoger la información y evidencias apropiadas. En función de la magnitud del plan y de la definición de sus planteamientos metodológicos, podremos decidirnos por

las distintas técnicas e instrumentos con que abordar las diferentes fuentes de información y los momentos o fases de evaluación.

Hay, sin embargo, un conjunto de aspectos relacionados con la metodología de la evaluación que afectan de un modo general a todo el planteamiento que se hace al considerar los distintos ámbitos o áreas, sobre todo porque los caminos de indagación que se abren para tal menester tienen como punto de partida una aproximación holística a la institución objeto de evaluación. En concreto nos referimos a la utilización de un planteamiento metodológico e instrumental plural, en donde no deben descartarse el uso de técnicas diversas, cualitativas y cuantitativas, para la recogida de información. Todo instrumento es perfectamente válido en la medida que nos proporcione una información que contribuya significativamente al conocimiento del hecho o fenómeno estudiado. Existen numerosos ejemplos que evidencian la utilidad de los datos cuantitativos y que ponen de manifiesto que pueden ser los mejores indicadores cualitativos de las políticas y actuaciones realizadas.

En cualquier caso, la utilización de técnicas diversas permite más fácilmente contrastar y validar la información recogida, así como conocer, en suma, los posibles sesgos de las percepciones y valoraciones hechas por las personas implicadas en la evaluación. Los procesos de triangulación, por tanto, deben extenderse no sólo a los posibles evaluadores y a las fuentes de información sino también al aparato técnico e instrumental que se utilice y a los distintos momentos de recogida de la información.

La metaevaluación: evaluación de la evaluación

Una vez agotado el proceso previsto, aunque supongo que se estará de acuerdo en el hecho de que la evaluación es una actividad continua y por tanto inconclusa, conviene proceder a una revisión crítica de lo realizado, tanto por honestidad intelectual con nosotros mismos, como para aprender de la experiencia vivida.

Los aspectos que deben someterse a revisión, son :

- La planificación del propio proceso de evaluación.
- La metodología y técnicas de evaluación utilizados.
- Las relaciones y aportaciones de los participantes.
- Los posibles sesgos y limitaciones del estudio.
- Las dificultades y problemas que han surgido.

- Los campos y temas que deben desarrollarse personal e institucionalmente.
- La posible institucionalización de algunos de los procedimientos.
- La incorporación de las aportaciones relevantes de la evaluación al trabajo del profesorado y a la organización y funcionamiento del centro.

5. BALANCE DE LAS EXPERIENCIAS Y PROPUESTAS DE AUTOEVALUACIÓN INSTITUCIONAL

A tenor de las propuestas y experiencias contempladas, es evidente que existen diferentes formatos y variedades de autoevaluación que, a su vez, son muy distintas entre sí si se tiene en cuenta los propósitos que persiguen, así como las concepciones y metodologías en que se sustentan. Este hecho incontestable es lo que hace que Lesley Saunders (1999: 429), a la hora de valorar para quién y para qué es la autoevaluación, se pregunte asimismo acerca de «*cómo y por qué* la autoevaluación escolar es un ‘animal’ de una especie diferente según el hábitat donde se encuentre».

Tanto en las experiencias de las administraciones como en las de los investigadores se defiende la necesidad de que los resultados escolares no constituyan la preocupación y el eje en torno a los cuales giren los procesos de evaluación. Sin embargo, en uno y otro grupo no son infrecuentes las evaluaciones de producto que se centran especialmente en los resultados, siendo la relación entre los objetivos y el rendimientos de los alumnos lo que determina el nivel de logro y la eficacia de los centros educativos. De esta manera, una institución educativa es mejor o tiene más calidad en la medida que obtiene unos buenos resultados escolares, esto es, cuando los niveles de logro o el rendimiento académico de sus alumnos es alto. Y dado que éstos están asociados a determinadas variables vinculadas con el aula, el curriculum, los profesores y la escuela, lo consecuente tanto discursivamente como en los campos de investigación y evaluación es promover el funcionamiento de los centros desde estos parámetros concretos.

Por lo general, salvo algunas excepciones, las propuestas de autoevaluación promovidas por las administraciones educativas son percibidas, no sin razón, como mecanismos de control de los centros escolares. Lo más frecuente es que la evaluación interna se subordine a las exigencias, requisitos, objetos e indicadores de la evaluación externa. Los propósitos que preten-

den este tipo de autoevaluaciones no pueden disociarse de los de los organismos oficiales que las promueven. Como hemos visto, en unas ocasiones, intentan recoger información de cara a fundamentar y orientar las políticas educativas, en otras, introducir y comprobar el grado de implantación de nuevos modelos de gestión institucional. En cualquier caso, los centros y los actores visualizan estas propuestas como un medio de control de las administraciones educativas y, de alguna manera, sus actitudes de recelo y rechazo ante las mismas se explican porque las perciben como una amenaza de fiscalización de sus quehaceres y actuaciones. Por ello, los datos e informaciones proporcionados por los centros, en bastantes ocasiones, son poco fiables y el valor de los resultados de los estudios es limitado.

Con ser importantes todos estos condicionamientos y limitaciones, quizá uno de los problemas de mayor magnitud y trascendencia de estas propuestas sea la cultura de evaluación que van generando en los centros y profesores. Acostumbrados a ser espectadores antes que actores, así como a filtrar los datos que facilitan, no resulta fácil revertir este estado de cosas haciendo que se conciba y asuma la autoevaluación como un proceso de conocimiento profesional e institucional que contribuya al crecimiento y desarrollo de los centros y profesores.

La autoevaluación auspiciada por determinados investigadores, en especial los vinculados a movimientos académicos como la teoría del cambio, Mejora de la escuela e investigación-acción, propugnan el protagonismo y compromiso de las instituciones y de sus actores para acometer por sí mismos proyectos de innovación y transformación de la enseñanza y organización de los centros. En este sentido, se asume y defiende que son ellos, en tanto que responsables de hacer realidad los objetivos de las políticas reformistas, quienes deben generar formas de trabajo y procesos de indagación para experimentar soluciones y respuestas concretas a los problemas singulares que viven cotidianamente en las aulas y centros.

A diferencia de otras propuestas y orientaciones, en la autoevaluación para la mejora son las propias instituciones y sus autores quienes deciden qué, para qué, cómo y cuándo evaluar, es decir, que se parte de su autonomía y capacidad de decisión para determinar todos estos elementos sustantivos en cualquier proceso de evaluación. Ello supone que no tienen que ajustarse a ningún esquema preestablecido, pudiéndose centrar los estudios en aquellos problemas o temas que sean considerados importantes o

prioritarios por los centros. En todos los casos, la autoevaluación es considerada como una estrategia de cambio en la que, a la vez, no se pierde de vista su contribución al desarrollo institucional y profesional.

Las críticas que recibe esta orientación se relacionan, sobre todo, con la posibilidad de que pueden aparecer conductas endogámicas, más dadas a la justificación que al análisis crítico de sus quehaceres y actuaciones; de que los estudios abordan problemas muy concretos y, por tanto, parciales, perdiendo así la visión global del centro; y que sus realizaciones y experiencias se centran más en las actuaciones y trabajo de los profesores que en los problemas y aprendizajes de los alumnos. Sin desconsiderar ninguna de estas objeciones que, sin duda, se dan y pueden aparecer en los estudios, no se puede tampoco dejar de resaltar las potencialidades que ofrece, si se toman las precauciones adecuadas, para generar espacios de aprendizaje institucional, para potenciar procesos de participación democrática, así como para facilitar la dimensión institucional del trabajo docente.

BIBLIOGRAFÍA

- AA.VV. (1999): «Un proceso de mejora. El modelo canario de autoevaluación», *Cuadernos de Pedagogía*, n. 283, pp.58-62.
- ADELMAN, C. y ALEXANDER, R.J. (1982):. *The SelfEvaluating Institution. Practices & Principles in the Management of Educational Change*. London, Methuen.
- ALCALDE, I. y RÚE, J. (1999): «La autoevaluación. Procedimientos y metodología desarrollados en Cataluña», *Cuadernos de Pedagogía*, n. 283, pp. 50-53.
- ALCALDE I SASTRE, I. (1999): «Llevar el modelo a la práctica. La experiencia del CEIP <<Granullarius>> de Granollers», *Cuadernos de Pedagogía*, n. 283, pp. 53-57.
- ÁLVAREZ MÉNDEZ, J. M. (1997): «La autoevaluación institucional en los centros educativos: una propuesta para la acción», *HEURESIS*, Revista Electrónica de Investigación Curricular y Educativa, <http://www2.uca.es/HEURESIS>.
- BELTRÁN, F. y SAN MARTÍN, A. (1992):. «Autoevaluación Escolar», *Cuadernos de Pedagogía*, n. 204, junio, pp.66-71.
- (1996): «¿Autoevaluación Institucional?», *Cuadernos de Pedagogía*, n. 244, pp. 80-85.

- CERDÁN, J.; SAHUQUILLO, L.; LUJÁN, J. y PUENTE, J. (1992): «La evaluación externa de centros», *Aula*, n. 6, pp. 65-68.
- ELLIOTT, J. (1986): «Autoevaluación, desarrollo profesional y responsabilidad», en M. Galton, y B. Moon: *Cambiar la escuela, cambiar el curriculum*. Barcelona, Martínez Roca, pp. 237-259.
- GATHER THURLER, M. (2002) «L'auto-évaluation de l'établissement scolaire comme moteur du changement», en M. Bois, (dir.) *Les systèmes scolaires et leurs régulations*, Lyon, CRDP, pp. 31-49.
- GREAT BRITAIN. PARLIAMENT. HOUSE OF COMMONS (1997). *Excellence in Schools* (Cm. 3681). London: The Stationery Office.
- HOUSE, E.R. (1992): «Tendencias en evaluación», *Revista de educación*, n. 299, pp. 43-55
- (1998): «Acuerdos institucionales para la evaluación», *Perspectivas*, vol. XXVIII, n. 1, marzo, pp. 123-131.
- INCE (1999): *Evaluación de la calidad en la enseñanza escolar: proyecto piloto europeo*. Madrid, Ministerio de Educación y Cultura, Instituto Nacional de Calidad y Evaluación.
- INNER LONDON EDUCATION AUTHORITY (1977): *Keeping the school under review*. London, ILEA.
- JAMES, M. (1987): «Self-Initiated Self-Evaluation», en P.S Clift, D. L. Nutall y McCormick (eds.): *Studies in School Self-Evaluation*. Lewes, The Falmer Press-Open University Set Book, pp. 172-189.
- KEMMIS, S.: (1986) «Seven principles for Program evaluation in curriculum-development and innovation», en E.R. House, (ed): *New directions in educational evaluation*. Lewes, The Palmer Press, pp. 117-140.
- MACBEATH, J. (1999): *Schools Must Speak for Themselves: The Case for School Self-evaluation*. London, Routledge.
- MACBEATH, J.; SCHRATZ, M., MEURET, D.; y JAKOBSEN, L.B. (2000): *Self-Evaluation in European Schools. A story of Change*. London, Routledge/Falmer.
- MARTÍN RODRÍGUEZ, E. (1988): «Profesión docente y autoevaluación institucional», *Revista de Educación*, n. 285, pp. 33-43.

- MARTÍN RODRÍGUEZ, E. (1989): «Proyección pedagógica de la autoevaluación institucional», en *IV Semana do Maxisterio Lucense*, Escuela Universitaria de Maxisterio, Diputación Provincial de Lugo.
- (1996): «Instrumentos para la autoevaluación institucional de los CEP», en MEC: *Materiales de Apoyo a la Formación*. Madrid, Ministerio de Educación y Ciencia, Documento 4, pp. 102-129.
- (2002): «La autoevaluación institucional», en Martín Rodríguez, E. (coord.): *Desarrollo de las instituciones educativas*. Madrid, UNED, pp. 174-202.
- MEC (1999): *Evaluación de la calidad en la enseñanza escolar. Proyecto piloto europeo*. Madrid, INCE-MEC.
- NEVO, D. (1997): *Evaluación basada en el centro: un diálogo para la mejora educativa*. Bilbao, Mensajero.
- (1998): «La evaluación mediante el diálogo: una contribución posible al perfeccionamiento de la escuela», *Perspectivas*, XXVIII, n. 1, marzo, pp. 87-100.
- NUTTALL, D. (1985): *School self-evaluation. Accountability with a human face?*. York, Longman.
- MCCORNICK, R. y JAMES, M.(1983): *Curriculum Evaluation in Schools*. London, Croom Helm (*Evaluación del curriculum en los centros escolares*. Madrid, Morata, 1996).
- NORRIS, N.(1990): *Understanding Educational Evaluation*. London, Kogan Page Ltd.
- OFFICE FOR STANDARDS IN EDUCATION (1998). *School Evaluation Matters* (Raising Standards Series). London: OFSTED.
- SAN FABIÁN, J. L. y GRANDA, A. (2007). *Guía de Autoevaluación y Mejora*. Principado de Asturias, Dirección General de Políticas Educativas y Ordenación Académica, disponible en Formato pdf en: http://www.educastur.es/media/institucional/calidad/autoevaluacion/manual_autoevaluacion2.pdf (consultado 15 de junio de 2013)
- (2013). *Autoevaluación de Centros Educativos*. Madrid, Editorial Síntesis.
- SANTOS GUERRA, M. A. (1993): *Estrategias para la evaluación interna de los Centros Educativos*. Madrid, MEC (Curso de formación para equipos directivos, Serie Cuadernos, n. 9).

- SAUNDERS, L. (1999): «Who or What is School 'Self-Evaluation for?», *School Effectiveness and School Improvement*, v. 10, n. 4, pp. 414-429.
- SIMONS, H.: (1985) «Against the Rules Procedural Problems in Self-evaluation» en *Curriculum Perspectives*, vol. 5, n. 2. Octubre.
- (1987). *Getting to know schools in a democracy: the politics & process of evaluation*. London, The Falmer Press. (*La evaluación democrática de las instituciones escolares*. Madrid, Edic. Morata, 1999).
- (1990): «Evaluation and the Reform of Schools», en Council of Europe: *The Evaluation of Educational Programmes: Methods, Uses and Benefits*. Amsterdam, Council of Europe.
- (1995a): «La autoevaluación escolar como proceso de desarrollo del profesorado: en apoyo a las escuelas democráticas», en AA.VV.: *Volver a pensar la educación*. Madrid, Edic. Morata, vol. II, pp. 220-242.
- STENHOUSE, L. (1984).: «Evaluating curriculum evaluation», en C. Adelman, (ed.): *The politics and ethics of evaluation*. London, Croom Helm.
- THOMAS, S. (1998): «Medir el valor añadido de la eficacia de las escuelas en el Reino Unido», *Perspectivas*, vol. XXVIII, n. 1, marzo, pp. 101-122.
- TIANA, A. y o (1999): *INAP. Innovative Approaches in School Evaluation*. Final version, march (informe-memoria).

Tema 6

La evaluación de las Instituciones Universitarias

Introducción

1. El papel de los actores en la evaluación educativa
2. La Evaluación Institucional Universitaria en el ámbito europeo
 - 2.1. La evaluación institucional en Gran Bretaña: The Quality Assurance Agency for Higher Education
 - 2.2. La evaluación institucional en Francia: Agence d'évaluation de la recherche et de l'enseignement supérieur
 - 2.3. La evaluación institucional en Italia: Agencia Nazionale di Valutazione del Sistema Educativo e di Ricerca Nucleo di Valutazione
3. Los Programas de Evaluación Institucional Universitarios Españoles

Bibliografía

TEMA 6

LA EVALUACIÓN DE LAS INSTITUCIONES UNIVERSITARIAS

Juan Fernández Sierra.
Susana Fernández Larragueta.
Universidad de Almería

INTRODUCCIÓN

Históricamente la evaluación nace y se desarrolla paralelamente al sistema de producción industrial; el concepto de aumento de la productividad de las fábricas y del control productivo de los obreros tiene su parangón, dentro del sistema educativo, en la medición del aprendizaje del alumnado, facilitado por el avance de los tests de rendimiento. Así pues, el éxito de los programas educativos, la bondad de los centros de enseñanza o la valía de los profesores y profesoras, se expresa en razón del grado de instrucción alcanzado por los estudiantes (Fernández Sierra, 2002). En países como el nuestro, en los que el desarrollo industrial llega con un significativo retraso, la evaluación, más allá de la de los estudiantes, brilla por su ausencia, de manera que los primeros intentos de evaluación de centros educativos tienen lugar, muy tímidamente, en la educación primaria y secundaria, como consecuencia de la puesta en vigor de la Ley General de Educación de 1970, intensificándose tras la LOGSE de 1990. Para la evaluación institucional de las universidades, hemos de esperar hasta mediados de la década de los 90 y la convergencia europea, para dar los primeros pasos.

Actualmente la pertenencia a la Unión Europea, el impulso de la globalización y el proyecto de convergencia de la enseñanza superior, permite que los diferentes países de nuestro entorno compartan sus experiencias en éste campo integrándose en la European Network of Quality Assurance in Higher Education (ENQA). En general, casi todos los países comunitarios tienen una serie de agencias o sistemas, que les permiten valorar la calidad de su educación superior; estando inmersos en procesos similares de búsqueda de propuestas evaluadoras que tienen bastantes elementos comunes, pero también diferentes y ricos matices, muchas veces muy sutiles, pero de gran importancia para sistemas de enseñanza tan diversos como son los europeos. Con la finalidad de ofrecer una panorámica de esta cuestión y dar pie al debate y reflexión, hemos comenzado describiendo tres ejemplos de

procesos institucionales de evaluación llevados a cabo en tres países de la Comunidad Europea —Gran Bretaña, Francia e Italia—; para pasar a analizar la evaluación institucional de la Universidades en nuestro país.

1. EL PAPEL DE LOS ACTORES EN LA EVALUACIÓN EDUCATIVA

Las variables, factores, dimensiones, ámbitos, componentes o aspectos que pueden ser evaluados o formar parte de la evaluación de una institución educativa son múltiples y variados. Su selección y especificación es una de las preocupaciones y tareas prioritarias de muchos teóricos que trabajan en la elaboración y desarrollo de modelos, técnicas e instrumentos de evaluación. En líneas generales, todos hablan de la dicotomía entre evaluaciones globales y evaluaciones de cuestiones específicas, y suelen coincidir en la conveniencia de conjugar ambas en un continuo, bien partiendo de lo general para ir focalizando y estudiando en profundidad las cuestiones más relevantes —efecto zoom que diría Santos Guerra (1990)—, o bien partir de problemas o asuntos concretos para ir ampliando el núcleo de acción a fin de llegar a comprender las complejas interrelaciones de la institución (Beltrán Llavador y San Martín Alonso, 1992). De una u otra forma, los teóricos establecen como función principal de la evaluación de las instituciones educativas, incluso cuando se plantean desde la *accountability*, la de orientar y reorientar la mejora de la calidad de las enseñanzas que se imparten y de la formación de sus estudiantes, así como garantizar la equidad de la educación (Bolívar, 2008).

Al mismo tiempo, es unánimemente aceptado que cualquier reforma y mejora educativa está condenada al fracaso si no contamos con la colaboración y buen hacer del profesorado. Por lo tanto, una cuestión esencial a la hora de planificar una evaluación en este ámbito será decidir cuál es el papel que han de jugar los profesores y profesoras, así como los demás sectores de la comunidad educativa, si se quiere que sea formativa, o sea, si va dirigida a la optimización o a la mejora. En esta dimensión podemos situar el debate sobre las ventajas o desventajas de la autoevaluación, la evaluación interna o las evaluaciones externas, a la hora de plantear la evaluación institucional universitaria.

Una evaluación realizada por expertos externos, por muy sistemática, rigurosa y holística que sea, tiene poco poder optimizador si no es capaz de

implicar e ilusionar al profesorado en las reformas o cambios necesarios para mejorar la institución educativa. Es más, el papel del especialista puede llegar a ser contraproducente si se entiende que es sustitutivo de los profesionales en cuanto a la búsqueda de alternativas a los problemas pedagógicos, potenciando concepciones docentes reduccionistas limitadas a la aplicación de técnicas didácticas de su materia. En esta línea, las evaluaciones externas, destinadas a crear competitividad, desconectadas de los intereses de los actores, llevará a establecer clasificaciones de las universidades, centros o titulaciones que, aparte de su dudosa eficacia para la mejora, tienen el peligro de que esto se haga con información precaria sobre ellos (Fernández Sierra, 2011).

Ahora bien, tampoco podemos rechazar sin más la posibilidad de evaluaciones externas con la finalidad de control, ya que hemos de tener en cuenta que la sociedad tiene el derecho a pedir el rendimiento de cuentas a sus instituciones y a los profesionales que en ellas trabajan. Otra cosa es aceptar cualquier tipo de evaluación sin que estén claras las razones y consecuencias que de ella se puedan derivar, el tipo de datos requeridos y la forma de obtenerlos. Y, sobre todo, sin debatir quiénes tienen el derecho democrático de pedir cuentas, a quiénes y hasta dónde han de rendirlas los mismos que las piden. Con estas condiciones, el experto externo puede ser de gran ayuda.

Respecto a las autoevaluaciones, si bien es cierto como dice Simons (1995, 1999) que tratan de educar nuestros juicios sobre la calidad de las experiencias o actuaciones educativas, y por tanto más adecuadas para el aprendizaje, implicación y compromiso profesional en la mejora de la práctica, es necesario que éstas sean asumidas libremente por la comunidad educativa. Si es impuesta, de forma más o menos sutil, externamente o por unos sectores o estamentos sobre otros, pierde su valor educativo y dinamizador de la vida institucional.

Así pues, ante las evaluaciones externas se suele argumentar que se dan separadas de los intereses y cultura singulares de las instituciones, que son descontextualizadas, que son demasiado instrumentalizadas, que están pensadas para controlar y que, en definitiva, tienen poca capacidad para implicar a los profesionales en la comprensión y solución de los problemas que evidencien. A las evaluaciones internas y las autoevaluaciones se les achaca falta de rigor, poca objetividad, tendencia autojustificadora, infor-

maciones demasiado «nativas», poca proyección y justificación externa y, en resumen, poca crítica y por tanto también poca potencia de cambio o mejora. Las primeras tienen el peligro de distanciarse de la comunidad educativa y las segundas de no conseguir el suficiente rigor y espíritu crítico por la proximidad de los hechos evaluados. Consecuencia de esto es que muchos teóricos han defendido el uso de planteamientos mixtos, es decir, la combinación de evaluaciones externas e internas más o menos simultáneamente. Ésta fue la opción tomada en nuestro país por el Consejo de Coordinación Universitaria en su I Plan Nacional de Calidad de las Universidades (PNECU, 1995-2000) y el II Plan de la Calidad de las Universidades (PCU, 2001), a partir del cual tomaría la responsabilidad evaluadora, la Agencia Nacional de Evaluación de la Calidad de las Universidades y las agencias nacidas al amparo de las Comunidades Autónomas.

No obstante, esta posición tampoco corta el debate, ya que no hay consenso respecto a la bondad de este «maridaje», o al menos se señalan algunos riesgos. Thélot (1993) expresa sus temores en este sentido, referidos al sistema educativo francés, en cuanto a los inconvenientes que tiene plantear evaluaciones externas o internas sin complementarse, pues, piensa, que si se llevan a cabo sólo evaluaciones puramente internas no públicas, incluso si toman la forma de autoevaluación, se corre el riesgo de perder pertinencia y utilidad para aportar información al exterior, pero también para la reflexión interna. La tendencia suele ir en el sentido de que es la evaluación interna la que se subordine a la externa, convirtiéndose en un complemento o estrategia para obtener información de la institución que sería imposible alcanzar a los externos. Otros autores piensan que esta igualación o asimilación de enfoques puede ser perjudicial para la mejora, pudiéndose caer en una monocultura (Saunders, 1999). Tampoco comparte esta combinación de colaboración Simons (1999), ya que opina que la autoevaluación ha de separarse de cualquier exigencia de rendimiento de cuentas, al menos hasta que la cultura evaluadora penetre más en las instituciones educativas, pues los autoinformes pueden convertirse en auditorías y los profesionales no saben muy bien qué se va a hacer con ellos.

En síntesis, la evaluación de una institución no puede plantearse al margen de su transformación, ni ésta es posible si no existe algún tipo de valoración reflexiva compartida por la comunidad educativa. Por lo tanto, lo importante de una propuesta es que sea capaz de implicar a una parte

importante de los diversos sectores y estamentos que componen esa comunidad, evitando o superando la posible resistencia inicial de ciertas personas o grupos. Tanto si la iniciativa es externa como interna, debe convenirse de la utilidad de la evaluación para la mejora de la calidad de la acción educativa. Esto, evidentemente, es más fácil y tiene mayores posibilidades de éxito si la iniciativa es interna y aún mejor si acaba en una autoevaluación asistida por un experto externo hábil y con mentalidad abierta (Elliott, 1990). Que la evaluación de las instituciones de enseñanza es un elemento de calidad, es hoy ampliamente aceptado, pero que la predisposición positiva, colaboración y participación de la comunidad educativa es imprescindible para que la evaluación genere cambio, mejora y calidad educativa, es igualmente evidente.

2. LA EVALUACIÓN INSTITUCIONAL UNIVERSITARIA EN EL ÁMBITO EUROPEO

La coordinación de las universidades europeas para adaptarse a los acuerdos de Bolonia (1999) y la realidad social, cultural, económica y de desarrollo que demanda la Comunidad Europea, ha promovido el surgimiento y desarrollo de diferentes agencias internas de Evaluación Institucional de la Educación Superior. No obstante, en el contexto europeo se establece el *European Network of Quality Assurance in Higher Education* (ENQA), que junto al *European University Association*, la *European Association of Institutions of Higher Education* y *The National Unions of Students in Europe*, asumen las responsabilidades de establecer un conjunto de normas, procedimientos y directrices sobre la garantía de la calidad; sin embargo, el incipiente desarrollo de agencias de evaluación internas, el grado de cooperación internacional y la poca implicación estudiantil, derivó en la creación en 2008 y a petición del Consejo y Parlamento Europeo, del *Registro Europeo de Garantía de Calidad para la Educación Superior* (EQAR) que «permite alcanzar el objetivo de promocionar la cooperación para asegurar la calidad, se apoya en una batería de cinco medidas para alcanzar el objetivo del reconocimiento mutuo en toda Europa de los resultados de las evaluaciones de la calidad universitaria» (Michavila y Zamorano, 2008: 239).

Más allá de los criterios y procedimientos establecidos por la ENQA y de las características idiosincrásicas de las agencias internas de evaluación de

la calidad en la Educación Superior de los diversos países europeos, que pasaremos a analizar en el caso de tres países de la eurozona en los siguientes apartados, observamos una serie de aspectos coincidentes que conviene presentar:

- La evaluación externa, que ha sido realizada por agencias de evaluación institucionalizada, se ha ido progresivamente complementando con una evaluación interna de las propias universidades e incluso por una autoevaluación.
- El objeto de estas evaluaciones varía en función de los países, pero parece generalizada la idea de un triple análisis: docencia, investigación y gestión. No obstante, cada país incidirá más en unos aspectos que en otros, así encontramos a Finlandia o Alemania que centran sus esfuerzos en evaluaciones de la gestión presupuestaria justificando este fin en la concepción de la universidad como una empresa que produce servicios con recursos limitados; y otros como Italia o Francia, quienes centran la atención en la evaluación institucional del proceso de enseñanza y aprendizaje.
- A la hora de la evaluación institucional del proceso de enseñanza y aprendizaje la mayoría de las agencias pone especial atención en la opinión del alumnado asistente, recogiendo habitualmente dicha información a través de cuestionarios, físicos u on-line.
- Prácticamente la totalidad de las agencias de evaluación disponen de un conjunto de guías o catálogos de indicadores que pretenden ser la base orientadora para la evaluación institucional y análisis de la calidad docente, considerando también la investigación y la gestión.
- Las repercusiones de la evaluación institucional dependerán de la toma de posición de los responsables universitarios; no obstante, suelen recorrer un triple camino, en ocasiones divergentes y en otras, convergentes. Así, pueden utilizarse los resultados como un acicate para la reflexión y el cambio; como un proceso farragoso y burocrático carente de sentido transformador; o, no necesariamente desligado de las anteriores, como un proceso de *accountability* encaminado a presentar clasificaciones que faciliten la redistribución de los fondos públicos.

2.1. La evaluación institucional en Gran Bretaña: The Quality Assurance Agency for Higher Education

La evaluación de la enseñanza superior en Gran Bretaña con objeto de mejorar la calidad de la misma, se institucionaliza con el Higher Education Quality Council —HEQC— (Consejo para la Calidad de la Educación Superior) cuyos trabajos, realizados desde 1992 a 1997 se centraban básicamente en la auditoría de la enseñanza superior, proveyendo los recursos económicos en función de cómo las instituciones organizaban y mantenían el nivel y la calidad de sus programas (Brown, 2004).

Actualmente y desde agosto de 1997 el organismo independiente encargado de velar por la calidad universitaria es la Quality Assurance Agency for Higher Education —QAA— (Agencia para la Calidad en la Educación Superior). Este organismo está compuesto por un equipo directivo de catorce miembros (cuatro propuestos por los rectores de las instituciones de educación superior; cuatro propuestos por los apoderados de estos centros; seis directores independientes con experiencia en el campo industrial, comercial, financiero, etc.) y tres observadores (un representante del alumnado; otro de la institución académica de educación superior; un representante del Departamento de Educación del Gobierno). Está encargado de la política de evaluación de la agencia, de supervisar el funcionamiento de los objetivos trazados y de gestionar las finanzas. A este núcleo central se le suman otros miembros como son los Examinadores Externos y los equipos que elaboran los indicadores de las diferentes guías.

Tras una primera etapa en la que la QAA introdujo una serie de cambios respecto al HEQC, como reducir la multiplicidad de evaluaciones que se acometían para con la educación superior, en el 2002 se produjo una reducción en la dotación económica de este organismo (Brown, 2004) que le llevó a concentrar sus esfuerzos en un proceso de auditoría y clasificación externa de las Instituciones de Educación Superior y elaboración de guías de descriptores-indicadores. Así, centra sus funciones en:

- Revisar la calidad de las diferentes instituciones de educación superior.
- Elaborar una infraestructura de guías de referencia (indicadores) que ayuden a definir y explicitar dichos niveles de calidad.

- Garantizar la claridad y transparencia acerca de los propósitos de las diferentes áreas, niveles o/y programas, para conseguir el éxito en cada uno de ellos.
- Presentar una clasificación de la educación superior atendiendo al puntaje obtenido en el proceso de evaluación de la calidad.

Sin embargo, en 2005 se invita a participar a la comunidad educativa en una auditoría institucional del modelo seguido por la QAA, que deriva en 2008 en una fuerte crítica del sistema de clasificación de carreras universitarias británicas llevado a cabo por la QAA. A partir de 2009, el consejo de dirección de la QAA, en un intento por aumentar su reputación, introduce en su agenda: aumentar la transparencia de la agencia, favorecer y fomentar la accesibilidad de la información y establecer un estilo menos formal en la presentación de los informes de evaluación.

Para la QAA es una parte importante de su responsabilidad, presentar con claridad y fiabilidad a la opinión pública cuáles son los niveles de calidad que arrojan sus evaluaciones y los indicadores de referencia sobre los que se basan. Entienden, que sólo pueden mantenerse los niveles de calidad si los implicados en el proceso (empresarios, comunidad universitaria, futuros estudiantes, etc) conocen cuáles son las exigencias básicas para considerar un proceso, un área o un nivel académico de calidad, para así poder actuar en consecuencia. En este sentido, elaboran tres guías básicas con diversos subapartados en los que aparecen dichos criterios de referencia.

La primera de estas guías, denominada The Framework for Higher Education Qualifications in England, Wales and Northern Ireland —FHEQ— (Requisitos para la Educación Superior), tiene como objetivo prioritario presentar a la opinión pública y, fundamentalmente al alumnado, un documento en el que aparezcan los requisitos mínimos para conseguir el éxito en las diferentes etapas de la educación superior. En este sentido, el FHEQ diferencia los cinco niveles en los que se estructura la enseñanza superior en Gran Bretaña, el Certificado (*Certificate*), el Intermedio (*Intermediate*), la Licenciatura (*Honours*), el *Masters* y el Doctorado (*Doctoral*), y en cada uno de ellos se desglosan una serie de ítems en forma de *objetivos de rendimiento* (McGhee, 2003: 12) acerca de qué grado de conocimiento y habilidades deberían ser superados con éxito para, al final del proceso, graduarse en ese nivel.

En las orientaciones de cada etapa puede encontrarse por un lado, la información acerca de las exigencias que necesita un alumno para conseguir el éxito y por otro, las fases por las que debería pasar y las competencias que debería tener un estudiante medio, para demostrar que ha conseguido tal éxito.

Como ejemplificación de los criterios que aparecen en esta guía a continuación presentamos algunos de los requisitos que deben demostrar los alumnos que quieren obtener el nivel *Certificate* (2008: 15):

- Conocimiento de los conceptos y principales argumentos de las áreas de estudio, y capacidad para evaluarlos e interpretarlos en el contexto de dichas áreas.
- Habilidad para presentar, evaluar e interpretar cualitativa y cuantitativamente los datos, así como desarrollar argumentaciones y juicios sólidos, en coherencia con las teorías y conceptos básicos de los temas de estudio.
- Evaluar la pertinencia de los diferentes enfoques para resolver problemas relacionados con sus áreas de estudio y/o trabajo
- Comunicar los resultados de sus estudios de manera formal, precisa y, con una estructura y coherencia argumentativa.
- Tener la capacidad para transferir las habilidades requeridas al desempeño laboral, como responsabilidad profesional.

La segunda de las guías que elabora el QAA es el Subject Benchmark Statements (SBSs), que podríamos traducir como ‘Organización de los puntos de referencia de las asignaturas-materias’. En ella aparecen los requisitos mínimos relacionados con el currículo de las diversas áreas. Describe qué hace a una disciplina coherente y con identidad, y define qué habilidades y destrezas cabe esperar y necesita un graduado para entender y superar la materia. El SBSs se organiza en torno a cinco grandes unidades que aglutinan, a su vez, diferentes ámbitos: los puntos de referencia de las licenciaturas con cuarenta y seis apartados (Honours degree benchmark statements); los del nivel Master con tres materias (Master’s level benchmark statements); los del departamento de salud con dieciséis áreas (NHS/Department of Health benchmark statements); los puntos de referencia en Escocia con cuatro subapartados (Scottish benchmark statements); y los requisitos fundamentales de graduación (Foundation Degree qualification benchmark).

La elaboración y organización interna de las guías queda en manos de los equipos de trabajo de cada área, sin embargo podemos encontrar características comunes entre ellas: habitualmente hay una definición de los principios u objetivos fundamentales de la materia académica sobre la que se elabora la guía; se exponen, a modo de orientación, los aspectos básicos del procedimiento de análisis; se presentan, los puntos de referencia relacionados con la enseñanza, el aprendizaje y la calificación.

Como ejemplificación del SBSs (2007: 4) presentamos una serie de criterios de análisis incluidos dentro del apartado de Licenciatura, en concreto la de Estudios de Educación:

5.6. Reflexión: La carrera universitaria deberá dar la oportunidad a los estudiantes para:

- Reflexionar sobre su propio sistema de valores, desarrollo y práctica.
- Cuestionar los conceptos y teorías que componen sus estudios.
- Interrogarse sobre los supuestos teóricos e investigar.
- Reflexionar sobre su propio desarrollo y actuación.

La última de las guías que presentamos, aunque no por ello menos importante, son The Code of Practice (Códigos de Actuación). En ella aparecen un conjunto de orientaciones generales de carácter práctico que permiten alcanzar y mantener la calidad y el nivel de las universidades y colleges. Está compuesto por diez secciones: 1. Programas de investigación de Posgrado. 2. Grado de colaboración y flexibilidad y distribución del aprendizaje (incluido el e-learnig). 3. Estudiantes con dificultades⁴. 4. Examinadores externos. 5. Quejas de los estudiantes ante las materias académicas. 6. Calificación de los estudiantes. 7. Aprobación, seguimiento y revisión de programas. 8. Orientación profesional, información y guía. 9. Aprendizaje profesional. 10. Contratación y acceso.

Para ejemplificar The Code of Practice hemos elegido una serie de criterios relacionados con la sección «Estudiantes con dificultades» (2010: 18-19).

– En la selección y admisión de estudiantes, la institución tendrá que asegurar la equidad en todas sus actuaciones, para ello y de manera general:

⁴ El término original es *disabled*.

- Asegurar que los criterios y procedimientos para la selección de los estudiantes, se establecen a partir de los requisitos de los programas académicos y profesionales, frente a la desventaja de los candidatos con dificultades.
- Disponer y ofrecer recursos apropiados para utilizar en las entrevistas y en otros métodos de selección.
- Proveer de una guía orientativa y formar a todos los tutores y administrativos envueltos en la selección de estudiantes.
- Disponer de lugares apropiados donde el candidato con dificultad tenga la oportunidad de demostrar su habilidad y conocimiento en el proceso de selección.

Como se puede apreciar, a través de estas guías de orientación, el QAA pretende presentar tanto a la opinión pública (empleadores, familias, estudiantes y Gobierno) como a la comunidad universitaria, toda una serie de criterios que faciliten información sobre los niveles del sistema universitario y sobre los requisitos que van a exigirse (tanto al alumnado como al proceso y a la institución) para mantener los niveles de calidad de la educación superior. Es por ello que estas guías se convierten en documentos informativos sobre las exigencias del sistema universitario.

El segundo gran ámbito de trabajo del QAA es la Auditoría Institucional (*Institutional audit*) de las Universidades y Colegios de Educación Superior de Inglaterra e Irlanda del Norte en colaboración con los Servicios de Calidad (*Quality Assurance*) de las universidades. El proceso de auditoría abarca diferentes objetivos:

- Contribuir, en coordinación con otros mecanismos, a la promoción y mejora de la calidad en la enseñanza y el aprendizaje.
- Asegurar a los estudiantes, empleadores y otros interesados, información pública, fiable y significativa.
- Asegurar un rápido proceso de intervención para la mejora, en caso de que la calidad de los programas de educación superior o los niveles de éxito sean pobres o tengan deficiencias serias.
- Proveer de un informe de *accountability* que facilite la distribución de los fondos públicos.

El proceso de auditoría intenta combinar dos niveles de análisis, el examen interno del Quality Assurance sobre el sistema de la institución y las evaluaciones referentes a cómo influye dicho sistema en el nivel disciplinar y en el nivel de programa⁵. Su plan de actuación se centra en analizar la institución a partir de los criterios establecidos en las diferentes guías del QAA, por ello, dependiendo de los indicadores de análisis que se esté trabajando en cada momento variará la estrategia de recogida de información (fundamentalmente la entrevista formal) y los informantes.

A partir de los resultados obtenidos en la Auditoría Institucional se establecen las carreras universitarias que pueden incluirse en el grupo de las Instituciones de Alta Calidad (HEIs). Además, se presenta un informe-clasificación de *accountability* para ser utilizado en la distribución entre las instituciones de Educación Superior de los fondos públicos.

A través de la evaluación institucional el QAA elabora una especie de catálogo donde se marcan los niveles de calidad de las diferentes universidades así, este organismo pretende relacionar a las Universidades y Colegios de Educación Superior suscritos a la asociación con los principales (posibles o *in facto*) proveedores-valedores de la educación superior.

2.2. La evaluación institucional en Francia: Agence d'évaluation de la recherche et de l'enseignement supérieur

Para entender la realidad de la educación superior francesa, debemos conocer dos aspectos fundamentales que la determinan y que pueden ejercer gran presión sobre la calidad del servicio que ofrecen. El primer elemento sería que el control administrativo de la educación superior está en manos del Ministerio de Educación, que es el que la financia y tiene la potestad de nombrar y promocionar al profesorado y autorizar las carreras universitarias nacionales; el segundo aspecto tendría que ver con la relación, no siempre coincidente, entre los intereses de los agentes cofinanciadores de la educación superior (estudiantes-familias, empresarios, etc) y la concepción de ésta como un servicio público para los ciudadanos.

Esta situación llevó a la Presidencia francesa a crear el Comité National d' Evaluation —CNE— (Comité Nacional de Evaluación) en 1984 como un

⁵ En este contexto, una disciplina está definida como un área de estudio académico y el programa es definido como el conjunto de módulos, investigaciones y seminarios, que favorecen el aprendizaje.

cuerpo independiente del Ministerio de Educación, para que se hiciera responsable de la evaluación de las instituciones públicas de carácter científico, cultural y profesional: universidades, *ecolés*⁶ y grandes instituciones bajo tutela del Ministerio de Educación Superior. Posteriormente, en 2007, se crea la Agence d'Évaluation de la Recherche et de l'Enseignement Supérieur —AERES— (Agencia de Evaluación de la Investigación y de la Enseñanza Superior) en la que se integra el CNE.

Actualmente, AERES está formado por veinticinco miembros franceses y extranjeros, elegidos por un periodo de cuatro años renovable en una ocasión: diecinueve representando a la comunidad académica y científica de la Educación Superior; cuatro a propuesta del Consejo Económico y Social; uno del Consejo de Estado; y uno del Tribunal de Cuentas.

Los principios rectores que guían la política de la agencia y de los procesos de evaluación son tres: independencia administrativa que garantiza el pleno funcionamiento sin ingerencias por parte de autoridades gubernamentales, evaluados u otros; transparencia en la gestión, en los procesos metodológicos, en la organización interna y en los códigos éticos, considerando los principios nacionales y europeos, para ello, se hacen públicos todos los informes de evaluación; e imparcialidad y equidad en el desarrollo de las evaluaciones a través de diferentes comités de expertos, los cuales son seleccionados atendiendo a criterios como la colegialidad, las habilidades complementarias, la multidisciplinariedad y la imparcialidad en la evaluación.

La agencia se estructura en torno a tres grandes programas: *Evaluation des Établissements* (evaluación de las instituciones), *Evaluation des unités de recherche* (evaluación de las unidades de investigación), y *Evaluation des formations & diplômes* (evaluación de la formación y los títulos). Cada una de estas secciones está a cargo de un Director elegido por el Consejo del AERES a propuesta del presidente de la agencia. Entre sus funciones está la de organizar los procesos de evaluación y, en su caso, la toma de decisiones que conlleve.

La evaluación institucional llevada a cabo en Francia siempre contempla una primera fase interna, en la que se realiza una autoevaluación insti-

⁶ Se trata de instituciones muy selectas, cuyo objetivo es la formación de los ingenieros, profesores, economistas, etc. de muy alto nivel.

tucional; en ésta, el centro de investigación o de enseñanza superior, elabora un informe a partir de los datos obtenidos con los diferentes instrumentos de recogida de información, destacando los cuestionarios al alumnado, exponiendo sus características básicas y las expectativas a las que aspiran. Y una fase externa, llevada a cabo por evaluadores externos que elaboran un informe confidencial. El informe final y de carácter público elaborado por AERES tiene en cuenta la información presentada tanto en los informes internos como en el informe de los expertos. Por otro lado, aunque la metodología utilizada en el proceso es tanto cuantitativa como cualitativa, desde AERES se remarca que la evaluación institucional de la educación superior es principalmente cualitativa, puesto que no sólo se tienen en cuenta los descriptores que aparecen en las guías de referencia, sino que se toma en consideración el contexto, la situación específica, la evolución y los objetivos particulares de cada institución.

A lo largo del proceso de evaluación institucional existe una coordinación y comunicación constante entre los evaluadores y los miembros de la comunidad universitaria a fin de: reflexionar sobre la metodología de evaluación y la guía de evaluación interna, consensuar las visitas in situ de los evaluadores externos, y negociar el informe final de la evaluación, puesto que los miembros de la institución evaluadora son también responsables de la validación de los datos que aparecen en el mismo.

La evaluación llevada a cabo por el AERES pretende, más que la sanción o la rendición de cuentas (no tiene por objeto calificar individuos, habilitar programas o distribuir las subvenciones del Estado), el análisis de la realidad (académica, de gestión y de relación con el contexto social) de la Educación Superior. Los informes de evaluación tienen por objeto enfrentar a las instituciones con su realidad, haciendo hincapié en si cumplen su misión como servicio público y alcanzan sus objetivos. De esta manera, el informe se constituye como un instrumento estratégico del que disponen las instituciones para mejorar la calidad del proceso de enseñanza y aprendizaje, investigación y gestión. Además, sirve como información a la opinión pública, fundamentalmente a los usuarios, al Estado y a otras instituciones-empresas financiadoras.

2.3. La evaluación institucional en Italia: Agencia Nazionale di Valutazione del Sistema Educativo e della Ricerca e Nucleo di Valutazione

En la concepción de la evaluación institucional de las universidades italianas, debemos tener en cuenta las características contextuales de la realidad organizativa y administrativa del Estado italiano. En su Constitución se delega a las Regiones las competencias

«[en materia] de enseñanza, salvo la autonomía de los establecimientos docentes y con excepción de la instrucción y formación profesional, las profesiones, la investigación científica y tecnológica y el apoyo a la innovación para los sectores productivos» (GU 298 del 27 de diciembre de 1947: art. 117).

Por ello, no se puede establecer un sistema único y directivo de evaluación desde el Estado general, debido a la independencia en materia educativa e investigadora de las Regiones y a la autonomía de gestión que tienen las universidades.

El sistema de análisis de la calidad de la educación superior debe adecuarse a esta situación promoviendo un doble plano de evaluación institucional a través de dos sistemas paralelos que se complementan mutuamente, así combinan: el plano externo, encargado desde su origen en 1993 al Comitato Nazionale per la Valutazione del Sistema Universitario (CNVSU) y, desde 2011, a la Agenzia Nazionale di Valutazione del Sistema Educativo e della Ricerca (ANVUR); y el plano interno, que queda en manos del Núcleo di Valutazione.

La evaluación externa de las universidades tiene carácter obligatorio, puesto que de ello dependen las subvenciones económicas estatales destinadas a cada universidad, y es llevado a cabo por un órgano institucional dependiente del Ministerio de Universidad, de Investigación Científica y Tecnológica; los cinco miembros del consejo de dirección, son designados desde el Ministerio y su mandato tiene una periodicidad de cuatro años sin renovación. No obstante, que el ANVUR forme parte del organigrama del Ministerio no significa que no se defina a sí mismo como un órgano independiente que interactúa autónomamente con la Universidad y el Ministerio.

La actividad de la *Agenzia Nazionale di Valutazione del Sistema Educativo e della Ricerca* (2013), de manera general, se centra en:

- Evaluación externa de la calidad de la actividad de las universidades y de los entes públicos y privados de investigación.
- Revisión de las evaluaciones que cada universidad lleva a cabo internamente.
- Valoración de la eficacia y eficiencia de los programas públicos de financiación y apoyo a la investigación.

Estas actividades se materializan prioritariamente en los siguientes programas:

- Programa AVA: encargado de la evaluación periódica de los planes de estudio instaurados en las diversas universidades, para comprobar su calidad y eficiencia, con la finalidad de calcular las subvenciones que éstas recibirán del gobierno.
- El VQR 2004-2010⁷: proyecto encargado de evaluar el nivel científico de las investigaciones llevadas a cabo en las universidad y centros de investigación privados, a fin de crear un ranking de centros de investigación.
- La comisión de los Dottorato di Ricerca: éstos se ocupan de evaluar periódicamente los estándares de calidad de las instituciones que ofertan doctorados.

Esta evaluación externa se complementa con el Nucleo di Valutazione que se establece como un organismo interno de las universidades⁸ encargado de la realización de la evaluación institucional interna.

El comité de Nucleo di Valutazione elabora un conjunto de ámbitos de evaluación interna que eleva al Comitato Nazionale per la Valutazione y que, generalmente y a colación con los intereses de la Conferencia de Rectores de las Universidades Italianas, está encaminado a promover la evaluación desde una triple óptica: didáctica, investigación y gestión administrativa (*didattica, ricerca y gestione amministrativa*).

Como ejemplo de Nucleo di Valutazione explicamos el del Ateneo di Bologna. Éste es un organismo independiente previsto en los Estatutos de

⁷ La fecha de cierre de la convocatoria fue diciembre 2012, y en este 2013 se prevé que salga el ranking

⁸ Es dependiente de cada universidad y tanto la organización como la elección de miembros dependerá de los Estatutos de cada una de ellas.

la Universidad de Bolonia y se instituye como un organismo de evaluación encaminado a, previa definición de los parámetros de referencia establecidos por el Gobierno de la Universidad y el Estamento Nacional, la evaluación institucional del sistema universitario, al análisis interno de la investigación y de la enseñanza, así como de otras funciones establecidas en la legislación. La constitución del Nucleo di Valutazione se produce en abril de 1994 y la selección de sus cinco miembros, con una periodicidad bianual, está en manos del Consejo de Administración una vez escuchado el Senado Académico.

Tomando como referencia los objetivos de la evaluación institucional relacionados con la enseñanza en las universidades italianas y como en la mayoría de los Nucleo di Valutazione, el Ateneo de Bolonia lleva desde el año 1997 elaborando *informes de opinión* del alumnado asistente sobre la enseñanza en las diferentes carreras universitarias y facultades que componen la universidad; a partir de la información obtenida, se prepara un informe piramidal de evaluación institucional de la enseñanza en el Ateneo que contenga un nivel básico relacionado con las asignaturas específicas, uno más general, sobre la enseñanza en las diferentes carreras y facultades, y otro más genérico sobre la enseñanza en la universidad.

La evaluación institucional llevada a cabo por el Comitato Nazionale per la Valutazione y por los Nucleo di Valutazione no sólo tienen por objeto informar a la comunidad universitaria y a la sociedad, sobre cuáles son las características del sistema universitario y la calidad de la educación superior o qué opinión tiene el alumnado de la enseñanza de los Ateneos para mejorar la calidad de la educación superior, sino que también provee de un informe de *accountability* que facilite el proceso de redistribución de los recursos económicos.

3. LOS PROGRAMAS DE EVALUACIÓN INSTITUCIONAL UNIVERSITARIOS ESPAÑOLES

Si la experiencia en evaluación de proyectos e instituciones educativas es reciente en nuestro país, en el ámbito universitario aún es más cercana en el tiempo. La primera actuación sistematizada la constituye el Programa de Evaluación Institucional de la Calidad de las Universidades Españolas que fue aprobado por el Consejo de Universidades a principios del curso 1995-96. Como experiencia previa propia, apenas si puede añadirse algo más del Programa Experimental de Evaluación de la Calidad del Sistema

Universitario⁹ aprobado por el Consejo en septiembre de 1992 y llevado a cabo durante el curso 1993-94, con el objetivo principal de poner a prueba algunas de las metodologías que se habían utilizado internacionalmente para este menester. Se consideró que, como punto de arranque, la experiencia había sido útil, pero que habrían de simplificarse los procesos, puesto que no habían cubierto todos los objetivos, especialmente en lo relacionado con la evaluación de la gestión.

«El nivel de cumplimiento de objetivos y tareas previstas en el programa se estimó, al finalizar el mismo, en un 80%. El déficit afectó principalmente a las evaluaciones de nivel III en el campo de la investigación y de la gestión, por no haber podido completarse la fase de evaluación externa» (Consejo de Universidades, 1995a: 18)

Esta primera experiencia se completó y cumplimentó con la participación en el Proyecto Piloto Europeo, que se desarrolló durante el curso 1994-95¹⁰, cuyos objetivos

«consistían en aumentar el grado de conciencia acerca de la necesidad de evaluar la enseñanza superior; enriquecer los procedimientos nacionales de evaluación existentes; promover los intercambios de experiencias y dotar a la evaluación de una dimensión europea» (Consejo de Universidades, 1995b: 1).

La evaluación se centró en la enseñanza y el aprendizaje, teniendo en cuenta las actividades de investigación en la medida en que afectaban al proceso educativo.

Tanto una como otra experiencia trataban de combinar lo que hoy es la línea prioritaria en la evaluación universitaria de los países europeos, la complementación entre la evaluación interna y la externa. Así lo recomendaba también el llamado informe Bricall (2000), o Informe Universidad 2000 en su capítulo VII de la Calidad y Acreditación. No obstante, esta coincidencia en el planteamiento general, no determina que todos los enfoques entiendan de la misma manera

⁹ En realidad, como se recoge en el Informe del Proyecto Piloto Europeo (Consejo de Universidades, 1995b), eran en ese momento escasas las experiencias de evaluación sistemática llevadas a cabo en la enseñanza superior; sólo Francia, Reino Unido, Dinamarca y Países Bajos contaban con un sistema total de evaluación.

¹⁰ De España participaron la Universidad Politécnica de Cataluña y la Universidad del País Vasco en la titulación de Ingeniería y Telecomunicaciones, y las de Granada y Carlos III, en el título de Documentación. Las áreas temáticas acordadas para este proyecto piloto eran Ciencias de la Ingeniería y Ciencias de la Comunicación/Información o Arte/Diseño.

el equilibrio entre evaluación interna y externa, pudiendo primar una u otra en cuanto se inclinen más por la mejora o por el control o establecimiento de estándares o ranking con intenciones más o menos comparativas.

Vistas estas experiencias, el Consejo de Universidades propuso al Gobierno y puso en marcha el Plan Nacional de Evaluación de la Calidad de las Universidades Españolas. La primera convocatoria se publica en el BOE nº 15 de 28/02/1996 y se adjudica el 30 de octubre del mismo año, así que el curso 1996/97 supone el primero en el que todas las universidades españolas tienen la oportunidad de participar en un plan general de evaluación y entrar en una nueva dinámica, hasta ese momento poco explorada, de Evaluación Institucional.

Este Plan Nacional, recoge bastante fielmente las líneas y conclusiones de los planes anteriormente mencionados, para diseñar y llevar a cabo el trabajo propuesto, tanto en cuanto a los propósitos como a la metodología.

Los objetivos pretendidos eran, según el decreto que aprobó el plan (BOE, nº 294 de 9/12/1995), promover la evaluación institucional de la calidad de las universidades, elaborar metodologías de evaluación integradas en la práctica vigente en la Unión Europea, y proporcionar información objetiva que pudiera servir de base para la adopción de decisiones.

Los procedimientos, igualmente usuales en el resto de Europa, son el autoinforme, la visita de expertos externos y la emisión del informe final. Cada universidad organiza su propio equipo de evaluación interna, que será el encargado de recoger, analizar los datos y elaborar el autoinforme, siguiendo la guía de evaluación (protocolos) facilitada por el Consejo de Universidades.

Una vez elaborado dicho autoinforme, se remite al comité de expertos externos que, a partir de su estudio y la guía de evaluación externa que también le facilita el Consejo de Universidades, prepara una visita de tres días en los que entrevistará y se reunirá, además de con el Comité Interno de Evaluación, con estudiantes de los diversos niveles, profesorado, personal de administración y servicios, responsables de los departamentos, centros o facultades, egresados, con cada sector independientemente, más una audiencia pública. Así mismo visitará las instalaciones y dependencias administrativas, docentes y de investigación utilizadas por la titulación evaluada. Dicha visita ha de terminar con un primer adelanto de informe oral y en el plazo de un mes, el Presidente del Comité Externo, ha de remitir al Comité Interno y al

Consejo de Universidades su informe correspondiente.

El análisis de los pros y los contras del proceso y resultados del Plan Nacional de Evaluación de las Universidades también cuenta con abundante información procedente, en gran parte, de los responsables del Comité Técnico y otras personas que de alguna forma han participado directamente en dichos procesos o en otros similares, en las universidades españolas (Apodaca y Grao, 1997; Michavila, 1998; *Revista de Educación*, n. 315 y n. 316, 1998)

En abril del 2001 se publicó el decreto que ponía en marcha el II Plan de la Calidad de las Universidades (RD 408/2001), que en esencia sigue las líneas generales del plan anterior; aunque con una mayor explicitación y predeterminación de los elementos a tener en cuenta en la evaluación —concretados a través de las guías internas y externas— que si por un lado facilitan los procesos, por otro los controla y los hace cada vez más directivos, disminuyendo la posibilidad de análisis e interpretaciones contextualizadas y diversas de los complejos sistemas de relación y actuación consustanciales a las instituciones universitarias. Este II Plan, que nace con una periodicidad de seis años (habría de prolongarse hasta 2006), es a su vez derogado en el 2003 (RD 1391/2003), frenándose el modelo de evaluación institucional puesto en escena hasta ese momento, preparándose así el cambio que introduce la LOU (2001).

En línea con lo marcado por esta última ley, durante el curso 2003/04, se pone en marcha un nuevo modelo experimental bajo la responsabilidad de la Agencia Nacional de Evaluación de la Calidad y Acreditación, que como el propio nombre de este organismo indica, la nueva filosofía se basa en la unión de la evaluación a la acreditación. Desde entonces, es la ANECA, quien, a través de diversos programas que han ido apareciendo, se ha encargado de desarrollar la evaluación y acreditación en varios niveles:

- Programa Verifica: que establece un marco normativo para la ordenación y verificación de los títulos oficiales.
- La evaluación institucional de enseñanza oficiales, que, como mencionan Michavila y Zamorano (2008), favorece la evaluación y mejora de la calidad de las enseñanzas oficiales a través de un procedimiento de doble vía: autoevaluación y evaluación externa, que finaliza en un informe final.
- Programa Docentia: evalúa la cualificación y competencia docente del

profesorado universitario.

- Programa Academia: encaminado a la acreditación del profesorado universitario.
- Programa Audit: encaminado a la organización, desarrollo y potenciación de Sistemas de Garantía Interna de Calidad, en el seno de las universidades.

Esos planes iniciales y los posteriores Programas de la ANECA han supuesto un gran impulso a la difusión y concienciación de la necesidad de evaluar la actuación y el servicio prestado por las instituciones de enseñanza superior a sus estudiantes y a la sociedad en general, dentro y fuera de la Universidad, aunque presentan algunas limitaciones o puntos débiles en los que es necesario incidir

Los diferentes planes y programas, si bien han tenido la virtud de, en cierto modo, conformar a todos o, al menos, no provocar rechazos frontales, ha sido a costa de la indefinición de sus propósitos. O sea, se ha situado en una dimensión de mejora interna con una metodología no totalmente apta para una comprensión profunda que entrara en el debate y la reorientación, y un diseño que parecía podría dar información externa para de alguna manera transmitir un ranking «sutil» que supusiera una presión para las universidades, pensando que eso les animaría a superar algunas de sus limitaciones o «vicios»; tendencia que se ha ido manifestando paulatinamente a lo largo de estos años.

En esa línea es verdad que, como alegan muchos de los miembros de los Comités Técnicos, este proceso ha provocado en las universidades (tal vez haya demostrado el déficit de partida) la necesidad de adecuar y actualizar sus sistemas de información y disposición de datos cuantitativos. No sabemos si para esto era necesaria la evaluación, pero parece que ésta de alguna forma ha avivado la urgencia. También consideran positivo que se haya forzado de alguna manera a marcarse objetivos que no solían estar explicitados en las titulaciones ni otras acciones; no obstante eso puede tener un doble efecto: si obliga a la discusión de perfiles profesionales y de los propósitos de las titulaciones, siempre resultará un avance, pero si la preocupación por buscar resultados «cuantificables» deriva en la definición de objetivos simples para poder medirlos con facilidad, puede distorsionar gravemente el debate y las alternativas.

En este sentido, hemos de señalar que, como advirtiera Elliott (1986) analizando algunos sistemas de autoevaluación ingleses, las listas o guías facilitadas a los comités internos, si bien les ayudan a realizar una labor a la que no están acostumbrados, no dejan de representar una forma de selección externa de lo que es lo importante y de obviar lo que no lo es, reconvirtiéndose en especies de recordatorios de lo que hay que hacer o formas de comprobar fácil y rápidamente las actuaciones que se llevan a cabo.

Por otro lado, los autoinformes, en general, tienden a dar descripciones en muchas ocasiones excesivamente suaves, por tanto superficiales, de problemas profundos, sea porque se perciben como «normales» algunas situaciones que se llevan viviendo muchos años o por mantener el equilibrio político en los centros. Aunque esto era de esperar puesto que ya pasó en el Proyecto Piloto Europeo.

Unido a lo anterior, es frecuente que las conclusiones y/o las enumeraciones de los puntos fuertes y débiles recogidos en dichos informes, más que de los datos empíricos, se derivan del conocimiento previo e interpretación más o menos subjetiva o interesada que los miembros del Comité Interno comparten sobre las situaciones e instituciones en las que desarrollan su trabajo. Así mismo, las alternativas o propuestas de mejora que realizan, se convierten en gran parte en una mezcla de sugerencias y reivindicaciones, normalmente dirigidas a estamentos superiores, políticos, académicos o sociales, entrando poco en las actuaciones que le son más cercanas y propias al profesorado y departamentos, independientemente de los datos y del valor para la mejora real que supongan.

Por su parte, los informes de evaluación externa han sufrido el mismo contagio. Les resulta muy difícil a los expertos externos deshacerse de su condición de «nativos» de las mismas áreas y similares vivencias en las actuaciones que se someten a evaluación a través de las titulaciones, centros, etc. Igual suele suceder ante problemas generalizados, reales o virtuales (como niveles de conocimientos y compromiso de los estudiantes, importancia de ciertas materias básicas en los planes de estudio, creencias docentes, apegos a rutinas didácticas obsoletas, etc.), que de alguna forma comparten, viven e interpretan desde un *profesionalismo* gremialmente solidario.

Más *debilidad* puede tener aún la casi nula focalización real hacia un aspecto esencial de la evaluación institucional como es la gestión, ya que

ésta no ha sido analizada ni superficialmente. Aquí se mezclan cuestiones sindicales, políticas, de representación, etc. que son complejas y prietas de intereses diversos difíciles de desentrañar, pero de gran importancia para comprender y tomar decisiones en pos de la mejora. La micropolítica (Ball, 1994) queda fuera del ojo analítico, incluso la *mesopolítica* universitaria que serían las que nos darían las claves de lo que pasa con las titulaciones (planes de estudio y posición en ellos de las diferentes áreas y materias, relaciones de fuerzas, perfiles profesionales explícitos o tácitos, sistemas de asignación de profesorado, etc.).

Para concluir decir que quizás lo más destacado de estos planes o programas de evaluación institucional universitaria, ha sido la posibilidad que ha dado a grupos de profesores, personal de servicios y estudiantes, especialmente a los que componen los comités internos, de tratar de recoger datos, organizarlos y elaborar un informe, con la posibilidad de debate e intercambio de información y opiniones, así como la discusión razonada de alternativas viables y contextualizadas. A ello ha de sumársele la oportunidad de contraste e intercambios de puntos de vista e interpretaciones de los hechos y fenómenos académicos con los comités externos que, a pesar de algunas limitaciones, representan una visión desde fuera de la problemática singular de cada lugar.

BIBLIOGRAFÍA

- AGENCE D' EVALUATION DE LA RECHERCHE ET DE L' ENSEIGNEMENT SUPÉRIEUR (2013) [en línea] [Fecha de consulta: 23 de mayo de 2013]. Disponible en: <http://www.aeres-evaluation.fr/>
- AGENZIA NAZIONALE DI VALUTAZIONE DEL SISTEMA EDUCATIVO E DELLA RICERCA <http://www.anvur.org/index.php?lang=it>
- APODACA, P. y GRAO, J. (coord.) (1997): *Calidad en la Universidad: Orientación y Evaluación*. Barcelona, Laertes.
- BALL, S.J. (1994): *La Micropolítica de la escuela: hacia una teoría de la organización escolar*. Barcelona, Paidós.
- BELTRÁN LLAVADOR, F. y SAN MARTÍN ALONSO, A. (1992): «Hacer posible la democracia organizativa». *Cuadernos de Pedagogía*, n. 204, pp. 66-71.

- BOLÍVAR, A. (2008): «Evaluación de la práctica docente. Una revisión desde España». en *Revista Iberoamericana de Evaluación Educativa*, vol. 1 n. 2: 56-74
- BRICALL y col. (2000): *Informe Universidad 2000*. Tomo I y II. Policopiado.
- BROWN, R. (2004): *Quality assurance in Higher Education. The UK experience since 1992*. Londres, RoutledgeFalmer.
- CASANOVA, M.A. (1992): *La evaluación, garantía de calidad para el Centro Educativo*. Zaragoza, Edelvives.
- CONSEJO DE UNIVERSIDADES (1995a): *Programa de Evaluación Institucional de la Calidad de las Universidades*. Madrid, Consejo de Universidades, Secretaría General.
- (1995b): *Informe Europeo. Proyecto Piloto Europeo para Evaluar la Calidad en el Enseñanza Superior*. Policopiado.
- (1997): *Guía de los evaluadores externos en el Plan Nacional de Evaluación de la Calidad de las Universidades*. Madrid, Consejo de Universidades, Secretaría General.
- ELLIOTT, J. (1986): «Autoevaluación, desarrollo profesional y responsabilidad», en M. Galton, y B. Moon, (eds.). *Cambiar la escuela, cambiar el currículum*. Barcelona, Martínez Roca.
- (1990): *La investigación-acción en educación*. Madrid, Morata.
- FERNÁNDEZ SIERRA, J. (coord.) (2002): *Evaluación del rendimiento, evaluación del aprendizaje*. Madrid, Akal
- (2011): *Formar para la economía del conocimiento vs educar para la sociedad del conocimiento: una visión desde la pedagogía*. Málaga, Ediciones Aljibe
- FREIRE, P. (2001): *Pedagogía de la Indignación*. Madrid, Morata.
- MARTÍNEZ BONAFÉ, A. (1992): «De la evaluación del alumnado a la evaluación del centro», *Aula de Innovación Educativa*, n. 6, pp. 52-59.
- MCGHEE, P. (2003): *The Academic Quality Handbook: Enhancing Higher Education in Universities and Further Education Colleges*. Glasgow: Routledge
- MICHAVILA, F. (ed.) (1998): *Experiencias y consecuencias de la evaluación universitaria*. Madrid: Fundación Universidad-Empresa.

- MICHAVIDA, F. y ZAMORANO, S. (2008): «Panorama de los sistemas de garantía de calidad en Europa: una visión transnacional de la acreditación». En *Revista de Educación*, núm. extraordinario 2008: 235-263.
- MINISTERIO DE EDUCACIÓN Y CULTURA (1998): «La evaluación de la universidad española», *Revista de Educación*, n. 315.
- (1998): «Sistemas educativos más allá de la OCDE», *Revista de Educación*, n. 316.
- (2008): «Tiempo de cambio universitario en Europa», *Revista de Educación*, núm. extraordinario.
- PÉREZ GÓMEZ, A. I. (1988): «Autonomía y formación para la diversidad», *Cuadernos de Pedagogía*, 161, 8-11.
- QUALITY ASSURANCE AGENCY FOR HIGHER EDUCATION (2007): *Subject Benchmark Statements (SBSs)*. [en línea] [Fecha de consulta: 23 de mayo de 2013]. Disponible en: <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Pages/Subject-benchmark-statement-Education-studies.aspx>
- (2008): *The Framework for Higher Education Qualifications in England, Wales and Northern Ireland (FHEQ)*. [en línea] [Fecha de consulta: 23 de mayo de 2013]. Disponible en: <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Pages/default.aspx>
- (2010): *The Code of Practice*. [en línea] [Fecha de consulta: 23 de mayo de 2013]. Disponible en: <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Pages/Code-of-practice-Section-3.aspx>
- SANTOS GUERRA, M.A. (1990): *Hacer visible lo cotidiano: teoría práctica de la evaluación cualitativa de los centros escolares*. Madrid, Akal.
- SAUNDERS, (1999): «Work in value-added analysis of schools performance at the National Foundation for Educational Research (NFER): El análisis del valor añadido en la evaluación de centros», en *Infancia y Aprendizaje*, n° 85 : 19-32
- SIMONS, H. (1995): «La autoevaluación escolar como proceso de desarrollo del profesorado: en apoyo a las escuelas democráticas», en AA.VV. *Volver a pensar la educación*, vol. II. Madrid, Morata.
- (1999): *Evaluación Democrática de Instituciones Escolares*. Madrid, Morata.
- THÉLOT, C. (1993): *L'évaluation du système éducatif*. Paris, Nathan.

Tema 7

El profesorado como objeto de evaluación

Introducción

1. Acerca de la evaluación del profesorado
2. Tendencias y características de la evaluación del profesorado
3. Los propósitos de la evaluación del profesorado
4. Perspectivas de evaluación docente
5. A modo de conclusiones

Bibliografía

TEMA 7

EL PROFESORADO COMO OBJETO DE EVALUACIÓN

Eustaquio Martín Rodríguez.

UNED

INTRODUCCIÓN

El profesorado al igual que los centros educativos es un ámbito de evaluación muy reciente en la historia de este campo de estudio e investigación. Una gran parte de los esfuerzos, en la mayoría de las reformas educativas llevadas a cabo hasta los años setenta, se canalizaron casi siempre a través de cambios centrados en el impulso de la mejora de las propuestas curriculares, en el desarrollo de programas escolares alternativos y en la introducción de nuevas formas de dirección de las instituciones. No será, sin embargo, hasta la década de los ochenta cuando se empiece a considerar al profesorado como agente de cambio, transformación e innovación en los desarrollos institucional y profesional. A partir de estos momentos la evaluación docente cobra un impulso especial.

Obviamente, el hecho de que hasta el último tercio del siglo XX no se haya acometido de forma sistemática la evaluación del profesorado no quiere decir que hasta ese momento no se hubieran arbitrado prácticas y procedimientos para valorar su competencia. La historia de la educación y de las escuelas nos proporcionan innumerables ejemplos de como los profesores han sido continuamente evaluados tanto para su selección y acreditación profesional como para su promoción, desarrollo y mejora. Las destrezas del profesorado así como los rasgos asociados a su personalidad, a su presencia física y a otras muchas características personales han sido, junto al rendimiento escolar de los alumnos, los elementos que más frecuentemente se han utilizado en la valoración docente y sobre los que se han fundamentado, en no pocas ocasiones, los informes de las visitas de inspección.

En este tema vamos a tratar de contextualizar, primero, cómo va surgiendo la evaluación docente; analizaremos después los propósitos que pretende la evaluación de este ámbito; y, por último, presentaremos, al igual que se hizo en el caso de la evaluación de centros, las diferentes opcio-

nes o formas de concebir la función docente, el trabajo y la profesionalidad del profesorado en tanto que elementos que sustentan y estructuran su evaluación. Es obvio que no es lo mismo situarse ante el profesorado considerando que su labor consiste en la aplicación regular de un conjunto de reglas y procedimientos técnicos invariables para la consecución de unos objetivos previamente establecidos, que cuando se acepta que son profesionales con capacidad y conocimientos para discernir, por sí mismos, los tratamientos y actuaciones que requieren los problemas singulares que surgen en un contexto cambiante y complejo como el del aula. Los enfoques técnico y de reflexión sobre la práctica, junto a la perspectivas tradicional académica y de la reflexión en la práctica para la reconstrucción social, suponen formas contrapuestas de definir el trabajo y la profesionalidad de los docentes y, por consiguiente, maneras distintas de plantear y llevar a cabo su evaluación.

1. ACERCA DE LA EVALUACIÓN DEL PROFESORADO

Tal y como señalábamos en la Introducción, la evaluación del profesorado no es algo que hay surgido recientemente, aunque sí es verdad que, a lo largo de la última década del siglo XX y la primera del siglo XXI, como consecuencia de las reformas emprendidas en favor de la mejora de la calidad de los sistemas educativos, muchos países de la OCDE establecieron sistemas de evaluación del desempeño docente como un elemento integral de las políticas docentes y escolares más generales, Isoré (2010: 4).

Para Darling-Hammond (1997: 23) son precisamente los conceptos y movimientos a favor de la «profesionalidad del profesorado y la reestructuración de la escuela», durante la década de los 80, los que van a resaltar y potenciar el protagonismo de las escuelas y los profesores en los nuevos planteamientos con los que se abordarán las reformas de los sistemas educativos. Shinkfield y Stufflebeam (1995: 9) señalan, por su parte, que la publicación en EE. UU. del Informe *A Nation at Risk: The Imperative of Education Reform*, en el año 1983, fue el catalizador que hizo posible la aparición de documentos políticos y normativos que reforzaban la importancia y regulaban las prácticas de evaluación del profesorado y que, por otra parte, fomentó el estudio de nuevos enfoques y métodos.

La evaluación de los docentes cobra, desde estas nuevas circunstancias, una importancia estratégica aun cuando no pueda evitar que el dilema

entre sus propósitos de control y de mejora se acentúe, ya que, según Popkewwitz (1990: 4), aunque parece claro que se establece una vinculación entre profesionalismo y calidad de la educación, las apelaciones al profesionalismo en las reformas americanas «incrementaron la carga de trabajo de los profesores y el control de las prácticas docentes». Y, por otra parte, si se analizan la retórica de los discursos reformistas sobre la autonomía de los profesores y los espacios y actuaciones docentes que realmente se han favorecido, puede observarse que lo que ha ocurrido, de acuerdo con Contreras (1996: 96), es que «el énfasis de la influencia externa a centros y docentes ha pasado de la prescripción de contenidos y métodos de enseñanza, a la prescripción de los procesos por los que aquéllos llegarán a elaborar sus propias innovaciones».

A. Shinkfield y Stufflebeam (1995: 88) señalan algunas de las razones que justifican la evaluación del profesorado, destacando las siguientes:

- Fomentar una mayor calidad en la atención a los estudiantes.
- Ayudar a los profesores a evaluar y mejorar sus competencias.
- Motivar y ayudar a los profesores para que evalúen y mejoren continuamente su práctica pedagógica.
- Favorecer la rendición de cuentas de los profesores.
- Reconocer y premiar a los profesores que destacan.
- Identificar y prestar ayuda a los profesores con dificultades o deficiencias.
- Proteger los intereses de los alumnos y de la comunidad de aquellos profesores que les pueden perjudicar.
- Despedir a aquellos profesores que de manera continua obtienen malos resultados.
- Supervisar y coordinar la colaboración entre los profesores.
- Valorar la práctica profesional del profesorado para planificar su desarrollo profesional.
- Fortalecer la credibilidad de la escuela.

Las discrepancias e incompatibilidades en torno a los propósitos, las dudas e incertidumbres que surgen en relación con los distintos elementos o componentes desde los que se definen las competencias y prácticas pedagógicas de los docentes, más que impedimentos para el desarrollo de la evaluación del profesorado no son sino el reflejo de las múltiples perspectivas desde las que se puede abordar el mismo. Nadie cuestiona su necesidad e importancia dado el papel relevante que los profesores juegan en una enseñanza de calidad. Pero es que, además, tal y como señala Gimeno Sacristán (1993: 23)

«para negar la posibilidad de la evaluación de profesores habría que partir del supuesto de que su trabajo es una labor intangible, etérea, sin criterios de calidad, que no es provocadora de efectos apreciables de alguna forma y, por tanto, imposible de someter a escrutinio público, sin poder contrastar cómo la ejecutan profesionales y centros distintos, sin poder mejorarla o formarse para su perfeccionamiento, etc.».

El profesorado como ámbito específico de evaluación constituye un campo de estudio complejo, en tanto que no es poco lo que nos falta por saber sobre la naturaleza e interrelación de los múltiples factores que intervienen en una enseñanza de calidad. Es, asimismo, polémico y controvertido por cuanto sus propósitos no están exentos de orientaciones políticas que, de una u otra forma, definen el para qué de la evaluación. Sus enfoques y modelos, por otra parte, dejan traslucir las concepciones que sobre la enseñanza y la profesión docente mantienen los promotores de los mismos. Las metodologías empleadas proyectan, también, las bases epistemológicas desde las que se construye el conocimiento sobre las prácticas pedagógicas o comportamientos docentes que son objeto de evaluación. Y todo ello, además, se despliega de acuerdo con una dimensión ética cuyos valores y principios no suelen ser unívocos para las distintas audiencias de la evaluación.

En este marco, no pueden desconocerse tampoco las resistencias generadas ante su evaluación por los mismos profesores. Existen múltiples referencias en que se recogen los conflictos y problemas planteados ante las propuestas de evaluación de los docentes. En nuestro país, distintos estudios realizados con profesores, recogen las actitudes controvertidas y, en general, las tendencias de rechazo al tema de la evaluación docente, así como al cuestionamiento del ejercicio de la función evaluadora por parte de la inspección, los consejos escolares, los padres y directores.

Concretamente, Mateo y otros (1996: 144) llegan a la conclusión de que en la Educación Secundaria

«los profesores manifiestan opiniones en relación con su evaluación que ponen de manifiesto que siguen considerando que las instituciones educativas constituyen organizaciones excesivamente cerradas, inspiradas bajo principios burocráticos y jerárquicos y, en consecuencia, consideran que este debe ser el marco desde donde cabe abordar los procesos de evaluación. No ha de extrañar que, en ese modelo, rechacen la función evaluadora más allá del nivel de los propios colegas (incluyendo el jefe de seminario o director/a de centro)».

Esta posición que en el contexto español es un reflejo de los temores de los docentes ante el control administrativo y, aunque controvertida, podría entenderse si las organizaciones tuvieran una cierta implicación en procesos de innovación y transformación institucionales, si como modalidad de evaluación su propósito fuera el desarrollo y el aprendizaje profesionales y si el profesorado estuviera realmente comprometido en la elaboración e implementación de proyectos pedagógicos consecuentes con los valores educativos de la enseñanza, pero todo ello es más que dudoso. Porque es obvio que si esto sucediera, los profesores dispondrían de un bagaje de experiencias y herramientas con los que hacer frente legítimamente a propuestas de evaluación externa con sesgos gerencialistas.

En cualquier caso, la realidad es la que es, y en lo que afecta a España como acertadamente señala Escudero Muñoz (2010) seguimos teniendo sistemas de selección del profesorado manifiestamente mejorables, sin que se haya abordado en años su reforma, y además las propuestas para la evaluación del profesorado siguen durmiendo el sueño de los justos, si nos atenemos a que en el sistema de enseñanza no universitario solamente se lleva a cabo para la acreditación de los directores de centros y para la concesión al profesorado de licencias de estudios.

2. TENDENCIAS Y CARACTERÍSTICAS DE LA EVALUACIÓN DEL PROFESORADO

La Red Europea de Información sobre Educación en la publicación de EURYDICE (2009) presenta los resultados del un estudio que llevó a cabo sobre las características de la evaluación del profesorado en Europa (los 27 países de la Unión Europea, más Islandia, Liechtenstein y Noruega). En el

mismo se habla de los cuatro rasgos que caracterizan a las propuestas y realizaciones acometidas en este ámbito de evaluación, a saber:

En primer lugar, se puede afirmar que existen diversas formas de acometer la evaluación que van desde la clásica inspección individual externa, preocupada especialmente por los procesos, hasta la autoevaluación del centro en la que se incluye un análisis del trabajo docente del conjunto de su profesorado, pasando por las entrevistas personales que realiza la dirección del centro.

Como segundo rasgo habría que destacar que hasta ahora se han llevado a cabo considerando el análisis de resultados junto con la descripción y observación de procesos educativos.

El tercer rasgo que define en esos momentos a la evaluación es la tendencia al equilibrio entre las dimensiones individual y colectiva. En los años 90 se hizo hincapié en la evaluación de los equipos docentes, mientras que a mediados de la primera década del 2000 se fueron reforzando los mecanismos de evaluación individual.

El cuarto y último rasgo pone de manifiesto que a pesar de que en la mayoría de los países europeos se han ampliado los cometidos del profesorado y se han implantado procesos de evaluación, estas actuaciones no han sido correspondidas con el desarrollo y aplicación de sistemas de incentivos.

Lo que el análisis de las propuestas de evaluación muestra es que las tareas tradicionalmente asociadas a las actividades de enseñanza y aprendizaje en el aula siguen siendo los elementos nucleares del trabajo educativo, conservando los docentes un notable margen de autonomía profesional en las decisiones que han de tomar sobre ellas. Pero conviene recalcar y no perder de vista que:

«Por encima de estas competencias profesionales, relacionadas estrechamente con el trabajo dentro del aula y no necesariamente ejercidas en colaboración con otros profesores y profesoras del centro, es posible observar una creciente ampliación del abanico de responsabilidades del docente, que le obligan a participar en actividades más allá de los límites de su relación directa con el alumnado. Esta extensión del campo de acción profesional puede achacarse a diversas causas. Las responsabilidades del profesorado evolucionan porque el propio contexto escolar en el que se desenvuelven han cambiado radicalmente en los últimos 20 años» (EURYDICE, 2009: 71).

En efecto, así como hoy nadie cuestiona que la docencia trasciende los cometidos propios de impartir o «dar clases», o que el conocimiento de una asignatura es una condición imprescindible pero no suficiente para ser un buen docente, también se ha ido aceptando que, aunque la capacidad y formación individuales del profesorado son prerequisites para una enseñanza de calidad, las actuaciones colegiadas en el trabajo docente son un factor no menos decisivo para dichos propósitos. Y ello a pesar de que la educación pública y el profesorado de este país esté sufriendo, en estos momentos de crisis y recortes, las consecuencias de unas políticas educativas que, fruto de la incompetencia e ignorancia, todo lo centran en la ampliación de las horas de clase y en la reducción de todo lo que implica el trabajo colaborativo y de coordinación del profesorado y del resto de los actores educativos.

En este sentido, convendría que los profesores no olvidáramos que la educación es un cometido que requiere del protagonismo, compromiso y actuación coordinada de los miembros de la comunidad educativa, en general, y del profesorado en particular; porque el desarrollo profesional y el crecimiento de los docentes no puede considerarse al margen de la consolidación y desarrollo institucional; así como que la escuela no deviene o se conforma como institución por la agregación individual de sus distintos actores sino por el hecho de que el proyecto que justifica su existencia se construye a través de la concertación de voluntades, de la suma de esfuerzos y del hecho de afrontar conjuntamente sus desafíos.

Todo esto lo que pone de manifiesto y en definitiva refuerza es que, por un lado, la educación es un proyecto que ha de ser compartido por los diferentes actores que intervienen en la misma, a través de su coordinación y trabajo conjunto. Y, por otro lado, que la escuela además de preocuparse por su cometido central, la enseñanza, no puede dejar de lado el hecho de ser una institución que también aprende, capaz de generar conocimiento y de incorporarlo a sus quehaceres y actuaciones educativas. En consecuencia estos cometidos y quehaceres deberían estar presentes en la evaluación docente.

Pero las resistencias, tal vez, no sean sino la expresión de que la evaluación docente, como advierte Isoré (2010:4), en un estudio sobre las prácticas vigentes en los países de la OCDE

«ha sido siempre un tema muy controvertido, con evidencia empírica mixta con respecto a sus efectos en el aprendizaje de los estudiantes y con conflic-

tos de interés entre los principales actores de los sistemas educacionales. Por consiguiente, la evaluación ha sido con frecuencia un ejercicio sin sentido, derivado de los rituales burocráticos requeridos en las escuelas y sobrellevado tanto por los docentes como por los evaluadores».

3. LOS PROPÓSITOS DE LA EVALUACIÓN DEL PROFESORADO

Robert Stake (1989: 13) iniciaba un artículo sobre la evaluación de la enseñanza planteándose los objetivos que justifican y dan sentido a la evaluación del profesorado, señalando que al menos son cuatro, a saber:

1. «Proporcionar datos para recompensar el mérito y corregir las deficiencias.
2. Ayudar a la selección de los profesores mejores cualificados para puestos nuevos y retener a aquellos más necesarios en puestos antiguos.
3. Apoyar la formación continua del profesorado.
4. Contribuir a que se entienda el funcionamiento de la escuela en su totalidad».

En estos cuatro objetivos se contienen los tipos de evaluación del profesorado que normalmente señalan la mayoría de los autores que trabajan en el campo. Podríamos hablar, en efecto, de una evaluación dirigida a la selección del profesorado para acceder al puesto de trabajo, de una evaluación que persigue valorar el rendimiento y cualificación de los docentes en ejercicio vinculada con la promoción personal pero también con decisiones sobre la continuidad laboral del profesor, de una evaluación de los profesores como parte de la evaluación integral de las escuelas que, en su orientación de mejora y desarrollo institucional, coincidiría con aquella otra que pretende fundamentar la formación permanente del profesorado.

Darling-Hammond (1997: 13) resalta que las distintas iniciativas que se han puesto en marcha en EE.UU. durante la década de los 80

«tratan de mejorar la educación captando, preparando y reteniendo dentro del ámbito docente al profesorado competente y utilizando sus conocimientos y talentos a lo largo de su nueva carrera profesional. Estas iniciativas se apoyan en la evaluación del profesor para una diversidad de propósitos, incluyendo la selección, la formación, la mejora y el progreso».

A tenor de cómo se abordan cada uno de estos propósitos, en la obra que coordina esta autora junto a Jason Millman, la evaluación para la selección de profesores contempla, a diferencia de la propuesta de Robert Stake, la valoración de los docentes tanto para su acreditación o habilitación profesional como para su contratación, analizándose además todo un conjunto de modalidades de evaluación relacionadas con el profesorado en prácticas, con la valoración de las licenciaturas o titulaciones de los profesores, con la permanencia o despido de profesores, etc. Otras tantas situaciones de evaluación se contemplan en los Estándares de Evaluación de Personal norteamericanos (Stufflebean, 2001), donde además de las ya señaladas se abordan, entre otras, las evaluaciones para la revisión de la práctica profesional, para el pago por méritos y para las decisiones sobre titularidad.

Convendría no perder de vista, sin embargo, que una de las características fundamentales de la evaluación del profesorado es la de formar parte de un proceso más amplio de valoración que, lógicamente, es el de la escuela. A esta importante consideración se refiere Stake (1989: 13)

«la evaluación de la enseñanza es una parte inseparable de la evaluación de la propia escuela. Una evaluación efectiva de la enseñanza requiere un estudio simultáneo de los objetivos institucionales, los entornos de la clase, la organización administrativa y el funcionamiento, el contenido curricular, el logro del estudiante y el impacto de los programas escolares en la comunidad. La enseñanza puede juzgarse correctamente sólo en el contexto de estos otros factores y, si no se hace un esfuerzo por estudiarlos, la evaluación de la enseñanza carecerá probablemente de validez».

Esta posición es también compartida por otros autores, como, por ejemplo, Natriello (1997: 49)

«el reconocimiento de que la evaluación del profesor tiene lugar dentro de un contexto organizativo es vital para comprender los propósitos fundamentales de los sistemas de evaluación del profesorado: la evaluación del rendimiento de los profesores individuales se lleva a cabo de forma constante por parte de los estudiantes, padres, otros profesores y el público en general. Pero sólo cuando el proceso de evaluación está ligado a la organización de la escuela lo convertimos en un sistema e intentamos estudiar sus propiedades formales y los propósitos deseados y sus posibles consecuencias».

Planteada de esta forma, ya sea en el nivel general de la escuela o dentro de su contexto organizativo, los propósitos y efectos de la evaluación del profesorado no quedan sólo reducidos a la dimensión individual o personal de los docentes, aunque ello sea inevitable y también conveniente, sino que se proyecta en una dimensión institucional que como tal es un marco de valoración más completo y comprensivo de la actuación del profesorado y que, además, puede facilitar formas de trabajo y compromisos colegiados para llevar adelante procesos de mejora. El problema siempre surge, no obstante, cuando ante la diversidad de propósitos que pueden perseguirse, la evaluación de los profesores se instrumentaliza para objetivos o finalidades distintas de las que inicialmente se pretendían. Esta «utilización» de la evaluación es, sin duda, una de las fuentes que genera los recelos de los profesores a ser evaluados y contribuye a desvirtuar los propios procesos de evaluación, atacando de lleno a la credibilidad y posible justicia de la misma.

Desde una perspectiva general, además de para la selección y promoción profesional, los propósitos de la evaluación del profesorado se enmarcan al igual que ocurría con la valoración de las escuelas entre las orientaciones que se inclinan por la responsabilidad o rendición de cuentas y las que promueven los procesos de mejora o de desarrollo profesional. Ambos propósitos han sido percibidos por las escuelas y los profesores, como ya tuvimos oportunidad de ver en el capítulo anterior, como incompatibles sobre todo en algunas de las experiencias evaluadoras llevadas a cabo en los años 70 y 80 por el movimiento de la *accountability*. Sin embargo, planteadas de una manera racional uno y otro tipo de evaluación no tienen por que ser asumidas como contradictorias u opuestas, ya que la valoración del trabajo de los docentes, y más en el sistema público de enseñanza, debería constituirse en el punto de partida desde el que abordar participativamente el desarrollo de los propios profesores, de las escuelas y, por supuesto, de los alumnos.

Tal vez los problemas de la desvirtuación, instrumentalización y usos diversos que se puede hacer de la evaluación podrían amortiguarse y controlarse si realmente se aclararán los auténticos propósitos de la evaluación. A este respecto Iwanicki (1997: 223) defiende

«[es] necesaria la existencia de una visión común acerca del por qué se está evaluando a los profesores en una escuela o un distrito escolar antes de que los esfuerzos para mejorar los procedimientos de evaluación tengan impac-

to alguno. A menudo la resistencia de algunos profesores a ser evaluados no está motivada por unos criterios de rendimiento en particular o por los procedimientos que se están aplicando, sino por los propósitos para los que se pueden utilizar los datos reunidos por medio del proceso de evaluación. Debe haber un acuerdo sobre estos propósitos antes de poder aplicar de forma eficaz un proceso de evaluación».

Compartir y/o acordar los propósitos de la evaluación es, sin duda, bastante más que informar o sensibilizar al profesorado de las virtudes de la evaluación, o que propiciar la participación del profesorado en los procedimientos evaluativos. Estaría más en la línea de una de las conclusiones a la que llegan Mateo, Escudero, De Miguel, Mora y Rodríguez (1996: 146) en una investigación sobre la Evaluación del Profesorado de Secundaria:

«igualmente cabe apuntar que en toda organización que se estructure bajo principios democráticos y participativos, los sistemas de evaluación del personal se establecen a través del consenso de las partes implicadas. [...] en el caso de las instituciones educativas, el modelo de evaluación de su profesorado debe establecerse partiendo del concurso de todas las partes implicadas (administradores y administrados)».

Negociaciones y acuerdos que no siempre resultan fáciles pero que son posibles, tal y como nos relata Elliott (1998) en su análisis sobre la experiencia vivida para la implantación de un sistema de evaluación docente en el Reino Unido entre 1983 y 1986, en el que se logró pasar de un modelo coercitivo a otro en el que se recogían elementos de las culturas profesionales de los profesores. Entre los cambios operados entre uno y otro modelo hay que resaltar que ya en 1985 se empezara hablar del desarrollo profesional como uno de los propósitos de la evaluación, cosa que hasta entonces no se había hecho, que se admitiera la presencia y voz de los profesores en la negociación de las formas de evaluación, y que, aun dentro de un sistema jerárquico, se aceptaran en 1986 fórmulas de revisión entre iguales y de reciprocidad. Elliott (1998: 179) resume, de una manera magistral, los dilemas a los que nos enfrentamos en la evaluación del profesorado cuando afirma que «toda forma de evaluación que limite las oportunidades de los docentes para reflexionar sobre sus prácticas en un discurso libre y abierto con los otros constituye una estrategia para eliminar la cultura progresista innovadora de las escuelas».

Ball (1993: 161) nos advierte de que la progresiva modificación de la dirección de las escuelas desde un estilo profesional-colegial a otro de ges-

ción-burocrático ha contribuido de manera decisiva a la reconstrucción del trabajo docente, pasando a ser considerados los profesores como técnicos y los directores como gestores. En este nuevo marco

«[la] evaluación se ha convertido en una de las características fundamentales de la reconstrucción política y de disciplina de los profesores como sujetos éticos en la década de 1980. Extiende la lógica del control de calidad y de los indicadores de la ejecución al centro pedagógico de la docencia. Lleva consigo la necesidad de soportar la mirada tutelar, haciendo que el profesor sea calculable, descriptible y comparable. Deja inermes a los individuos ante el ojo evaluador y el poder disciplinario».

Lo que parece claro, tal y como han puesto de manifiesto Duke y Stiggins (1997) a través de una serie de estudios e investigaciones sobre los sistemas de evaluación del profesorado norteamericano, es que es muy difícil de compatibilizar las propuestas que tratan de aunar evaluaciones para la responsabilidad (*accountability*) y para el desarrollo profesional. En estas últimas, además, es necesario considerar las características y posiciones de los sujetos de la evaluación, profesores y evaluadores, las características de los sistemas de evaluación y el medio y las políticas con las que se tratan de desarrollar estos sistemas.

Pese a la base de conocimientos que se tiene sobre todos estos aspectos, en el sentido de que están mejor predisuestos los profesores con una actitud abierta a los cambios, a asumir riesgos, a afrontar la retroalimentación de la evaluación; de evaluadores que son más aceptados cuando son percibidos como personas con credibilidad, con relaciones de cooperación, con conocimientos no sólo de evaluación sino de la enseñanza en general y con experiencia pedagógica, etc.; Duke y Stiggins (1997: 184) están convencidos de que es necesario seguir profundizando en aspectos como:

«¿Deberían los profesores con experiencia ser libres de no participar en programas de desarrollo profesional y simplemente tomar parte en las evaluaciones para propósitos de responsabilidad? ¿Qué clase de actividades de crecimiento encajan bajo la rúbrica del desarrollo profesional? [...] ¿Deben todas las actividades de desarrollo profesional representar objetivos individuales o pueden algunas de ellas realizarlas grupos de profesores?».

El discurso del desarrollo profesional puede, por tanto, ser falaz y en muchas ocasiones legitimar procedimientos de control que, lejos de capacitar a los docentes en un proceso de autodeterminación y profesionalización, pre-

tenden comprobar hasta que punto los profesores observan y cumplen las tareas y cometidos prescritos previamente. Holly (1989: 113) precisa que los enfoques de evaluación del profesorado consistentes con los principios del desarrollo profesional tienen, entre otros, los siguientes beneficios:

1. «Los procedimientos y resultados de la evaluación aumentan la capacidad del educador para la toma de decisiones y emisión de juicios sobre su labor docente.
2. La evaluación es considerada como un proceso educativo.
3. La evaluación potencia la práctica profesional de varias maneras:
 - Mediante la reflexión.
 - Haciendo que el conocimiento tácito se haga explícito para que pueda ser cuestionado.
 - Eliminando el aislamiento y reforzando el contacto social y el análisis crítico.
 - Facilitando el establecimiento de una cultura y un lenguaje profesional.
 - Permitiendo que los educadores aprendan de su práctica.
 - Fomentando que se formulen interpretaciones distintas a partir de perspectivas y contextos más amplios.
 - Proporcionando una concepción de la historia y de las direcciones a seguir.
4. La evaluación promueve la indagación y la investigación que, a su vez, contribuyen al avance de la pedagogía».

4. PERSPECTIVAS DE EVALUACIÓN DOCENTE

En un interesante y documentado ensayo sobre la evaluación del profesorado, House y Lapan (1989) señalan que en todos los enfoques de evaluación subyacen una serie de supuestos relacionados con las concepciones que los evaluadores mantienen acerca de en qué consiste la docencia, así como cuáles son las características de la enseñanza. Dichos supuestos se

hacen visibles, aunque no siempre sean explícitos, tanto a la hora de concretar el objeto de evaluación como en el establecimiento de los criterios que se utilicen para valorar las actuaciones de los profesores.

Wise *et al.* (1984) —en House, E. R. and Lapan, S. D. (1989: 56)—, en una investigación realizada en 32 ciudades de EE.UU. sobre las prácticas de evaluación del profesorado establecieron la relación que existía entre éstas y las formas de concebir la enseñanza en tanto que profesión. Identificaron cuatro formas de concebir la enseñanza: como trabajo, como destreza, como profesión y como arte. En cada una de ellas se planteaba la evaluación de un modo distinto. En la consideración de la enseñanza como trabajo se entiende que los cometidos del profesor consisten en poner en juego una serie de procedimientos estándar y rutinizados que normalmente son evaluados a través del control directo del trabajo del profesor en aspectos como el seguimiento de la planificación de las clases que imparte, el rendimiento de los alumnos y otros indicadores de resultados. Los cometidos de los docentes deben ajustarse a los procedimientos predeterminados como ocurre en los procesos de fabricación industrial, siendo el administrador del colegio el que supervisa a los profesores.

La identificación de la enseñanza como destreza supone que los profesores disponen de un conjunto de técnicas y de un conocimiento especializado que se refiere no sólo a los contenidos de la enseñanza sino también a la forma cómo pueden ser aplicadas las destrezas. La evaluación, en este caso, se centra en determinar de modo indirecto si el profesorado posee las técnicas y habilidades requeridas para la docencia que, de alguna manera, constituyen el conjunto de estándares que sirven para que los gestores establezcan los niveles de rendimiento general del profesorado, a la vez que para especificar el objeto y los criterios de evaluación.

La concepción de la enseñanza como profesión implica que el profesorado además de las técnicas y destrezas posee conocimientos para saber poner en práctica, en cada momento y circunstancia, las que tienen que ser aplicadas en función de los procesos y actividades singulares de que se trate. Es decir, los conocimientos teóricos y destrezas de los profesores se proyectan en su desempeño profesional más en la forma en que actúan los médicos que como lo hacen los técnicos. Los estándares de rendimiento son establecidos entre pares, por lo que la evaluación se realiza entre colegas y son ellos los que determinan la calidad de las prácticas pedagógicas.

La enseñanza como arte significa, sobre todo, que los procedimientos y técnicas se personifican en la figura de cada profesor en vez de ser estandarizadas e iguales para todos. La imprevisibilidad y especificidad de las dinámicas educativas no pueden resolverse acudiendo a respuestas preestablecidas sino que exigen soluciones personales e individuales de los profesores. Desde estos supuestos su evaluación no va a consistir en un análisis de sus características y comportamientos específicos sino que se centrará en las dimensiones holísticas y actuaciones globales de los profesores mediante las modalidades de autoevaluación y de evaluación entre iguales.

Estas cuatro formas de conceptualizar las labores de los docentes y, por tanto, de definir cuales son las competencias y cometidos de los profesores se corresponden históricamente con las líneas de investigación didácticas desarrolladas desde los distintos paradigmas contemporáneos (Pérez Gómez, 1985; Shulmann, 1989), o con las diferentes perspectivas desde las que han ido evolucionando las funciones de los docentes (Pérez Gómez, 1994; Salinas, 1993; Contreras, 1997). Las consideraciones de la enseñanza como trabajo y destreza son propias de las distintas fases por las que han transitado la tradición científico-racional, la perspectiva técnica y los enfoques tecnológicos de la innovación, encuadrables en los paradigmas presagio-producto y proceso-producto. En dichas fases, y desde principios del siglo XX hasta bien avanzados los años 70, han sido evidentes la influencia del movimiento de gestión científica y la hegemonía de los planteamientos conductistas y su contribución a la defensa de concepciones de la educación como un proceso de producción, de características similares a los de la industria, o como un proceso tecnológico guiado por objetivos terminales.

La aplicación mecánica de los modos de gestión científica de la industria a la escuela fue una práctica frecuente en los primeros años de nuestro siglo. Se trataba simplemente de emular los procedimientos y formas mediante los que la empresa lograba incrementar la eficiencia de su producción. En el ámbito escolar la tarea no parecía difícil. Era cuestión de establecer unos estándares y diseñar, a la vez, unos mecanismos que permitieran medir cuantitativamente el grado en que los alumnos y profesores satisfacían dichos estándares.

La eficiencia escolar, desde estos planteamientos, pasaba ineludiblemente por la determinación de unos objetivos comportamentales específicos, definidos operativamente, y el establecimiento de unos mecanismos de control que

pudiesen medir el grado de cumplimiento de los mismos. Todo ello descansa en los supuestos de que los *outputs* educativos pueden ser evaluados en términos de elementos discretos, que el problema educativo es esencialmente un problema tecnológico y que el trabajo de profesores y supervisores es el de ser unidades que operan en un proceso convergente y establecido.

El paso del tiempo permitió la evolución de este planteamiento hacia propuestas más acordes a cada momento. Desde luego, tal y como afirman Tanner y Tanner (1987: 172), el legado moderno del movimiento de producción de principios de siglo, lo que queda son una serie de mecanismos, tales como:

«... los contratos de rendimiento, los objetivos comportamentales, la dirección por objetivos, las pruebas de competencia mínima, el “*task-on-time*” (tiempo en las tareas o tareas controladas temporalmente), la investigación proceso-producto, el análisis de sistemas y la tecnología instructiva».

Toda esta línea de investigación de la enseñanza, propia del modelo racionalista y del paradigma proceso-producto —véase Ángel Pérez (1985: 88 y ss)— es equiparable, en bastantes ocasiones, a la preocupación y movimientos óptimos de los obreros en las cadenas de producción de cara a efectivizar al máximo su productividad. El modelo de producción se ha fijado preferentemente, como ya hemos apuntado con anterioridad, en la investigación que trataba de poner a prueba la incidencia del tiempo en el aprendizaje de tareas. Inicialmente, la investigación del tiempo en la tarea (*time-on-task*), y con posterioridad, los estudios sobre el tiempo de aprendizaje académico (*Academic Learning Time*) han tratado de relacionar estas variables con el rendimiento escolar superior.

Los estudios del tiempo en la tarea lo que han hecho es fijarse y controlar la cantidad de tiempo que los estudiantes están dedicados a tipos específicos de tareas. Pero el tiempo dedicado a la tarea, para cualquiera que está familiarizado con la actividad de una clase no tiene porqué relacionarse necesariamente con el rendimiento, entre otras cosas porque muy bien pudiera ocurrir que el estudiante esté perdiendo el tiempo (puede parecer que atiende o trabaja pero está mentalmente en otras cosas). El tiempo de aprendizaje académico como línea de investigación más perfeccionada ha tratado de ver la cantidad de tiempo que el estudiante está dedicado a una actividad académica a la vez que da respuestas correctas. Los estudios, en este caso, han puesto de manifiesto que estas medidas se relacionan significativamente con el rendimiento superior. Sin entrar a cuestionar el tipo de relación entre las

variables, la crítica hacia estas investigaciones coinciden en señalar que el centro de estos estudios han sido la actividad en clase y los resultados de aprendizaje que son más susceptibles de cuantificación, desconsiderando los aprendizajes colaterales y los de orden cognitivo superior.

Tanto las investigaciones del tiempo en la tarea como el modelo del tiempo de aprendizaje académico, no han proporcionado, en contraposición, ninguna medida de la incidencia de las variables cuando profesores y alumnos están dedicados a tareas en las que es necesario poner en juego procesos cognitivos superiores o de orden afectivo o social. Paralelamente, el modelo de producción de la enseñanza y supervisión concibe al profesor más como una unidad de producción o un técnico que como un profesional que disfruta de un alto grado de autonomía. Su papel se reduce a la distribución de la instrucción (*delivery of instruction*). Su objetivo primordial es la mejora de la enseñanza, circunscribiéndose al aula, para que los alumnos, en un tipo de enseñanza dirigida por y para el examen, puedan llegar a alcanzar los niveles estándares que han sido establecidos y serán controlados por instancias ajenas al profesor.

La descualificación del profesor dentro de estas posiciones es a todas luces evidente. Existe una desconexión tajante entre currículum y enseñanza, escapándose del control del profesor aspectos tan fundamentales como los del diseño, desarrollo y evaluación del currículum, que quedan en gran parte en manos de los responsables políticos y de los especialistas en estas materias. Las líneas de investigación docente se fijan, preferentemente, en las «competencias del profesor», es decir, en cualquiera de las actividades realizadas por éste que contribuyan al desarrollo cognitivo, afectivo o psicomotriz del alumnado, así como en el análisis de la interacción en el aula a través del que se intenta relacionar las conductas docentes con el rendimiento escolar de los estudiantes. Pero, sin duda, el hecho de que se reduzca el análisis, Pérez Gómez (1985: 111), a los comportamientos observables

«[es] el punto central de la crítica externa, y el motivo fundamental del rechazo del paradigma proceso-producto. Al considerar sólo las manifestaciones observables y mensurables del comportamiento, tanto del profesor como del alumno, se pierde el significado de los procesos de interacción y el análisis, tratamiento e interpretación de datos puede concluir en una grotesca simplificación, en una caricatura del funcionamiento de lo real».

La identificación de la enseñanza como profesión y arte, en los términos planteados en la investigación de Wise *et al.*, conlleva el reconocimiento para el profesorado de un status y unos roles muy distintos. Los docentes, en

función de su cualificación y autonomía profesional, hacen uso de su capacidad para tomar las decisiones que las circunstancias y dinámicas singulares de la enseñanza exigen en cada momento. La educación y la enseñanza son asumidos como procesos de crecimiento continuo en vez de como un conjunto de fases u operaciones guiados por objetivos terminales.

El profesor, a diferencia de las concepciones precedentes, no es un agente de distribución de la instrucción que lleva a cabo unos fines prefijados de un currículum establecido por una fuente de autoridad externa. Como educador profesional trabaja de forma continua y en estrecha relación con los especialistas, supervisores e investigadores que se responsabilizan directamente del campo del currículum. La consideración de la enseñanza como campo científico, técnico y artístico, precisa de un profesor que sea antes que nada un investigador; capaz de dar respuestas y soluciones a los problemas educativos y favorecer el crecimiento de la persona que aprende como un ser que piensa autónomamente y que es un miembro responsable de la sociedad.

Para House y Lapan (1989: 62) la tipología de concepciones de la enseñanza como trabajo, destreza, profesión y arte, identificadas por Wise *et al.*, no pueden ser consideradas como una jerarquía por la que todos los profesores pasan. Entienden cada denominación como «una forma diferente de percibir lo que puede ser la docencia. Así, si los evaluadores perciben la enseñanza como trabajo o destreza, utilizarán unos estándares muy distintos de los que usarán los evaluadores que se guían por asunciones profesionales o artísticas». En el Cuadro 7.1 puede verse una síntesis de las características de las mismas.

Cuadro 7.1. Perspectivas y asunciones sobre la enseñanza y la evaluación*

Perspectiva docente	Asunciones fundamentales		
	Modelo de rendimiento aceptado	Fuente de los estándares	Evaluación del rendimiento
Trabajo	Habilidades existentes pre-determinadas	Autoridad externa y especificaciones detalladas	Inspección directa y supervisión continua
Destreza	Habilidades prescritas, adquiridas mediante la práctica y la imitación de otros maestros	Autoridad externa y reglamento general	Inspección indirecta y supervisión periódica
Profesión	Diferentes modelos de rendimiento llevan al éxito	Desarrollado por compañeros y posiblemente uno mismo	La calidad es evaluada por los compañeros
Arte	Cada modelo de rendimiento es único	Desarrollado por uno mismo y posiblemente por los compañeros	La calidad es evaluada por los compañeros, uno mismo y la reputación

* Fuente: House, E. R. and Lapan (1989): adaptado de Wise et al. (1984)

Esta pluralidad de posiciones no hace sino reflejar, como señala Escudero Muñoz (1998: 116),

«[en] la literatura sobre el profesor suele aludirse a una serie de calificativos que suponen otros tantos modelos de docentes: profesional técnico, reflexivo, investigador en la acción, abierto a la colaboración, crítico y militante [...] Pues bien, lo que aquí importa destacar es que no hay sistema de evaluación que no suponga un determinado tipo de compromiso, explícito o no, con un modelo de profesor».

La función del docente no puede desvincularse, en este sentido, de las concepciones acerca de cómo debe enseñar y de cómo se plantea la práctica educativa en el aula. Las metodologías se centran en el saber hacer del profesor. Y éste, en su capacidad para enseñar su saber a los alumnos utiliza las estrategias necesarias para hacer que se apropien de los conocimientos, considerados como valiosos, y que realmente se produzca un aprendizaje lo más significativo posible.

A este respecto, Pérez Gómez (1994) presenta y analiza distintas perspectivas de la práctica educativa y de la metodología didáctica que se corresponden con diferentes tipologías de docentes. Para él, habría que distinguir una perspectiva tradicional académica que entiende la enseñanza como un proceso de transmisión de los conocimientos y la cultura históricamente acumulada. El profesor es especialista en las distintas áreas o disciplinas que debe transmitir y cuantos más conocimientos posea mejor podrá llevar a cabo dicha transmisión. En el desempeño docente no se da una diferenciación entre el saber y saber hacer, de manera que para enseñar basta con tener un conocimiento experto de la materia que se imparte. Su estrategia didáctica se basa en seguir la secuencia lógica y la estructura epistemológica de la disciplina que enseña, utilizando una metodología básicamente expositiva-magistral cuyo interlocutor es un supuesto alumno medio, que forma parte de cursos o de grupos homogéneos en cuanto a su edad cronológica y/o sus conocimientos iniciales.

Cabe hablar, asimismo, de una perspectiva técnica en la que el profesorado se caracterizaría por el dominio de las aplicaciones del conocimiento científico producido por otros y convertido en reglas que han de observarse en las actuaciones educativas (Pérez Gómez, 1994: 402). En esta concepción tecnológica, los conocimientos profesionales se sustentan en los componentes aportados por una ciencia básica sobre la que descansa la práctica (psicología, sociología), por una ciencia aplicada de la que se derivan los

procedimientos de diagnóstico y solución de problemas (Didáctica) y por un conjunto de competencias y actitudes que se ponen en juego en la intervención y actuación al servicio del cliente (competencias y habilidades para la comunicación en el aula). Este conjunto de saberes implica una división claramente establecida entre expertos, investigadores, teóricos y prácticos, con un reconocimiento social vinculado al rol que cada uno desempeña.

Esta perspectiva, que se encuadra dentro de la racionalidad técnica con una determinante influencia positivista en la separación entre la teoría y la práctica, reduce la intervención docente a una mera actividad instrumental. El profesorado, en ocasiones, puede actuar en el aula con una metodología activa e investigadora, pero de una manera rutinaria y mecánica en la que apenas hay reflexión y conceptualización críticas en sus decisiones y quehaceres. Por lo general, se limita a enseñar estrategias y técnicas procedimentales sin detenerse en los conceptos, marcos globales y relacionales

En contraposición a las anteriores, existe una perspectiva práctica en la que, a su vez, puede distinguirse dos enfoques. Uno de carácter tradicional y conservador en donde la enseñanza se concibe como una actividad artesanal y, otro, al que se conoce como enfoque reflexivo sobre la práctica. En el primero, el conocimiento profesional del docente se basa en una cultura que se ha ido configurando de forma acumulativa, de generación en generación, en contacto directo con profesores con un gran bagaje práctico que han adquirido por sí mismos en el mundo privado del aula, de manera intuitiva y pragmática. Su metodología dominante es la repetición y, en consecuencia, los alumnos deben reflejar lo más fielmente posible la información y conocimiento que el profesor ha impartido.

La reflexión sobre la práctica, en cambio, tiene una posición definida sobre la necesidad de analizar lo que realmente hace el profesorado cuando se enfrenta a los problemas complejos de la vida del aula, elementos ignorados o ausentes en los presupuestos de la racionalidad técnica y el enfoque tradicional. Los docentes han de situar en el primer plano su reflexión y pensamiento sobre, o antes que, sus estrategias de actuación, de modo que puedan desarrollarse los valores éticos con los que se está comprometido en la actividad pedagógica. Desde su reflexión ante los problemas singulares que emergen en el proceso de enseñanza-aprendizaje, tomará decisiones sobre las estrategias didácticas más adecuadas para cada momento pudiendo emplear desde la exposición magistral a todo el grupo de aula, a

la organización de la clase en grupos de trabajo y estudio, pasando por procedimientos de trabajo individual y autónomo.

La cuarta perspectiva desde la que puede concebirse la enseñanza y la actuación del profesorado es la que se conoce como la reflexión en la práctica para la reconstrucción social. La enseñanza se concibe como una actividad crítica, una práctica social saturada de opciones de carácter ético, en la que los valores que presiden la intencionalidad deben traducirse en principios de procedimiento que rijan y se realicen a lo largo de todo el proceso de enseñanza y aprendizaje. El profesor es concebido como un profesional autónomo que reflexiona críticamente sobre la práctica cotidiana de modo que su actuación reflexiva facilite el desarrollo autónomo y emancipador de quienes participan en el proceso educativo. En esta perspectiva se encontrarían los que se manifiestan como defensores de trabajar en la escuela y el aula una propuesta ética concreta de justicia, igualdad y emancipación social en los procesos de enseñanza (Freire, Giroux, Smith, Apple, Kemmis...) y los que desde posiciones más liberales defiende la coherencia ética entre los principios, intencionalidades y procedimientos educativos democráticos (Stenhouse, Elliott, Macdonald).

La metodología va íntimamente ligada a los procesos de reflexión sobre la acción, y en ella quedan implicados los alumnos, reproduciendo en sí mismos los elementos de participación en grupos, integrando a todos en un proceso de indagación y dialogo permanente, donde aprender a aprender se convierte en la clave de la construcción del alumno como sujeto que ressignifica la realidad junto a los demás en un proceso permanente e inacabado. El trabajo grupal y la responsabilidad personal son las claves que darán vigor a cualquiera de las metodologías que se utilicen en el proceso de enseñanza y aprendizaje desde esta perspectiva.

En sus actuaciones, no obstante, los docentes combinan comportamientos y estrategias de las distintas perspectivas consideradas. Es bastante infrecuente encontrar profesores que trabajen desde los principios y presupuestos de una sola de estas concepciones.

Hoyle, tal y como se presenta en el Cuadro 7.2, ha diseccionado las características de dos tipos de profesionalidad posibles en el desempeño docente, distinguiendo entre una restringida y otra desarrollada. Las diferencias entre ambas podrían corresponderse a las perspectivas de la enseñanza como destreza, práctica artesanal o enfoque tradicional académico frente a la conside-

ración de la docencia como profesión, arte, o en las dos vertientes desde la que puede concebirse la enseñanza y la actuación del profesorado en la reflexión en la práctica, de acuerdo con el análisis realizado con anterioridad.

Cuadro 7.2. Características de la profesionalidad restringida y desarrollada de los docentes, propuesta por Hoyle (1974).

PROFESIONALIDAD RESTRINGIDA	PROFESIONALIDAD DESARROLLADA
Destrezas profesionales derivadas de experiencias.	Destrezas derivadas de una reflexión entre experiencia y teoría.
Perspectivas limitadas a lo inmediato en tiempo y espacio.	Perspectivas que abarcan el más amplio contexto social de la educación.
Sucesos y experiencias del aula percibidos aisladamente.	Sucesos y experiencias del aula percibidos en relación con la política y con las metas que se tracen.
Metodología fundamental introspectiva.	Metodología basada en la comparación con las de los compañeros y contrastada con la práctica.
Valoración de la autonomía profesional.	Valoración de la colaboración profesional.
Limitada participación en actividades profesionales no relacionadas exclusivamente con la enseñanza en el aula.	Alta participación en actividades profesionales adicionales a sus enseñanzas en el aula (p.e. participación en actividades de los Centros de profesores, asociaciones profesionales, investigación ...)
Lectura poco frecuente de literatura profesional.	Lectura regular de literatura profesional.
Participación en tareas limitadas de perfeccionamiento a cursos prácticos.	Participación considerable en tareas de perfeccionamiento que incluyen cursos de naturaleza teórica.
La enseñanza es vista como una actividad intuitiva.	La enseñanza es vista como una actividad racional.

Las características de uno y otro tipo de profesionalidad ponen en evidencia, al igual que sucede en las distintas concepciones de la enseñanza y docencia, que las propuestas y desarrollo de la evaluación del profesorado variarán sustancialmente desde las diferentes opciones desde las que se definen sus desempeños y quehaceres educativos. Y lo que es no menos importante, que los enfoques y/o modelos de evaluación, en todos los casos, proyectan de manera tácita o explícita dichas concepciones.

5. A MODO DE CONCLUSIONES

El profesorado es un ámbito de evaluación complejo y problemático. Una parte considerable de las dificultades que han surgido en torno al mismo provienen especialmente de las resistencias de los docentes a ser evaluados, sobre todo en lo que se refiere a su desempeño profesional. Buena prueba de ello es que, por ejemplo, en el contexto español no se ha logrado concretar ninguna propuesta en las enseñanzas no universitarias, más allá de las evaluaciones que se llevan a cabo para la acreditación de los directores de centros o para la concesión al profesorado de licencias de estudio. En el contexto internacional la casuística es variada, aunque las prácticas vigentes en los países de la OCDE ponen de manifiesto que es un tema muy controvertido y con conflicto de intereses entre los principales actores de los sistemas educativos.

En cualquier caso, la evaluación del profesorado es sumamente necesaria, al menos para acometer la selección de los docentes más cualificados, además de para retener a los que tienen un desempeño profesional relevante; para apoyar y vincular la formación permanente del profesorado; para recompensar méritos y corregir deficiencias; y para contribuir a que se entienda el funcionamiento de la escuela en su totalidad.

Conviene, no obstante, que quede claro que la evaluación del profesorado no es una actividad neutra. En cualquiera de las propuestas que se proponen, de una u otra forma, subyacen concepciones acerca de en qué consiste la docencia y cuáles son las características de la enseñanza. Consecuentemente, para evaluar es necesario conocer las distintas perspectivas desde las que surgen las propuestas y proyectos de evaluación, ya que en ellas se vislumbran las diversas concepciones y características de la docencia y de la enseñanza.

BIBLIOGRAFÍA

- BALL, S. J. (1993): «La gestión como tecnología moral. Un análisis ludista», en S.J. Ball (Comp.): *Foucault y la educación. Disciplinas y saber*. Madrid, Morata, pp. 155-166.
- CONTRERAS DOMINGO, J. (1996): «Teoría y práctica docente», *Cuadernos de Pedagogía*, n. 253, diciembre, pp. 92-100
- DARLING-HAMMOND, L. (1997): «Evolución en la evaluación de profesores: nuevos papeles y métodos», en J. Millman y L. Darling-Hammond (eds.): *Manual para la evaluación del profesorado*. Madrid, La Muralla, pp. 23-45.
- DUKE, D. L. y STIGGINS, R. J. (1997): «Más allá de la competencia mínima: Evaluación para el desarrollo profesional», en J. Millman y L. Darling-Hammond, (eds.): *Manual para la evaluación del profesorado*. Madrid, La Muralla, p. 184 .
- ELLIOTT, J. (1998): «Saber, poder y evaluación de los docentes», en W. Carr (ed.): *Calidad de la enseñanza e investigación-acción*. Sevilla, Díada Editora, 3ª edición revisada, pp.167-188.
- ESCUADERO MUÑOZ, J.M. (1998): «La Evaluación Docente: Naturaleza y propósitos de la evaluación del profesorado», en J.M. Escudero Muñoz, M^a.T. González y J. del Cerro (eds.): *Evaluación de programas, centros y profesores*. Murcia, ICE Universidad e Murcia-Diego Marín Ed., pp. 107-119.
- (2010). «La selección y la evaluación del profesorado», en *Revista Interuniversitaria de Formación del Profesorado*, 68 (24,2), pp. 201-221.
- EURYDICE (2009). *Autonomía y responsabilidades del profesorado en Europa*. Madrid, Ministerio de Educación.
- GIMENO SACRISTÁN, J. (1993): «El profesorado: ¿Mejora de la calidad o incremento del control?», *Cuadernos de Pedagogía*, 219, noviembre, pp. 22-27.
- HOLLY, M. L. (1989): «Perspectives on teacher appraisal and professional development», en H. Simons y J. Elliott (eds.): *Rethinking appraisal and assessment*. Milton Keynes, Open University Press, pp.100-118.
- HOUSE, E. R. y LAPAN, S. D. (1989): «Teacher appraisal», en H. Simons y J. Elliott (eds.): *Rethinking appraisal and assessment*. Milton Keynes, Open University Press, pp. 55-63.

- HOYLE, E. (1974): «Professionalism, professionalism and control in teaching», *London Educational Review*, 3, 13-19.
- ISORÉ, M. (2010). *Evaluación de docentes: prácticas vigentes en los países de la OCDE y una revisión de la literatura*. Chile, Editorial San Marino, OCDE-PREAL
- IWANICKI, E. F. (1997): «Evaluación del profesorado para la mejoras de las escuelas», en Millman, J. y Darling-Hammond, L. (eds.): *Manual para la evaluación del profesorado*. Madrid, Editorial La Muralla, pp. 222-239.
- JIMÉNEZ, B. (1999): «Evaluación de la docencia», en B. Jiménez (ed.): *Evaluación de programas, centros y profesores*. Madrid, Síntesis Educación, pp. 173-206.
- LAPAN, S. D. (1989): «Teacher appraisal», en H. Simons y J. Elliott (eds.): *Rethinking appraisal and assessment*. Milton Keynes, Open University Press, pág. 56.
- MATEO ANDRÉS, J. y otros (1996): *Evaluación del profesorado de secundaria*. Barcelona, Cedecs.
- MATEO ANDRÉS, J. (2000): «La evaluación del profesorado y la gestión de la calidad. Hacia un modelo comprensivo de evaluación sistemática de la docencia», *Revista de Investigación Educativa*, vol. 18, n. 1, pp. 7-34.
- (2000 a): *La evaluación educativa, su práctica y otras metáforas*. Barcelona, ICE- Horsori.
- MCKENNA, B. y otros (1998): *Guía profesional para la mejora de los sistemas de evaluación del profesorado*. Bilbao, Mensajero.
- NATRIELLO, G. (1997): «Consecuencias deseadas e imprevistas: propósitos y efectos de la evaluación del profesorado», en J. Millman y L. Darling-Hammond (1997): *Manual para la evaluación del profesorado*. Madrid, La Muralla, pp. 49-64.
- PÉREZ GÓMEZ, A. I. (1994): «La función y formación del profesor/a en la enseñanza para la comprensión. Diferentes perspectivas», en J. Gimeno y A. Pérez (eds.): *Comprender y transformar la enseñanza*. Madrid, Edit. Morata, Tercera edición, pp. 398-429.
- (1.985): «Paradigmas contemporáneos de investigación didáctica», en J. Gimeno Sacristán y A. Pérez Gómez (eds.): *La enseñanza: su teoría y su práctica*. Madrid, Akal (2ª ed.), pp. 95-138.

- POPKWITZ, THOMAS S. (1990): «Profesionalización y formación del profesorado», *Cuadernos de Pedagogía*, n. 184, septiembre, soporte CD-Rom.
- SHINKFIELD, A. AND STUFFLEBEAM, D. L. (1995): *Teacher Evaluation: Guide to Effective Practice*. Boston, Kluwer Academic Publisher.
- SHULMAN, L.S. (1986): «Paradigma and Research Programs in the Study of Teaching: A Contemporary Perspectives», en M.C. Wittrock: *Hanbook of Research on Teaching. Third Edition*. New York, Macmillan.
- STAKE, R. E. (1989): «The Evaluation of Teaching», en H. Simons y J. Elliott (eds.): *Rethinking Appraisal and Assessment*. Buckingham, Open University, pp. 13-19.
- WISE, A. E.; DARLING-HAMMOND, L.; McLAUGHLIN, M.W. y BERNSTEIN, H.T. (1984): *Teacher Evaluation: A Study of Effective Practices*. Santa Mónica, Rand Corporation. Referencia tomada de E. R. House y S. D. Lapan (1989): "Teacher appraisal", en H. Simons y J. Elliott (eds.): *Rethinking appraisal and assessment*. Milton Keynes, Open University Press.

Tema 8

La evaluación del desempeño docente

Introducción

1. Cuestiones en torno a la evaluación del desempeño docente
2. Los contenidos de la evaluación del profesorado
3. La evaluación de la competencia y desempeño del profesor
4. La evaluación de los docentes a través de los rendimientos del alumnado
5. Los procedimientos de la evaluación de profesores
6. A modo de conclusión

Bibliografía

TEMA 8 LA EVALUACIÓN DEL DESEMPEÑO DOCENTE

Eustaquio Martín Rodríguez.

UNED

INTRODUCCIÓN

Los enfoques, modelos o propuestas de evaluación del profesorado reflejan, de una u otra manera, los planteamientos, presupuestos y concepciones sobre las características de la enseñanza como profesión, las respuestas que se dan a las preguntas sobre cómo debe enseñarse y de qué manera se afronta la enseñanza en el aula, así como las distintas facetas que definen y configuran la profesionalidad docente. La consideración de estos aspectos implica situar al objeto de evaluación, el profesorado, en toda su complejidad, más allá de soluciones pragmáticas o reduccionistas que ignoran o simplifican las decisiones que deben tomarse para dilucidar el qué de la evaluación

Lo que diferencia la evaluación de los docentes, sobre todo en el impulso que recibe en la década de los años noventa con respecto a otros momentos históricos, no sólo va a ser el desarrollo de nuevos enfoques y métodos que van a contribuir a su sistematización, sino también el auge e importancia que irá adquiriendo como consecuencia de los procesos de descentralización y autonomía que se producen en los sistemas educativos con el consiguiente aumento de la responsabilidad de las escuelas y profesores a la hora de abordar su trabajo y la solución de sus problemas. Si el peso y el eje de gravedad de las innovaciones se desplaza de las propuestas curriculares al compromiso y trabajo de los profesores es lógico que se refuerce la necesidad de que salga a la luz lo que ocurre dentro de las instituciones educativas, y de que las escuelas y los docentes sean valorados tanto en función de sus logros como en relación con sus contribuciones a la mejora de la enseñanza.

En este tema, seguimos profundizando en el estudio de la evaluación del profesorado analizando, en primer lugar, algunas de las cuestiones relacionadas con la concepción y características de la evaluación del desempeño docente. Al objeto de clarificar cuáles son las fuentes de inspiración de los

estándares docentes, analizamos a continuación los contenidos de evaluación, es decir, las distintas dimensiones en que habría que centrar la recogida de datos e información. Presentamos después la Supervisión Clínica como una modalidad de evaluación de desempeño que puede dar respuestas al control y al desarrollo profesional del profesorado. Recogemos, asimismo, las limitaciones y dilemas que implica el uso del rendimiento de los alumnos como criterio único de evaluación del profesorado Y, para terminar, nos hacemos eco de la necesidad de que la evaluación docente cuente con un marco normativo, ético y de procedimientos para acometer la misma.

1. CUESTIONES EN TORNO A LA EVALUACIÓN DEL DESEMPEÑO DOCENTE

A diferencia de la evaluación para la selección o contratación del profesorado, la evaluación del desempeño docente se centra en el despliegue de las obligaciones y responsabilidades que los profesores ponen en juego en sus actuaciones, cometidos y quehaceres en los distintos espacios en que se desarrolla su trabajo, es decir, en los procesos de enseñanza y aprendizaje en el aula, en la coordinación y puesta en común de los planteamientos pedagógicos y didácticos del equipo docente, en los menesteres propiamente institucionales que están vinculados, en general, a la comunidad educativa y, en particular, al centro educativo, etc.

La preocupación por la calidad de la enseñanza y la proyección que en ella tienen los profesores han ido allanando el camino para que se vaya aceptando, poco a poco, la necesidad de instaurar la evaluación de su desempeño docente con al menos dos propósitos. Por un lado, para conocer y garantizar que su trabajo contribuye de manera clara y evidente al aprendizaje de los alumnos y, por otro, para abordar la mejora continua de su propia práctica profesional desde los datos, informaciones y evidencias que ponen de manifiesto sus logros y deficiencias. En definitiva, estos propósitos traslucen las mismas dos orientaciones que analizábamos en la evaluación de centros, la de rendición de cuentas o control de resultados y la de la mejora interna y desarrollo profesional que conllevan, a su vez, el uso de las modalidades de evaluación sumativa y formativa, respectivamente.

En el tema anterior hablábamos sobre las controversias y resistencias que en general ha suscitado en la mayoría de los países la evaluación del

profesorado y, en particular, en España en donde no se ha logrado aún su implantación a pesar de que sus propuestas se contemplan ya en la Ley Orgánica de la Participación, la Evaluación y el Gobierno de los Centros Docentes (LOPEG) de 1995 y que, sin solución de continuidad, las distintas leyes educativas hasta llegar a la Ley Orgánica de Educación (LOE), de 2006, han ido haciendo propuestas al respecto. Como acertadamente señala Bolívar (2008: 67) es preciso que en nuestro país contemos

«con algún tipo de dispositivo (externos, además de internos) que garanticen la equidad de “todos” los alumnos en su derecho a la educación. La cuestión es cómo hacerlo en formas que motiven a los que ya lo hacen bien y, a la vez, contribuyan a mejorar a aquellos establecimientos de enseñanza y profesorado que consiguen bajos niveles en su alumnado, como actualmente se plantea en la “nueva” responsabilidad por los resultados (en algunos países) [...] Asegurar que todo ciudadano está recibiendo la educación que desarrolla sus posibilidades no puede hacerse dejando el asunto al arbitrio (y suerte) contingente de cada centro escolar y su profesorado».

La evaluación externa, de carácter sumativo, del desempeño docente que se va implantando en los países de la OCDE, Isoré (2010), es sobre todo una forma de valorar en qué medida los docentes estén adoptando las acciones y prácticas *óptimas* para mejorar los resultados de los estudiantes. En este sentido, es esencial, acceder a información con respecto a las prácticas que el docente está aplicando, así como al desempeño que está teniendo dicho docente en relación con los estándares establecidos de *buena enseñanza*. Conducida de esta manera la evaluación sumativa es considerada como una fuente de documentación indispensable para responsabilizar al profesorado de su desempeño profesional. Y además fundamenta la toma de decisiones al poderse utilizar para el reconocimiento y retribución de la competencia y el desempeño de los profesores, (OCDE 2005), o para la contratación y permanencia en los cargos, o las oportunidades de ascenso o la toma de medidas ante la ineficiencia, como también ocurre en algunos países.

Como podrá observarse la evaluación del desempeño docente que estamos describiendo tiene como referente central los estándares de la *buena enseñanza*, es decir, el logro de niveles de consecución deseables en los procesos y resultados de la tarea pedagógica, Ferrer (2006). Teniendo en cuenta estos aspectos, la evaluación basada en estándares es un tipo de evaluación vinculada con la tradición psicométrica de la medición y cuantificación

frente a la evaluación cualitativa o interpretativa que especialmente se ocupa de llegar a la comprensión de lo evaluado. Precisamente, el último libro en español de Robert Stake (2006) lleva como título *Evaluación comprensiva y evaluación basada en estándares*, en un intento por parte de nuestro autor de tender puentes entre las dos tradiciones de evaluación que abordan de maneras distintas los estudios evaluativos.

La evaluación del desempeño docente basada en estándares parte, a este respecto, de que hay una relación entre lo que se denomina *buen enseñanza* y las actuaciones eficientes del profesor para la facilitación del aprendizaje del alumnado. De manera que la mejora de los resultados se asocia al dominio de las competencias y/o a los niveles de desempeño de la docencia que los profesores ponen de manifiesto en su ejercicio profesional. Tal y como sintetizan Ingvarson y Kleinhenz (2006: 269),

«[los] estándares profesionales de la enseñanza reflejan el conocimiento, las aptitudes, opiniones y disposiciones de los profesores que pueden ofrecer una enseñanza de alta calidad. Expresan explícitamente los valores centrales de la educación que guían a los profesores eficaces e identifican las características y tareas exclusivas de la enseñanza como profesión. Al describir la trayectoria del desarrollo profesional, los estándares para la enseñanza de alta calidad dejan claro qué necesitan saber y hacer los profesores para mejorar a largo plazo».

La evaluación interna del desempeño docente, desde la orientación del desarrollo profesional, tiene en principio características que la diferencian de la modalidad de evaluación externa que hemos analizado hasta aquí. En principio, se trata de una evaluación formativa que, sobre todo, pretende consolidar el crecimiento profesional de los profesores en todas aquellas capacidades docentes de probada efectividad en los procesos de enseñanza y aprendizaje, así como la búsqueda de soluciones y respuestas a las deficiencias y limitaciones que forman también parte de su desempeño. Por otro lado, no es menos importante que sean los mismos profesores quienes decidan la realización de evaluaciones de diagnóstico y/o de autoevaluaciones a los efectos de conocer cuáles son las luces y sombras presentes en sus cometidos y actuaciones profesionales. En consecuencia, no pretende emitir juicios de valor sobre los niveles de desempeño docentes sino interpretar y comprender qué ocurre y por qué a lo largo de la intervención de los profesores y equipos docentes en el desarrollo del currículum, en la enseñanza y el aprendizaje y en todo lo relacionado con la organización institucional.

En las propuestas de evaluación de ciertos países de la OCDE, como Estados Unidos, Chile, Inglaterra, etc. se han adoptado y adaptado las listas de criterios y estándares que se han ido experimentando en diversas experiencias. Así por ejemplo, Danielson es el autor y promotor de un Marco para la Enseñanza en el que las responsabilidades del profesorado se agrupan en cuatro grandes dominios que, a la vez, se dividen en diversos componentes. Tal y como expresa Isoré (2010: 10), el Marco para la Enseñanza puede

«servir al mismo tiempo como un *mapa de ruta* para guiar a los maestros principiantes a través de sus experiencias iniciales en las aulas, una estructura para ayudar a que los profesionales experimentados se vuelvan más eficaces y un medio para focalizar las iniciativas de mejoramiento».

Veamos, Danielson (2011:14), cuáles son sus cuatro dominios y sus distintos componentes:

- **Dominio 1: Planificación y preparación.** Demostración de conocimientos del contenido y la pedagogía; demostración del conocimiento de los estudiantes; selección de los objetivos pedagógicos; demostración de los conocimientos de los recursos; diseño de una instrucción coherente; diseño de evaluaciones del aprendizaje de los estudiantes.
- **Dominio 2: El ambiente del aula.** Creación de un ambiente de respeto e interacción; establecimiento de una cultura propicia para el aprendizaje; manejo de los procedimientos de aula; manejo de la conducta de los estudiantes; organización del espacio físico.
- **Dominio 3: Instrucción.** Comunicación con los estudiantes, uso de técnicas de interrogación y debate; compromiso de los estudiantes en el aprendizaje; uso de evaluaciones en la instrucción, demostración de flexibilidad y sensibilidad.
- **Dominio 4: Responsabilidades profesionales.** Reflexión sobre la enseñanza; mantenimiento de registros rigurosos; comunicación con las familias, participación en una comunidad profesional (centro y distrito escolar); crecimiento y desarrollo profesional; muestra de profesionalismo.

Cada uno de los componentes contiene distintos elementos que han de ser evaluados. El conocimiento de los estudiantes por parte del profesor, por ejemplo, comprende elementos como el conocimiento de los grupos, el conocimiento de los diversos estilos de aprendizaje de los estudiantes. Cada

elemento de un componente está asociado a cuatro niveles de desempeño: *insatisfactorio*, *básico*, *competente* y *destacado*. Cada uno de estos niveles pueden emplearse, según Danielson, para la supervisión y evaluación, así como para ayudar a la autoevaluación o apoyar las relaciones de tutoría, para informar un debate profesional y para identificar y sugerir las áreas que requieren perfeccionamiento.

Básicamente el proceso seguido se corresponde con los tres pasos que requiere el diseño de estándares de enseñanza para evaluar una práctica, según Ingvarson y Kleinhenz (2006: 270), a saber:

- Definir qué aspectos de la enseñanza de alta calidad deben considerarse en los estándares (¿qué es la enseñanza de alta calidad?)
- Decidir cómo serán medidos estos aspectos (es decir, ¿es también importante cómo se reúnen los méritos relevantes?)
- Identificar qué se necesita para cumplir el estándar, o qué nivel debe alcanzarse.

2. LOS CONTENIDOS DE LA EVALUACIÓN DEL PROFESORADO

En todos los estudios que abordan la evaluación del profesorado, uno de los aspectos claves es determinar y establecer los contenidos de dicha evaluación. A este respecto, las posiciones mantenidas por los distintos autores y escuelas de evaluación pueden ser muy diferentes entre sí. Así, por ejemplo, el Center for Research on Educational Accountability and Teacher Education (CREATE), que comenzó sus trabajos en 1990 en el Centro de Evaluación de la Universidad de Western Michigan, ha seguido casi siempre los planteamientos defendidos por Michael Scriven (1986, 1994, 1997) en relación a que la evaluación del profesorado debería tener muy en cuenta los siguientes aspectos:

- No deberían basarse en elementos o variables seleccionadas a través de estudios o investigaciones correlacionales, ya se refieran éstas a características o atributos de los docentes (sexo, raza, minusvalías, etc.) o a sus estilos de enseñanza, ya que la utilización de estos criterios supone que se da más valor a estas variables que al resultado del proceso de aprendizaje, pudiéndose dar el caso de profesores con bajas puntuaciones en estas medidas de predicción que, sin embargo, contribuyen eficazmente al aprendizaje de los alumnos;

- Tienen que dejar al margen o desconsiderar, igualmente, criterios que independientemente de la magnitud de su correlación pueden minimizar las responsabilidades de los docentes que siempre van a ser consideradas relevantes para la enseñanza como el conocimiento de la materia, o las habilidades para la transmisión de los conocimientos, para realizar actividades de aula, para evaluar el progreso de los alumnos, etc.

En definitiva, Scriven (1994), y con él otros autores como Mckenna, Nevo, Stufflebean y Tomas (1998) y Farland y Gullickson (1998), es partidario de que la evaluación del profesorado se centre en la realización de sus deberes y tareas. Y para él, los deberes y tareas principales, a partir de las que se pueden ir añadiendo otras, son:

1. Conocimiento de la materia
 - Campo de especialización
 - Temas fundamentales del currículo
2. Competencias docentes
 - Habilidades de comunicación
 - Gestión de la clase
 - Desarrollo del curso
 - Evaluación del curso
3. Evaluación de alumnos
 - Exámenes
 - Calificaciones
 - Informes
4. Profesionalidad
 - Ética
 - Actitudes
 - Servicio
 - Conocimiento de funciones
 - Conocimiento de la escuela y su contexto
5. Otros servicios que a título individual preste a la escuela y la comunidad.

Estos deberes y tareas que han de presidir la evaluación de los profesores en ejercicio son más amplios y hasta un poco distintos de las cualidades en las que debe centrarse la evaluación para la selección de profesores. Según Scriven (1997: 124-125) cinco son las cualidades que han de poseer los docentes:

- A. *Conocimientos sólidos de áreas de especialización*, incluyendo materias del currículo transversal (escritura, ciencias sociales, informática); de forma que lo que se enseñe y se aprenda sea correcto.
- B. *Competencia sólida con conocimientos sobre evaluación* y/o observación estructurada; de forma que el profesorado averigüe qué aprendizaje es necesario que tenga al principio del trimestre, si su forma de enseñar es adecuada y proporciona una retroalimentación oportuna en el tiempo y correcta en cuanto a su contenido a alumnos, profesorado y dirección sobre los progresos realizados por los primeros. Los exámenes elaborados, que no elegidos, por el candidato deberían someterse a discusión.
- C. *Capacidad para enseñar en clase*, incluyendo capacidades para la dirección de clase (dirección de procesos —disciplina, controles, etc.—, dirección del progreso, —dar las materias en el tiempo disponible—, y dirección de crisis —modos de proceder ante emergencias: fuegos, inundaciones, etc.—), audiovisuales y otras tecnologías educativas así como producción de materiales de clase (esto es aplicable incluso para el profesorado principiante). Debería demostrarse *in situ* si no se pueden nombrar dos observadores independientes y experimentados.
- D. *Cualidades intelectuales y personales relacionadas con la enseñanza* (profesionalidad, una mente despierta, puntualidad, deseo de ayudar a otros profesores, compromiso con la enseñanza). Nótese que un *entusiasmo aparente con respecto a la enseñanza* no es permisible, aunque el entusiasmo sea una característica de los buenos profesores; se trata de una variable en cuanto al estilo.
- E. *Valor*; la utilidad para la escuela que procede de cursos que no sean de competencia profesional (esto varaiará de una situación a otra, pero a menudo incluye factores como la versatilidad, la familiaridad con el tipo de comunidad, etc.).

Para Scriven (1986) los criterios utilizados para la evaluación sumativa de los docentes deben tener una conexión lógica o legal con el rendimiento de la enseñanza, es decir, que deben ser necesarios para una buena enseñanza. Por ejemplo, el contacto visual o la cantidad de tiempo que un profesor hable en la clase son considerados indicadores secundarios de una buena enseñanza, sobre todo cuando existen otros como su capacidad para controlar la clase o la utilización de procedimientos adecuados para generar una enseñanza de valor. Scriven plantea la necesidad de tener en cuenta cuatro dimensiones de valor para la evaluación sumativa, a saber: la calidad del contenido impartido, el éxito en impartir/inspirar el aprendizaje, el dominio de las habilidades profesionales y los principios éticos. Dichas dimensiones son importantes tanto por separado como conjuntamente, de tal manera que los profesores tienen que ser competentes en todas ellas y cualquier carencia grave o acusada en alguna de ellas no puede ser compensada por su capacidad en las otras pudiendo ser, por tanto, objeto de despidos.

La calidad del contenido impartido no equivale ni es lo mismo que la cantidad del contenido conocido. El primero está directamente relacionado con la enseñanza, mientras que el segundo lo está de una manera indirecta, de modo que son discutibles los tests o pruebas que traten de comprobar los conocimientos de los profesores porque, aunque es cierto que para enseñar es necesario disponer de conocimientos sólidos, también es verdad que no es menos importante verificar cómo se aplican y de qué manera dichos conocimientos. Pero además, el éxito en impartir/inspirar el aprendizaje de los alumnos no debe limitarse a los contenidos de carácter cognoscitivo sino que debe abarcar su desarrollo personal entendido en su más amplio sentido. De igual modo, las habilidades profesionales como los procedimientos de evaluación o el control de la clase deben ser valoradas adecuadamente dada su relevancia para articular y desarrollar una enseñanza de calidad. Y todo ello carecería de valor si la actuación de los docentes no está inspirada por una serie de principios éticos que eviten comportamientos racistas, favoritismos, etc.

House y Lapan (1989) defienden que los enfoques de evaluación del profesorado deberían tener muy en cuenta cómo los docentes aprenden a enseñar, ya que hay determinados aprendizajes que son más evidentes que otros para su valoración. Por ejemplo, es mucho más fácil evaluar el control y la disciplina de la clase, aprendizajes a los que el profesor se enfrenta sólo en los inicios de su experiencia profesional, que la forma en que los docen-

tes conducen los debates en el aula o como seleccionan determinados temas o problemas para mantener la motivación y el progreso de los estudiantes. Estos últimos son tipos de conocimientos profesionales mucho más complejos y sutiles y distinguen a los buenos profesores en su disposición a seguir aprendiendo permanentemente.

Por otro lado, tampoco puede olvidarse que, aunque una parte del conocimiento práctico de los profesores es explícito por lo que no es difícil que puedan explicar u ofrecer pautas para saber cómo lo hacen, la mayor parte de sus conocimientos son implícitos, es decir que los docentes saben como hacerlo aunque tengan dificultades o sean incapaces de explicar como lo hacen, lo cual representa siempre un inconveniente para demostrar su capacidad para enseñar; sobre todo, porque los observadores no perciben fácilmente este tipo de conocimientos ni el profesor puede explicarlo.

Teniendo en cuenta las características de los conocimientos profesionales del profesor y de que la enseñanza depende en gran medida de la estructura deductiva de cada docente, House y Lapan (1989: 60) son partidarios de que su evaluación se centre en dicha estructura deductiva, considerando las siguientes tres posibilidades:

«Una es averiguar si el profesor todavía experimenta de una forma activa con la enseñanza. Los auténticos buenos profesores siempre están probando nuevas cosas para mejorar la calidad de su enseñanza y para adaptar a sus estudiantes a unos tiempos en continuo cambio. Los profesores mediocres, por otro lado, desarrollan un modelo fijo y no lo cambian a pesar de lo variable que puedan ser las circunstancias.

Una segunda posibilidad es investigar si el profesor puede explicar y justificar los movimientos docentes que hace. Pese a que hemos argumentado que gran parte del conocimiento docente es implícito, creemos que los buenos profesores serán capaces de dar razones que argumenten lo que hacen en la clase y los malos profesores no podrán hacerlo. Es verdad que gran parte del conocimiento docente no es explícito pero una buena parte se puede obtener mediante técnicas de inventiva. El reto está en desarrollar dichas técnicas.

La tercera posibilidad es comprobar las deducciones del profesor frente a aquéllas de sus compañeros. Estos últimos compartirán muchas de las asunciones fundamentales que tiene el profesor y, aunque esto no ayude a

que el conocimiento implícito se convierta en explícito, sí que ayudará al profesor a probar sus ideas frente a aquéllos que se encuentren en la misma situación. Podemos prever un conjunto de procedimientos de evaluación que se basarían en las deducciones causales del propio profesor y que ayudarían a los profesores a comprobar sus propias deducciones».

3. LA EVALUACIÓN DE LA COMPETENCIA Y DESEMPEÑO DEL PROFESOR

Una de las propuestas que se recogen en todos los estudios sobre evaluación de los profesores es la Supervisión Clínica. Este enfoque cuenta en su haber con una larga tradición en sus experiencias de desarrollo y cualificación del profesorado. Como se sabe, la misma tiene su origen en la Escuela de Educación de la Universidad de Harvard, en los últimos años de la década de los 60. Concretamente nace dentro del <<Master of Arts in Teaching>>, promovido por el Fondo de la Fundación Ford para el Progreso de la Educación. La idea original del programa consistía en el intento de formación de un profesor de élite, con una gran cualificación, que estuviera al tanto de todos los avances de la educación y que se dedicara con exclusividad a la enseñanza práctica de los futuros profesores.

El término <<clínico>> aplicado a este tipo de Supervisión tiene evidentemente connotaciones médicas, pero se refiere, en este caso, en contraposición a los estudios experimentales de laboratorio, a la posibilidad de una práctica basada en el tratamiento y observación real de los pacientes. Desde esta posición, la Supervisión Clínica ha sido definida por Morris Cogan (1973: 9), uno de sus mayores defensores, como la:

«... fundamentación y práctica diseñada para la mejora de la actuación del profesor en el aula. Toma sus datos principales de los acontecimientos del aula. El análisis de estos datos y la relación entre profesor y supervisor forman la base del programa, de los procedimientos y de las estrategias diseñadas para la mejora del aprendizaje de los alumnos a través de la mejora de la conducta de los profesores en el aula».

Sergiovanni y Starratt (1983: 305), hablan de que la Supervisión Clínica «se refiere a los encuentros cara a cara con los profesores acerca de la enseñanza, usualmente en las aulas, con el doble propósito del desarrollo profesional y el mejoramiento de la instrucción».

Tanto el origen como las definiciones dadas de la Supervisión Clínica vienen a confirmarnos que se trata de un tipo especial de Supervisión, centrada en la observación del comportamiento del profesor en el aula. Los datos que se recogen de estas observaciones se utilizan como base para trabajar con los profesores de cara a una mejora de su enseñanza, en una relación personal entre profesor y supervisor.

El proceso de Supervisión Clínica requiere que se recorran una serie de fases o etapas. Cogan (1973: 10-12) hace una propuesta, aún más casuística, describiendo las siguientes etapas:

1. Establecimiento de la relación entre profesor y supervisor.
2. Planificación con el profesor (una lección, series de lecciones o una unidad —normalmente incluye la especificación de resultados—, previsión de los problemas con los que se va a encontrar la enseñanza, previsión de materiales y sistemas que se prevén también de retroalimentación y evaluación).
3. Planificación de las estrategias de observación.
4. Observación de la enseñanza.
5. Análisis de los procesos de enseñanza-aprendizaje.
6. Planificación de la estrategia para los intercambios, reuniones del supervisor y profesor.
7. Realización de la Conferencia de Supervisión.
8. Nueva planificación (planificación renovada para las lecciones siguientes, abarca los cambios en que supervisor y profesor se han mostrado de acuerdo).

Como puede apreciarse se trata de un proceso sistemático, pero no rígido, ya que no hay un único estilo de Supervisión Clínica y, por otra parte, se pueden encontrar diversidad de enfoques. Es de resaltar, por ejemplo, el énfasis que pone en el desarrollo profesional del profesor y la naturaleza más democrática y cooperativa de la relación entre profesor y supervisor. El hecho, asimismo, de ser una supervisión *in situ* garantiza la significación práctica que para el profesor puede tener este tipo de procesos.

No puede desconocerse, tampoco, su impronta de seriedad y rigor puesta de manifiesto en una estructurada planificación de cada una de las fases o etapas del trabajo a realizar por el profesor y de la relación de éste con el supervisor.

La finalidad más importante de una conferencia (reunión) de supervisión es la de favorecer el desarrollo de la práctica pedagógica del profesor en un sentido efectivo. Tienen, a la vez, una función diagnóstica y prescriptiva y se proponen la mejora de la calidad de la educación en la escuela. Shinkfield y Stufflebeam (1995: 202), son partidarios de que todas las conferencias de supervisión instructivas tengan un objetivo básico. Esto no significa que pueda haber otros adicionales, pero nunca se deberían incluir objetivos divergentes del objetivo básico. Por ello, distinguen hasta cinco tipos de conferencias (reuniones) de supervisión, a saber:

- «A. El observador examina con el profesor los aspectos positivos de la lección vinculados con la conducta del profesor, que a su vez repercute en la conducta del alumno; categoriza las conductas; y proporciona la explicación que refuerce esa conducta..
- B. El observador ayuda al profesor a desarrollar una gama más amplia de actuaciones docentes eficaces.
- C. El profesor identifica los aspectos de la clase que han sido satisfactorios y en colaboración con el observador planifica estrategias para eliminar los no satisfactorios.
- D. El observador identifica y cuestiona actuaciones concretas, y si es oportuno sugiere alternativas para aquellos aspectos menos satisfactorios de la clase que pueden haber pasado desapercibidos para el profesor.
- E. El observador analiza la lección de un profesor muy competente que es quien luego establece los pasos a seguir para lograr un mayor enriquecimiento profesional».

No descubrimos nada nuevo si decimos que es una modalidad de supervisión de enfoque personalizado y de asesoramiento individual. Este aspecto es sumamente positivo en la medida que permite un conocimiento más profundo de las características de la enseñanza en un aula concreta y de la actuación singular de cada profesor. Pero, en un mismo orden de cosas, tampoco es menos verdad que las actuaciones de esta naturaleza se desin-

teresan de los problemas que afectan a toda la institución escolar, y desconocen los sistemas de carácter grupal y participativos con los que podría abordarse sistemática y cooperativamente las deficiencias institucionales de una o más escuelas.

Al centrar su atención en el análisis pormenorizado de la enseñanza en el aula, es también evidente que se dejan de lado toda una serie de aspectos relacionados con el diseño, desarrollo y evaluación del currículum. Esta perspectiva se deja en manos de otras instancias superiores, como la propia Administración, por lo que se desencadena un divorcio real y con escasas posibilidades de superación entre currículum y enseñanza.

Además de estas lagunas en que incurre la Supervisión Clínica, las críticas a la misma se han polarizado preferentemente en las limitaciones derivadas del estudio del comportamiento observable de la clase y en la supuesta tendencia a una relación jerárquica entre supervisor y profesor.

Con respecto a la primera de las críticas, es decir, a las limitaciones derivadas del estudio del comportamiento observable en la clase, podemos decir que van en la misma dirección que las que se han hecho, dentro de la investigación de la enseñanza contemporánea, al paradigma proceso-producto en lo que se refiere al movimiento de la «observación sistemática en el aula» (Ángel Pérez, 1985, y Lee S. Shulman, 1986).

Sintéticamente, cabe decir que en este tipo de investigaciones y actuaciones subyace una concepción reduccionista de la ciencia, dado que ésta se identifica exclusivamente con hallazgos <<objetivos>>, <<observables>> y <<cuantificables>>. Y, como afirman Tanner y Tanner (1987: 184) la recogida y categorización del comportamiento fácilmente observable deja una gran cantidad de aspectos muy significativos de los procesos de enseñanza-aprendizaje sin detectar, como por ejemplo los enfoques de carácter inferencial... (y) no se puede suponer sin más que el análisis interactivo profesor-estudiante en la clase vaya a conducir al logro de resultados de aprendizaje deseados, como por ejemplo el aumento de la curiosidad intelectual, el pensamiento crítico, la mejora de las aptitudes para la solución de problemas, el realce de los valores y procesos democráticos.

Las relaciones entre supervisor y profesor han sido analizadas, también, desde distintas perspectivas. La verdad es que los pioneros de la Supervisión Clínica, como Cogan, han repetido una y otra vez que las relaciones

que se dan entre profesor y supervisor son más del tipo de las que se establece entre colegas que de asesoramiento o de orientación. Con ello han querido resaltar que este tipo de relaciones tienen un carácter totalmente horizontal, sin que exista el más mínimo vestigio que permita hablar de una interacción vertical y jerárquica.

St. Maurice (1990) aborda, en un interesante y sugerente artículo, los regímenes de poder social y político que subyacen a las relaciones entre supervisor y profesor. Desde los planteamientos foucaultianos de los regímenes de poder y verdad, St. Maurice afirma con rotundidad que

«[el] profesor, a diferencia de otros profesionales como los médicos o los abogados, tiene poca autonomía para diseñar su actividad. La administración le dice lo que debe enseñar, cómo debe enseñarlo, y cómo se va a evaluar su trabajo. Lo paradójico de lo que promete la supervisión clínica es que la ‘mejora’ está definida totalmente por el régimen, y por tanto sólo puede servir al régimen. El régimen establece los términos en que se debe desarrollar el trabajo, hace que éste resulte duro, y después pide a los trabajadores que digan por qué los supervisores clínicos se sienten rechazados: a quienes viven debajo de un puente no les gustan los timadores que intentan venderles el puente».

En esta misma línea de evaluación del profesorado, hay que recoger las aportaciones de Richard Manatt (1983) con su propuesta de *Teacher Performance Evaluation (TPE)*. La evaluación del desempeño o rendimiento de los docentes va más allá de la Supervisión Clínica en la medida que queda un registro de los logros que permite la toma de decisiones sobre como un profesor va a desarrollar sus clases en el futuro. Es también diferente a la Supervisión Clínica porque analiza cómo imparten sus clases los profesores, atendiendo no sólo a los requisitos de la organización sino también tomando en cuenta los datos del rendimiento escolar. Tiene por finalidad incorporar mecanismos de control de calidad. A semejanza de la Supervisión Clínica resalta el currículum de aula, pero se diferencia de ella en la medida que compara los resultados de los profesores.

Shinkfield y Stufflebeam (1995: 272) señalan que este tipo de evaluación se centra en un análisis de las mediciones de los progresos que el profesor va logrando en la consecución de unas metas predeterminadas, o como los llama Manatt, objetivos del puesto de trabajo. No se sigue la línea tradicional de los enfoques proceso-producto (o entradas-resultados), sino que se centra en un proceso que basa, en gran medida, su éxito en que tanto el

profesor como el evaluador sepan qué es lo que constituye una actuación en el aula verdaderamente eficaz. También se insiste en una utilización efectiva y eficaz del tiempo.

En una guía para el monitor, que acompaña a una cinta de video utilizada para el perfeccionamiento del profesorado de plantilla, Manatt (1983: 3) indica que para que pueda tener éxito la TPE exige:

1. Escalas de valoración que se ajusten a criterios basados en la investigación sobre enseñanza eficaz.
2. Combinar los análisis de las clases con observaciones cualificadas.
3. Técnicas de tutoría y orientación que inciten a los profesores al cambio
4. Tomar medidas para que tanto en el procedimiento como en el contenido se siga el proceso debido, para proteger al evaluador y al profesor.

4. LA EVALUACIÓN DE LOS DOCENTES A TRAVÉS DE LOS RENDIMIENTOS DEL ALUMNADO

El rendimiento escolar de los alumnos ha sido tradicionalmente un criterio bastante utilizado para la evaluación de las escuelas y el profesorado. Su uso ha generado no pocas controversias y debates, y, como ya se señalaba en el Tema 1, quizá constituya la cuestión más polémica de la historia de la evaluación. Desde luego, no cabe duda, que es más que discutible sostener que el aprendizaje de los alumnos se debe exclusivamente a la competencia y prácticas pedagógicas de los profesores, aunque la influencia de éstos sea algo que tampoco nadie cuestiona. La multideterminación del rendimiento escolar ha sido puesto en evidencia por estudios e investigaciones que señalan la importancia de variables personales, de los propios alumnos, institucionales, referidas a las políticas educativas, a las características de los centros, familiares, como las vinculadas al origen social, expectativas culturales de los padres, etc. El profesorado es, por tanto, uno más de los elementos que contribuyen al aprendizaje de los alumnos, cuya responsabilidad sí es la de crear y poner en juego estrategias y oportunidades facilitadoras del crecimiento intelectual y autonomía personal de sus estudiantes.

Sea como fuere, y a pesar de todos los argumentos y evidencias aportadas, no puede desconocerse que desde los años 70 hasta hoy el rendimiento de los alumnos sigue utilizándose, cada vez con tecnologías más refinadas en el campo psicométrico y estadístico, para la evaluación del profesorado. De una manera indirecta, la publicación en los medios de comunicación de las puntuaciones obtenidas por los alumnos en los exámenes nacionales, caso de países como Francia, Inglaterra y EE.UU., o la divulgación de los estudios sobre la evaluación de los sistemas educativos nacionales, centrados la mayoría de ellos en los logros o resultados de los estudiantes, contribuyen a crear opinión sobre la calidad de los docentes.

De una manera directa, y partiendo de la más clásica tradición tyleriana, se han ido desarrollando procedimientos de evaluación del profesorado tomando el rendimiento de los alumnos desde distintas perspectivas y para diferentes finalidades. Así, por ejemplo, a finales de los 60 y principios de los 70 en la Universidad de California (UCLA) se realizaron distintas experiencias con tests de rendimiento (McNeil, Alkin y Klein, 1974; Pophan, 1968, 1971) para la evaluación de los docentes. La secuencia que se seguía, en estos casos, consistía en que los profesores tenían que trabajar con unos materiales instructivos específicamente diseñados para su aplicación en el aula y de acuerdo con unos objetivos comportamentales explícitos que los alumnos debían alcanzar al final del proceso. Las puntuaciones obtenidas y las diferencias existentes entre las logradas en el pretest y las alcanzadas en el postest eran consideradas como el referente o la medida de la eficacia de los profesores.

Glass (1997) analiza en un excelente trabajo las experiencias llevadas a cabo de evaluación del profesorado sobre la base del rendimiento escolar de los alumnos en seis distritos escolares. Su descripción y valoración de este sistema de evaluación del profesorado, asociado a los incrementos salariales de los docentes, las fundamenta en las entrevistas realizadas a personas que siguieron de cerca las experiencias y en el análisis de la documentación y publicaciones de los programas. Dichas experiencias consistieron en establecer unas determinadas compensaciones económicas (pago según méritos) para todos aquellos profesores que lograsen alcanzar ciertos objetivos de crecimiento de sus alumnos, aunque finalmente en el Programa Salarial según Méritos se contemplaron también bonificaciones para los docentes que mejoraran su asistencia a las aulas, que impartieran asignaturas de áreas de especialización con escasez de profesores y que enseñaran en centros con una elevada proporción de minorías.

Glass relata con detalle y claridad las distintas negociaciones que se llevaron a cabo para determinar los estándares y criterios de crecimiento de los alumnos, así como los procedimientos acordados para el reparto de las compensaciones. El rendimiento de los alumnos se obtenía a través de unas pruebas estandarizadas de lectura y matemáticas, en primaria, y de una prueba global de conocimientos en secundaria, estableciéndose como criterio de crecimiento que los primeros logaran un aprendizaje igual al del curso anterior añadiéndole el equivalente a dos meses más de enseñanza, mientras que en secundaria el criterio se determinó fijando el porcentaje de respuestas acertadas en la prueba. El procedimiento de reparto de compensaciones económicas no se hizo por cuotas, es decir en función de los resultados obtenidos por los alumnos de cada profesor, sino que todo los profesores de un mismo centro recibían las mismas bonificaciones si el 75% de los de una misma escuela lograban alcanzar los objetivos marcados.

En la valoración de las experiencias se señala que durante los primeros años la mayoría de la escuelas participantes pasaron con éxito las pruebas al alcanzar la mayoría de ellas los objetivos de crecimiento, aunque se observa que las compensaciones económicas a los profesores se debían más a la disminución de su absentismo, al hecho de que se bonificara la presencia de minorías en los centros y a la impartición de ciertas materias especializadas, que a lo que realmente representaba la cuantía de lo recibido por el progreso de los alumnos. Cosa que, por otra parte, era un objetivo no explícito de los responsables y administradores de los distritos, dado que el monto total de estos programas no superaba, en bastantes casos, el dinero que anualmente se desembolsaba para pagar las suplencias de los profesores en los centros.

Con el paso del tiempo, después de cinco o seis años de su inicio, el programa llegó a sus límites ya que cada vez era más difícil lograr unos estándares de progreso que anualmente se iban elevando, dándose casos de manipulación de resultados por directores y profesores. Glass (1997: 329-330) concluye que utilizar “los datos sobre rendimiento académico de los alumnos para evaluar profesorado:

1. Casi siempre se lleva a cabo a nivel escolar (o todos o ninguno de los profesores de una escuela son recompensados de igual forma) en vez de hacerse a nivel individual porque: a) no existen exámenes autorizados en la mayoría de las áreas del currículo de secundaria, ni tampoco para

los papeles más especializados que juegan los profesores de primaria; b) los profesores rechazan la noción de que deberían competir con sus colegas para obtener subidas, privilegios, etc.

2. Siempre se combina con otros criterios (tales como el absentismo o el trabajo extra) que resultan ser los verdaderos discriminadores entre quién es recompensado y quién no.
3. Es demasiado susceptible a la distorsión y a la manipulación para engendrar confianza en los datos; además, profesores y otros creen que ningún tipo de test o de análisis estadístico puede compararse con la dificultad de las tareas del profesor en la diversidad de circunstancias en las que trabaja.
4. Eleva los exámenes en sí a la categoría de objetivos del currículo, dificultando la distinción entre aprendizaje y rendimiento en los exámenes.
5. Es a menudo un acto administrativo simbólico llevado a cabo para asegurar al público que se valora el aprendizaje del niño y se persigue asiduamente.”

Con el desarrollo de las metodologías estadísticas, el rendimiento de los alumnos vuelve a cobrar de nuevo un inusitado protagonismo en la evaluación de los sistemas educativos, centros y profesores. Desde mediados de la década de los 80 en que surgen, los sistemas de medida de valor añadido han ido captando el interés de los responsables de las políticas educativas (Saunders, 1997, 1999), de profesores universitarios de campos psicopedagógicos (Thomas, 1998; Marchesi y Martín, 1999; Tiana, 1999), de estadísticos (Sanders y Horn, 1995), etc. Precisamente a estos últimos se debe el desarrollo del Sistema de Evaluación de Valor Añadido de Tennessee (TVAAS) que Shinkfield y Stufflebeam (1995) recogen en su libro como uno de los posibles modelos de evaluación del profesorado.

El TVAAS de Sanders y Horn ha sido desarrollado como consecuencia de los resultados obtenidos en distintos estudios sobre la influencia de los profesores en el rendimiento escolar de los estudiantes. McLean y Sanders (1984), en una investigación llevada a cabo durante tres años con alumnos y docentes en la que utilizaron el Test de Rendimiento de California y un sistema de análisis estadístico basado en la metodología de modelo mixto de Henderson, pusieron de manifiesto que: existían diferencias apreciables entre escuelas y profesores en relación a sus efectos sobre los indicadores

de aprendizaje de los alumnos; las estimaciones de los efectos de las escuelas y profesores tendían a mantenerse año a año; la influencia del profesorado era independiente del contexto o entorno del centro escolar; había una alta correlación entre la influencia del profesor, tal y como muestran los datos, y las evaluaciones subjetivas de los supervisores; los índices de progreso de los estudiantes no tenían relación alguna con los niveles de capacidad o rendimiento con los que se incorporaron a la escuela, Sanders y Horn (1995: 338).

Tomando como base estos datos e inspirándose en la tradición tyleriana de comparación entre objetivos y resultados, Sanders y Horn (1995: 341), el TVAAS puso en marcha un programa en el curso 1992-93 al objeto de recoger anualmente información detallada sobre la identificación de cada profesor y de los estudiantes a los que enseñaba; las materias que los profesores impartían a cada alumno y la proporción de tiempo que cada alumno pasaba con su profesor, si se daba trabajo en equipo o departamental junto con las partes de cada materia de las que los profesores son responsables; la asistencia de los alumnos a las clases de sus profesores (sólo se tienen en cuenta a los estudiantes que han asistido un mínimo de 150 días en el curso). Estos datos y las puntuaciones obtenidas por los alumnos en los tests es lo que permite, en periodos de tiempo de entre tres y cinco años, estimar el valor añadido que la influencia del profesor tiene en el rendimiento escolar de los estudiantes.

5. LOS PROCEDIMIENTOS DE LA EVALUACIÓN DE PROFESORES

Como hemos tenido oportunidad de ver en el tema anterior, la evaluación de los docentes no puede desvincularse de los propósitos que se persiguen con la misma, ni de las concepciones que se mantengan sobre la enseñanza y la profesión docente. No obstante, en cualquiera de los casos el éxito de la valoración de los profesores, ya se haga ésta a efectos de su desarrollo profesional o de cara a identificar su grado de competencia, depende en gran medida de los procedimientos empleados. En general, en la literatura sobre el tema es bastante usual encontrarse con referencias que ponen en cuestión los modelos de evaluación del profesorado que hasta ahora se han venido utilizando.

La insatisfacción y el descontento con la evaluación de los docentes se centran en distintos motivos, según señalan Scriven, Wheeler y Haertel

(1992, 1993), aunque los más frecuentes en cuanto a las limitaciones y carencias señaladas son las de no estar basadas en normativas y criterios claros, ni centradas en criterios defendibles, no ser fiables y creíbles ni tampoco sensibles a situaciones docentes muy concretas, carecer de prestigio, y caracterizarse por su tendenciosidad, superficialidad y capacidad para crear desmoralización.

Ante este panorama, los esfuerzos desplegados en EE.UU. durante toda la década de los años 90 han estado dirigidos a la mejora de los sistemas de evaluación del profesorado, potenciándose publicaciones como la de los *Estándares de Evaluación de Personal* (Stufflebeam, 2001), o de determinadas guías profesionales (Mckenna, Nevo, Stufflebean y Tomas, 1998) o de metamanuales (Farland y Gullickson, 1998), de cara a evaluar, desarrollar y mejorar dichos sistemas no solamente desde una perspectiva técnica sino teniendo en cuenta también sus consecuencias legales, dada la importancia de éstas en los procedimientos de empleo norteamericanos.

Las tres publicaciones citadas son un reflejo de la sensibilidad existente por contar con modelos y procedimientos de evaluación docentes que reúnan unas mínimas garantías. Las dos primeras tratan de articular sus propuestas en torno a los cuatro atributos básicos —propiedad, utilidad, viabilidad y precisión— sobre los que se han desarrollado los 21 estándares para la evaluación de personal.

Los estándares son principios que han sido acordados por los profesionales que vienen trabajando en el campo de la evaluación a los efectos de registrar, estimar y/o medir la valía y calidad de una evaluación. Su función es doble ya que, por un lado, se constituye en una guía que es necesario observar para acometer la evaluación del profesorado, en nuestro caso, y, por otro, es un marco de referencia inevitable para valorar cómo se ha llevado a cabo el propio proceso de evaluación, es decir, pueden ser herramientas de metaevaluación —evaluación de la evaluación—.

A continuación presentamos los 21 estándares para la evaluación del personal que, seleccionados por un Comité Conjunto de dieciséis miembros de Estados Unidos, fueron elaborados y redactados por dicho Comité en su formato actual, una vez revisados por expertos y profesionales de la evaluación en diferentes contextos nacionales e internacionales. En los Cuadros del 8.1 al 8.4 se recogen los estándares de propiedad, utilidad, viabilidad y precisión.

Cuadro 8.1. Estándares de propiedad para la evaluación del personal.

Orientación hacia el servicio	Las evaluaciones de los profesores deberían promover principios pedagógicos bien fundamentados, el cumplimiento de los fines institucionales y el desempeño eficaz de las responsabilidades profesionales con el fin que se cubran las necesidades educativas de los alumnos, la comunidad y la sociedad.
Orientaciones formales de evaluación	Para que las evaluaciones sean consistentes, equitativas y conformes con la legislación correspondiente y con los códigos éticos, las orientaciones para la evaluación de los profesores deberían registrarse en declaraciones políticas o en acuerdos negociados.
Conflicto de intereses	Los conflictos de intereses deberán detectarse y abordarse abierta y honestamente para que no pongan en peligro el proceso y los resultados de la evaluación.
Acceso a los informes de evaluación del personal	El acceso a los informes de evaluación de los profesores deberían limitarse a aquellas personas con un interés legítimo de revisar y actualizar los informes, garantizando de esta manera un uso apropiado de la información.
Interacciones con los evaluadores	La evaluación debería tratar a los evaluados de una forma profesional de manera que sus autoestima, motivación, actitud profesional y actitud hacia la evaluación mejoren o por lo menos no se vean perjudicadas

Los estándares de propiedad hacen referencia a la orientación de los principios y los procedimientos de carácter ético y legal con los que debe llevarse a cabo la evaluación del profesorado para proteger sus derechos. En ellos subyace la importancia de que la evaluación contemple la contribución de los docentes al logro de los propósitos que justifican la existencia de las instituciones escolares y las necesidades de los alumnos.

Los estándares de utilidad, por su parte, se refieren a los requisitos con los que se debería realizar la evaluación. En ellos se considera la orientación de las evaluaciones para que sean informativas, oportunas en el tiempo e influyentes.

Cuadro 8.2. Estándares de utilidad para la evaluación de personal.

Orientación constructiva	Las evaluaciones deberían ser constructivas para que ayuden a los centros a desarrollar los recursos humanos y animar y ayudar a los evaluados a prestar un servicio excelente.
Usos definidos	Para que la evaluación pueda abordar las cuestiones adecuadas a los usuarios y todos los usos pretendidos de una evaluación de los profesores deberían ser identificados.

Credibilidad del evaluador	La evaluación debería ser dirigida y ejecutada por personas con las cualificaciones, destrezas y autoridad necesarias, y los evaluadores comportarse profesionalmente para que los informes de evaluación sean respetados y utilizados.
Informes funcionales	Los informes deberían ser claros, oportunos, precisos y pertinentes para que sean de valor práctico para el evaluado y otros destinatarios pertinentes.
Seguimiento e impacto	Debería hacerse un seguimiento de la evaluación para que se ayude a los usuarios y evaluados a comprender los resultados y a adoptar las acciones adecuadas.

Los estándares de viabilidad están relacionados con las condiciones necesarias para que se lleve a efecto la evaluación de una manera asequible en cuanto a sus procedimientos, a la colaboración de las partes interesadas y a la economía de tiempos y recursos.

Cuadro 8.3. Estándares de viabilidad para la evaluación de personal.

Procedimientos prácticos	Los procedimientos deberían ser planificados y llevados a cabo de forma que produzcan la información necesaria al tiempo que reducen el costo y las molestias.
Viabilidad política	La evaluación debería desarrollarse y supervisarse en colaboración para que todas las partes implicadas se involucren de forma constructiva en hacer que funcione el sistema.
Viabilidad fiscal	Se deberían asignar tiempo y recursos adecuados para las actividades de evaluación de manera que ésta pueda ser aplicada eficaz y eficientemente.

Los estándares de precisión, por último, se fijan en los procedimientos de carácter metodológico que deben presidir la evaluación, tratando de que la información obtenida esté bien fundamentada en cuanto a aspectos como la pertinencia y validez de la misma.

Cuadro 8.4. Estándares de precisión para la evaluación de personal.

Definición de funciones	La función, responsabilidades, objetivos de actuación y las cualificaciones necesarias del evaluado deberían ser definidas claramente para que el evaluador pueda determinar criterios de evaluación válidos.
Entorno de trabajo	Para que las influencias y limitaciones del entorno sobre la actuación docente sean recogidas en la evaluación, deberían identificarse, describirse y registrarse el contexto en el que trabaja el evaluado.
Documentación de procedimientos	Para que los evaluados y otros usuarios puedan evaluar los procedimientos reales en relación con los pretendidos, los procedimientos de evaluación seguidos en realidad deberían ser documentados.

Medición válida	Los procedimientos de medición deberían ser elegidos o desarrollados y aplicados sobre la base de las funciones descritas y el uso pretendido para que las inferencias que tienen relación con el evaluado sean válidas y precisas.
Medición fiable	Los procedimientos de medición deberían elegirse o desarrollarse para asegurar la fiabilidad con el propósito de que la información recogida proporcione indicaciones consistentes y precisas de la actuación del evaluado.
Control sistemático de datos	La información utilizada en la evaluación debería guardarse en lugar seguro.
Control del sesgo	Para que las cualificaciones del evaluado sean justamente valoradas, el proceso de evaluación debería garantizar la inexistencia de sesgos.
Seguimiento de los sistemas de evaluación	Los sistemas de evaluación de los profesores deberían revisarse periódica y sistemáticamente de forma que pudieran hacerse las modificaciones oportunas.

6. A MODO DE CONCLUSIÓN

La evaluación del desempeño docente trata de valorar la trayectoria del profesorado a lo largo del tiempo, de modo que se disponga periódicamente de los logros y deficiencias de su práctica profesional. Es obvio que esta evaluación tiene una orientación de control que es, precisamente, el principal inconveniente hasta ahora para su implantación y desarrollo en general y que, en particular, en nuestro país ha hecho que las propuestas contempladas en distintas leyes desde mediados de los años noventa hasta hoy permanezcan encalladas. Es cierto que llegar a acuerdos en un tema como éste no es nada fácil, pero también es verdad que no se puede continuar por más tiempo sin disponer de información que permita a la comunidad educativa y a la sociedad conocer hasta qué punto y de qué manera se hace realidad el principio de equidad de todos los alumnos en su derecho a la educación. A día de hoy, la información de lo que ocurre en las aulas y centros tiene demasiadas zonas de penumbra que hay que iluminar con la evaluación.

Por otro lado, además de la orientación de control también la evaluación del desempeño docente puede ser una palanca poderosa para la mejora y el desarrollo profesional. Una y otra orientación no son incompatibles y bien gestionadas las experiencias de evaluación entre iguales que veremos

en el tema 9, especialmente las vinculadas a la investigación acción y la autoevaluación de la tradición interpretativa, práctica-deliberativa, pueden ser cauces para la socialización y crecimiento profesional del profesorado en todos los retos y desafíos que los profesores hemos de asumir en nuestros desempeños docentes.

Lo que si se ha de tomar conciencia, ya sea en la orientación de control o de desarrollo profesional, es que las propuestas y proyectos de evaluación son el fiel reflejo de las concepciones que en las mismas subyacen sobre qué es ser un profesor y, consecuentemente, cuáles han de sus competencias básicas. Aquí es donde residen las diferencias entre las distintas propuestas que desde unas u otras posiciones y paradigmas se hacen. En todos los casos las posiciones serán legítimas si en la propia evaluación se tienen en cuenta todo un conjunto de principios y normas de carácter ético y moral.

BIBLIOGRAFÍA

- BOLÍVAR, A. (2008). «Evaluación de la práctica docente. Una revisión desde España», en *Revista Iberoamericana de Evaluación Educativa*, vol. 1, n. 2, páginas 57-74. <http://www.rinace.net/riee/numeros/vol1-num2/art4.pdf> . Consultado el 1 de marzo de 2013
- COGAN, M. (1973): *Clinical Supervision*. Boston, Houghton Mifflin.
- DANIELSON, CH. (2011). *Competencias docentes: desarrollo, apoyo y evaluación*. Santiago de Chile, Programa de Promoción de la reforma Educativa en América Latina y el Caribe (PREAL), serie Documentos n. 51.
- ESCUDERO MUÑOZ, J.M. (1998): «La Evaluación Docente: Naturaleza y propósitos de la evaluación del profesorado», en J.M. Escudero Muñoz, M^a.T. González y J. del Cerro, (eds.): *Evaluación de programas, centros y profesores*. Murcia, ICE Universidad de Murcia-Diego Marín Edit., pp. 107-119.
- FARLAND, DALE S. y GULLICKSON, ARLEN R. (1998): *Guía para la elaboración de un manual de evaluación del profesorado: un metamanual*. Bilbao, Mensajero.
- FERRER, G. (2006). *Estándares en educación. Implicancias para su aplicación en América Latina*. Perú, PREAL, disponible en <http://www.preal.org> , Consultado 20-06-2013

- GLANZ, J. y BEHAR-HORENSTEIN, L. S. (2.000): *Paradigm debates in curriculum and supervision: modern and postmodern perspectives*. Westport; London: Bergin & Garvey.
- GLASS, G. V. (1997): «Utilización de resultados de exámenes de alumnos para evaluar profesorado», en J. Millman, y L. Darling-Hammond (eds.): *Manual para la evaluación del profesorado*. Madrid, La Muralla, pp. 316-331.
- HOUSE, E. R. y LAPAN, S. D. (1989): «Teacher appraisal», en H. Simons y J. Elliott (eds.): *Rethinking appraisal and assessment*. Milton Keynes, Open University Press, pp. 55-63.
- INGVARSON, L. y KLEINHENZ, E. (2006). «Estándares profesionales de práctica y su importancia para la enseñanza», en *Revista de Educación*, 340, mayo-agosto, pp. 265-295.
- MANATT, R.P. (1983): *Competent evaluators of teaching: Their knowledge, skills, attitude* (Occasional Paper 83-1). Ames, IA: School Improvement Model Project. Iowa State University.
- MAURICE, H. ST. (1990): «Supervisión clínica y poder: regímenes de la gestión docente», en T. S. Popkewitz (ed.): *Formación del profesorado. Tradición, teoría, práctica*. Valencia, Servei de Publicacions de la Universitat de València.
- MAYOR, RUIZ, C. (1997). «La supervisión clínica como estrategia de asesoramiento», en C. Marcelo García y J. López Yáñez (coords.). *Asesoramiento curricular y organizativo en educación*. Barcelona, Ariel Educación, pp. 361-379.
- MURILLO, F. J. y HERNÁNDEZ, R. (2012). «Hacia una Evaluación del Desempeño Docente no Universitario en España. Lecciones Aprendidas de la Evaluación de Directores/as Escolares», en *Revista Iberoamericana de Evaluación Educativa*, 5 (1e), pp. 363-373. Disponible en http://www.rinace.net/riee/numeros/vol5-num1_e/art28.pdf Consultado el 20 de marzo de 2013.
- PÉREZ GÓMEZ, A. (1.985): «Paradigmas contemporáneos de investigación didáctica», en J. Gimeno Sacristán y A. Pérez Gómez (eds.): *La enseñanza: su teoría y su práctica*. Madrid, Akal (2ª edic.), pp. 95-138.
- SANDERS, W. L. AND HORN, S. P. (1995): «The Tennessee Value-Added Assessment System (TVAAS): Mixed Model Methodology in Educational Assessment», en A. Shinkfield y D. L. Stufflebeam (eds): *Teacher Evaluation: Guide to Effective Practice*. Boston, Kluwer Academic Publisher, pp. 337-350.

- SAUNDERS, L. (1997): «Value-added principles, practice and ethical considerations», en A. Harris, N. Bennett y M. Preedy (eds.): *Organizational Effectiveness and Improvement in Education*. Buckingham: Open University Press.
- (1999): «Work in <<value-added>> analysis of schools' performance at the National Foundation for Educational Research (NFER)», *Infancia y Aprendizaje*, n. 85, pp. 19-32.
- SCRIVEN, M. (1988): «Duties-based teacher evaluation», *Journal of Personnel Evaluation in Education*, I, pp. 9-23.
- (1990): «Can research-based teacher evaluation be saved», *Journal of Personnel Evaluation in Education*, 4(1), pp. 19-32.
- (1991): *Evaluation Thesaurus*. Newbury Park, Sage Publications, first printing.
- (1994): «Duties of the Teacher», *Journal of Personnel Evaluation in Education*, 8, pp. 151-184.
- (1997): «Selección del profesorado», en J. Millman y L. Darling-Hammond (dirs.): *Manual para la evaluación del profesorado*: Madrid, La Muralla, pp. 105-146.
- SERGIOVANNI, T. J. (1982). *Supervision of Teaching*. Alexandria, Association for Supervision and Curriculum Development (ASCD).
- SERGIOVANNI, T.J. y STARRAT, R. J. (1983): *Supervisión: Human Perspectives*. New York, MacGraw Hill Book, 3rd. edn.
- SHINKFIELD, A. y STUFFLEBEAM, D. L. (eds): *Teacher Evaluation: Guide to Effective Practice*. Boston, Kluwer Academic Publisher, pp. 337-350.
- SHULMAN, L.S. (1986): «Paradigma and Research Programs in the Study of Teaching: A Contemporary Perspectives», en M.C. Wittrock: *Handbook of Research on Teaching, Third Edition*. New York, Macmillan.
- STUFFLEBEAM, D. L. (2001): *Estándares de Evaluación de Personal. Cómo evaluar sistemas de evaluación de educadores*. Bilbao, Ediciones Mensajero.
- TANNER, D. y TANNER, L. W. (1987): *Supervision in Education: Problems and Practices*. New York, MacMillan.
- TEJEDOR, F. J. (2012). «Evaluación del desempeño docente», en *Revista Iberoamericana de Evaluación Educativa*, Volumen 5, n. 1e, páginas 318-327. Disponible en http://www.rinace.net/riee/numeros/vol5-num1_e/art24.pdf Consultado el 1 de marzo de 2013.

Tema 9

La evaluación docente entre iguales

Introducción

1. La autoevaluación individual de los profesores
2. El profesorado y las diversas posibilidades de la autoevaluación docente
3. La autoevaluación docente y la investigación-acción
4. El uso del portafolio para la autoevaluación del profesorado
5. A modo de conclusión

Bibliografía

TEMA 9 LA EVALUACIÓN DOCENTE ENTRE IGUALES

Eustaquio Martín Rodríguez.

UNED

INTRODUCCIÓN

La autoevaluación del profesorado se presenta, del mismo modo que ocurría en el ámbito más amplio de la escuela, como una propuesta alternativa a las modalidades de evaluación externa. Su énfasis formativo y su vinculación expresa con el desarrollo profesional del profesorado y la mejora de la escuela hacen que sea un enfoque más apreciado por los docentes para la valoración de sus prácticas pedagógicas.

Bajo el rótulo de autoevaluación del profesorado se pueden encontrar, sin embargo, en la literatura pedagógica enfoques, experiencias y procedimientos muy diferentes entre sí, (James, 1987; Wilson, 1992; Laffitte, 1994; Solabarrieta, 1996). Es verdad, que todos ellos defienden la autoevaluación como una propuesta de evaluación formativa dado que sus propósitos siempre se encaminan a potenciar el crecimiento y la mejora profesional del profesorado. Pero no es menos cierto, también, que los supuestos y concepciones de la profesión docente y la enseñanza que dejan traslucir se identifican, en ocasiones, con las competencias y destrezas que debe poseer un profesional técnico en función de los objetivos extrínsecos que ha de perseguir la enseñanza; en otros casos, se proyectan en el compromiso y responsabilidad individual de los profesores en el marco concreto de su actividad en el aula; y en algunas ocasiones, las menos, los docentes son considerados como prácticos reflexivos que investigan sobre su enseñanza ya que la enseñanza en sí es un proceso de investigación-acción.

Paralelamente la metodología así como los métodos y técnicas de evaluación se corresponden, normalmente, con los supuestos y concepciones planteadas. La capacidad de los profesores para determinar los aspectos o elementos que van a ser objeto de evaluación, desde la problematización de la enseñanza como un acto puramente técnico o como prácticas morales; las decisiones sobre las fuentes, técnicas y momentos de la evaluación; la vinculación entre los resultados de la evaluación y las formas y modos de

intervención consecuentes; etc., son, entre otros, algunos de los aspectos que diferencian a unas y otras formas de concebir y acometer los procesos de autoevaluación.

1. LA AUTOEVALUACIÓN INDIVIDUAL DE LOS PROFESORES

Barber (1997: 315) defiende la autoevaluación individual de los profesores y la autoevaluación en que participan colegas, concebida ésta como sistema híbrido que combina técnicas de autoevaluación y de revisión realizada por otros profesores, como modalidades de evaluación formativa que pueden suponer una alternativa a las formas punitivas de evaluación sumativa del profesorado. A su juicio, tres son las fases necesarias para llevar a cabo una autoevaluación eficaz, a saber:

«Primero, debe identificarse correctamente el comportamiento pedagógico actual. Los profesores deben ser conscientes de lo que se está haciendo, tanto verbal como no-verbalmente, en el aula. Muchos profesores perciben su rendimiento de un modo diferente a cómo es en realidad. Así, para que tenga lugar una mejora, es de vital importancia que durante la autoevaluación sean conscientes de cuál es su verdadero comportamiento. Segunda, las áreas problemáticas han de ser identificadas y mejoradas, deben mantenerse los puntos fuertes y ponerse en práctica nuevos comportamientos. Tercero, debería someterse los nuevos comportamientos a una evaluación adicional para determinar su eficacia, y comenzar todo el proceso de nuevo. De este modo, la autoevaluación se convierte en un proceso continuo en vez de en un evento único e irrepetible».

A tenor de las palabras de Barber parece claro su defensa de la autoevaluación del profesorado como una forma de evaluación individual de cada docente que juega, sobre todo, un papel diagnóstico de su comportamiento pedagógico y que como proceso debe proseguir con la valoración de la eficacia de las actuaciones emprendidas para solventar y mejorar las áreas problemáticas detectadas. Globalmente considerada, la propuesta de Barber es consecuente con los propósitos formativos que declara y resalta la autoevaluación como un proceso continuo de análisis y valoración. Mucho más problemático es su definición de la autoevaluación como una forma de evaluación exclusivamente individual y de que identifique el desempeño profesional de los profesores con sus comportamientos pedagógicos en el aula.

En efecto, aun cuando el ensayo de autoevaluación de Barber se centra en general en los aspectos metodológicos y, especialmente, en un repertorio de métodos, técnicas e instrumentos, y no entra a analizar cuáles serían los componentes o elementos sustantivos que vertebran el comportamiento pedagógico de los profesores, sí es posible a través de aquéllos identificar cuál es su posición con respecto a éstos.

La mayor parte del repertorio técnico-instrumental recogido por Barber (1997: 304-312) se basa en técnicas que como las grabaciones de audio y vídeo intentan, al estilo de los métodos de microenseñanza o de la planificación y enseñanza de un breve segmento de clase, registrar los comportamientos de los profesores que *se desvían de la norma* o que, por el contrario, son los que se corresponden con los deseados. Las técnicas de escalas de valoración u hojas de autclasificación, los métodos de *materiales de autoestudio* a través de textos programados para valorar y analizar los estilos de enseñanza, o los de *modelaje* que consisten en la observación de modelos de una enseñanza de alta calidad con objeto de imitarla, son otros tantos ejemplos de formas de *autoevaluación* que tienen como centro un comportamiento pedagógico de los docentes articulado sobre competencias, destrezas y habilidades que se consideran imprescindibles para una buena enseñanza. Y que, además, se proyectan hasta en el uso de técnicas más flexibles o abiertas como en los llamados *informes elaborados por el propio profesor* en los que también debe responderse a una serie de preguntas relacionadas con la instrucción.

Por otra parte, la técnica de sistema híbrido que se recoge en el ensayo que venimos comentando (Barber, 1997: 311-312) es muy similar a los procedimientos de supervisión clínica, si tenemos en cuenta los pasos que se siguen, a saber:

Paso 1: Identificación de propósitos, objetivos o intenciones

Este paso lo realiza el profesor en solitario o con la ayuda del compañero u otro observador objetivo.

Paso 2: Preparación de un plan de enseñanza

Este paso casi siempre lo lleva a cabo el profesor en solitario. Aquí, el profesor identifica el dónde, el qué, el cuándo y el por qué, así como técnicas y estrategias de las clases que ha de dar.

Paso 3: Reunión previo a la observación

Este paso lo llevan a cabo el profesor y el colega que le ayuda a revisar su trabajo. Su propósito es que el profesor pueda identificar claramente el comportamiento que debe observar aquel, así como ha de realizarse el registro de datos y qué condiciones de confidencialidad deben hallarse presentes.

Paso 4: Dar la clase, registrar comportamientos, obtener material acordado en la reunión previa

En casi todos los casos, este paso utiliza grabaciones de la instrucción de modo que el profesor pueda observar su trabajo y realizar su autoevaluación. Asimismo, permite que el colega observe el trabajo del profesor evaluado como lo desee.

Paso 5: Autoevaluación del comportamiento pedagógico

En este paso, el profesor utiliza una o más técnicas de autoevaluación (normalmente vídeo, hojas de autoclasificación, entrevistas con estudiantes) para determinar la calidad de la enseñanza y también si los objetivos del paso uno se han cumplido: el profesor escribe entonces sus hallazgos y conclusiones.

Paso 6. Provisión de datos al colega que realiza la revisión

En este paso, el colega que realiza la revisión recibe todos los datos que se acordó recoger en la reunión previa a la observación, paso tres, así como los hallazgos y conclusiones del profesor durante el paso cinco. La tarea del compañero consiste en determinar el grado de corrección de los hallazgos y conclusiones del profesor a partir de los datos proporcionados. En general, la función del compañero que realiza la revisión *no* consiste en determinar la calidad de la enseñanza. Los hallazgos y conclusiones de éste se escriben y se entregan al profesor.

Paso 7: Reunión posterior a la observación

Este paso implica una reunión *vis à vis* (o telefónica) entre el profesor y el colega encargado de llevar a cabo la revisión con objeto de discutir aquellas discrepancias que pudieran darse entre los hallazgos y conclusiones de uno y otro.

Paso 8: Resolución de discrepancias

Este paso lo lleva a cabo el profesor para determinar si es necesario realizar modificaciones en sus técnicas, estrategias, procedimientos y materiales después de analizar las discrepancias existentes entre sus conclusiones y hallazgos y los de su colega. Este paso también podría considerarse el principio de otra ronda de pasos, desde el número uno al presente.

Los pasos de este sistema híbrido, resultado de la combinación de la autoevaluación y la revisión realizada por un colega, se diferencia de la supervisión clínica, a nuestro juicio, únicamente por las distintas figuras profesionales que intervienen en una y otra, y sobre todo porque en esta última, como ya señalamos en su momento, la presencia de un supervisor imprime un sentido jerárquico a la revisión por lo que cabría cuestionar si no pervierte el sentido de lo que debe ser una autoevaluación y evaluación entre iguales.

No obstante, también nos surgen muchas dudas en torno a si el propósito de estas autoevaluaciones es realmente formativo o de desarrollo profesional, ya que en el caso de la evaluación de los profesores en ejercicio en la que se aborda los procedimientos y estrategias utilizados, Barber (1977: 312) señala expresamente que se consigue mediante los procesos de autoevaluación y revisión siguientes:

«a) el profesorado que ocupa una plaza en propiedad lleva a cabo una autoevaluación anual y es objeto de revisión periódica; b) éste es responsable de mantener informes actualizado que se utilizarán durante la revisión (sumativa) con objeto de mostrar sus esfuerzos para mejorar su enseñanza; c) la revisión la lleva a cabo el supervisor al caso y tiene lugar al menos una vez cada tres años; d) el proceso implica calibrar el rendimiento respecto de obligaciones previamente identificadas».

2. EL PROFESORADO Y LAS DIVERSAS POSIBILIDADES DE LA AUTOEVALUACIÓN DOCENTE

En otra publicación traducida al castellano, Peter Airasian y Arlen Gullickson (1999: 12) presentan un conjunto de herramientas para la autoevaluación del profesorado. Es, por ejemplo, muy significativo que

desde el principio justifican el contenido y título de su libro señalando que las soluciones propuestas deben verse

«como algo que debería ser ampliado, retocado y adaptado por los profesores a medida que encuentren sus propias ‘herramientas’ que les ayuden a entender y mejorar su práctica. Por esta razón no denominamos a este paquete ‘Manual de Autoevaluación’ porque implicaría cerrazón y rutinización. Además, tal denominación degradaría el proceso mismo que pretendemos alentar, la **autoevaluación**».

Esta declaración inicial es consecuente con los principios defendidos por los propios autores, en el sentido de que la autoevaluación debe ser una propuesta asumida como una práctica general en todo el centro e integrada en la estructura escolar del mismo; de carácter voluntario, es decir autogenerada y no impuesta; apoyada en cuanto a las necesidades de formación y recursos que los docentes puedan necesitar para iniciar y proseguir un proceso de esta naturaleza; llevada a cabo, por un lado, en un ambiente que facilite la imparcialidad y reduzca la subjetividad y, por otro, con garantías suficientes para la protección de los derechos de los profesores; y asegurada en cuanto a su continuidad e integridad.

Airasian y Gullickson (1999: 13) defienden una modalidad de autoevaluación en la que resaltan que el profesorado es el sujeto y objeto de su propio proceso de evaluación, con plena capacidad para decidir los criterios y estándares con los que valorar sus principios, conocimientos, destrezas y eficacia. Como ellos mismos afirman: *la autoevaluación del profesor/a es evaluación del profesor, por y para el profesor*. Vinculan la autoevaluación con aquel tipo de toma de decisiones que conduzcan al desarrollo profesional y a la mejora continua de los profesores. Y en su publicación (pp. 28-29) proporcionan ocho estrategias generales de autoevaluación:

1. **Herramientas de autorreflexión**, que están diseñadas para que las complete el profesor lo que le permitirá evaluar la actuación en las áreas identificadas en las herramientas de auto-reflexión. Estas herramientas están generalmente en forma de una lista, escala de evaluación o cuestionario.
2. **Registro y análisis de medios** que utiliza procedimientos como casete de audio o cinta de video para grabar un ejemplo de la actuación del profesor para su consiguiente análisis por el profesor y/o sus colegas. La microenseñanza es un ejemplo de grabación de medios.

3. **Retroalimentación del alumnado**, que es similar a la herramienta de autorreflexión excepto en que son los alumnos en vez del profesor los que rellenan el cuestionario.
4. **Portafolios del profesorado**, que preparan los profesores para aportar datos sobre sus principios, conocimientos, destrezas o eficacia. El proceso de recoger y analizar partes del portafolio aporta otra forma o indicio de autoevaluación.
5. **Datos sobre la actuación del alumnado**, lo que incluye todas las producciones del alumnado que pueden utilizarse para ayudar al profesor a evaluar su propia eficacia instructiva. Los resultados de exámenes (estandarizados o elaborados por el profesor), relaciones, proyectos de clase y demás son ejemplos de datos sobre la actuación del alumnado.
6. **Observación externa de los colegas**, la cual requiere tener un compañero o colega que observe, evalúe y aporte sugerencias sobre un aspecto de la práctica docente.
7. **Redacción de diarios**, que requiere que el profesorado mantenga un registro de acontecimientos de clase o actividades y reflexione sobre ellas con objeto de reconocer problemas recurrentes, tópicos, cosas que han salido bien y las necesidades.
8. **Solución de problemas basada en el diálogo / compartir experiencias de forma conjunta y colegiada**, todas las cuales contribuyen a la colaboración entre el profesorado para discutir problemas comunes, compartir procedimientos y estrategias y comparar sus percepciones. La exposición a las ideas y la práctica de otros colegas es una potente estrategia para la reflexión del profesorado y el cambio.

Dichas estrategias, en la selección de áreas de evaluación, se basan en aspectos concretos de la enseñanza (relacionados con los principios, conocimiento, prácticas y resultados) en vez de en planteamientos generales, intentando recoger información sobre los aspectos que previamente han sido limitados y bien definidos.

La propuesta de estos autores, globalmente considerada, nos parece un avance con respecto a la que hemos tenido oportunidad de analizar anteriormente. Tiene, sin embargo, facetas criticables, aspectos ambiguos y, en algunos casos, argumentaciones un tanto confusas. No se sabe muy bien,

por ejemplo, si optan por una modalidad de autoevaluación individualizada, es decir en la que son los profesores individualmente quienes se autoevalúan, o de tipo grupal o colegiado, ya que en el texto hay abundantes referencias en ambos sentidos, aunque la sensación final es que es más personal que colectiva si nos atenemos, en última instancia, a que todos los ejemplos de herramientas recogidos se refieren a actividades del profesor en el aula¹¹. Algunas de las herramientas propuestas, como es el caso de una escala de valoración de principios (que los autores concretan en el aprendizaje en el aula), que se complementa con respuestas de opciones que van desde muy de acuerdo (MA) a muy en desacuerdo (MD), las afirmaciones que se hacen son las siguientes (pág.31):

1. La mayor parte de las sesiones de clase muestran una amplia variedad de habilidades y estilos de aprendizaje.
2. Para aprender mejor los alumnos deben participar activamente en el proceso de aprendizaje.
3. Los alumnos consiguen mejores resultados cuando conocen el objetivo, ven modelos que les muestran la práctica correcta y conocen los criterios por los que serán juzgados.
4. El profesorado debería determinar las normas de clase, los objetivos, los contenidos didácticos, métodos de enseñanza y las evaluaciones.
5. El principal objetivo de la enseñanza es ayudar a que el alumnado acumule información acerca de la asignatura.
6. La mayor parte de los alumnos/as pueden ser enseñados de forma parecida.

Vistas las afirmaciones, a la 4, por ejemplo, se le asignan puntuaciones decrecientes conforme el grado de acuerdo con este comportamiento sea mayor. La explicación de esta valoración es difícil de entender hasta tanto se comprueba que los autores, más tarde, señalan que una de las caracte-

¹¹ El único ejemplo de estrategia de diálogo colegiado sobre la práctica, se refiere a la elección por parte del profesor, en una especie de respuesta a un cuestionario no estructurado, de la estrategia de enseñanza que le gustaría perfeccionar, de la descripción de los aspectos de esa destreza concreta, de identificar a dos profesores del claustro que, a su juicio, son expertos en esa destreza y a la sugerencia de que hable con ellos o de que le observen una clase (pág. 50).

rísticas de la enseñanza es que los profesores procedan a una discusión de los objetivos, criterios y estándares con los alumnos antes de la evaluación. Modo de proceder que, a nuestro entender, se puede compartir dependiendo de la edad y madurez de los alumnos. No obstante la afirmación no se refiere sólo a la evaluación, sino también a la determinación de las normas de clase, los objetivos, contenidos didácticos y métodos de enseñanza que, por lo general, suelen ser cometidos propios de los docentes.

Este tipo de hechos, la simplicidad de algunas de las herramientas y la insistencia en que la autoevaluación tiene que centrarse en aspectos formales y tangibles suponen también una determinada opción en cuanto a sus planteamientos implícitos sobre la enseñanza y la profesión docente que, al menos, requerirían una cierta clarificación y análisis. Especialmente, cuando tratando de diferenciar la autoevaluación de la reflexión en la acción, de la investigación-acción, de la reflexión sobre la práctica, etc. no precisan, por una parte, que aunque para ellos este tipo de estudios y orientaciones no puedan considerarse una forma de autoevaluación si tienen, por el contrario, una posición definida sobre la enseñanza y los profesores, y, por otra, desconsideran como efímero e intrascendente el conocimiento tácito de los profesores a la hora de su evaluación.

Pero si la autoevaluación no se hace con una finalidad sumativa, es decir, con propósitos de rendición de cuentas, ¿por qué ha de centrarse en la actuación individual de cada profesor? ¿Acaso esta forma de proceder no contribuirá a reforzar una cultura institucional y profesional de carácter atomizado e individualista? Hargreaves (1993: 52) nos recuerda que las consecuencias de una interacción profesional fragmentada, esporádica y superficial, por parte del profesor, le conduce a «dudas sobre su propia competencia (ausencia de una realimentación positiva), resistencia a compartir ideas con los demás (miedos a exponerse a errores en su práctica), falta de voluntad de aceptar riesgos (miedo a juicios duros si cometen errores)».

3. LA AUTOEVALUACIÓN DOCENTE Y LA INVESTIGACIÓN-ACCIÓN

Parece claro que el profesor tiene que acometer y asumir personalmente el conocimiento de su práctica profesional, la identificación de sus habilidades docentes, la clarificación de las oportunidades de aprendizaje que es capaz de crear para sus alumnos, etc. Wilson (1992: 105), a este respecto,

señala la importancia de la autoevaluación del profesor como una estrategia para la mejora de la calidad de la enseñanza dentro de un contexto de apoyo a los docentes. Y precisa que, en este sentido, una “idea clave en el empleo de la investigación-acción, como base de un desarrollo profesional, es la de que los profesores poseen un inmenso depósito de conocimientos profesionales tácitos o no reconocidos que constituyen la plataforma sobre la que actúan..... El desarrollo profesional implica tornar conscientes tales conocimientos profesionales tácitos, aclarando la relación entre las acciones y sus efectos, para lograr una mayor congruencia entre ambos”.

De acuerdo con McKernan (1999: 34-38) cabe hablar de tres tipos diferentes de investigación-acción en tanto que vinculados con la visión técnico-científica, con la práctica deliberativa y con la educación crítica emancipadora. El primero de ellos, promovido inicialmente por Kurt Lewin (1947), utilizaba el método científico de resolución de problemas para estudiar los procesos sociales, introduciendo cambios y observando científicamente los efectos en dichos procesos. Como científico social, Lewin estaba convencido de la necesidad de formular leyes generales de la vida social, de la relación entre la teoría y los hechos, y de la operatividad del conocimiento práctico logrado a través de estudios de campo situacionales que condujeran a la acción y a las mejoras sociales. Concebía la investigación-acción como un proceso de toma de decisiones fundamentadas sobre ciclos sucesivos de análisis, reconocimiento, reconceptualización del problema, planificación, puesta en práctica de la acción social y evaluación de la efectividad de la acción.

A diferencia de esta propuesta, la investigación acción como práctica deliberativa, surgida desde el ámbito académico del Centre for Applied Research in Education (CARE) de la Universidad de East Anglia (Inglaterra), y la investigación acción crítica emancipadora, potenciada desde la Universidad de Deakin (Australia), presentan características bastantes distintas. En principio, ambas cuestionan el marco de la tradición científica racional o positivista para acogerse a los postulados de las tradiciones interpretativo-simbólica y crítica, propugnando el protagonismo del profesorado en los procesos de indagación e investigación que surgen directamente de los problemas que genera su práctica profesional en el aula y centro escolar. Asumen, por el contrario, que son las preguntas e interrogantes que emergen en el contexto de su desempeño docente los que necesitan investigarse situacionalmente para encontrar respuestas y soluciones

con los que fundamentar sus actuaciones y estrategias y, en definitiva, lograr un conocimiento práctico incardinado con su compromiso moral con la enseñanza.

John Elliott (1996: 131) caracteriza la investigación-acción basada en el profesor de la siguiente manera:

1. Se centra en el descubrimiento, aclaración y resolución de los problemas a los que se enfrentan los profesores para llevar a la práctica sus valores educativos. En cuanto forma de investigación, se trata de una ciencia práctica y moral, en vez de teórica y técnica.
2. Supone una reflexión simultánea sobre los medios y los fines. Como fines, los valores educativos se definen por las acciones concretas que selecciona el profesor como medio para realizarlos. Algunos valores se ponen en práctica en las interacciones del docente con los alumnos y no como resultado extrínseco de ellas. Las actividades de enseñanza constituyen interpretaciones prácticas de los valores. Por tanto, al reflexionar sobre la calidad de su enseñanza, el profesor debe hacerlo sobre los conceptos de valor que la configuran y moldean.
3. Es una práctica reflexiva. Como forma de autoevaluación o autovaloración, la investigación-acción no consiste sólo en que el profesor evalúe sus acciones desde cierto punto de vista, por ejemplo, el de la eficacia técnica. Ante todo, consiste en que el profesor evalúe las cualidades de su propio «yo» tal como se manifiestan en sus acciones. En esta perspectiva, esas acciones se conciben como prácticas morales más que como simples expresiones técnicas. En el contexto de una práctica moral, la autoevaluación supone un tipo determinado de autoreflexión: la reflexividad.
4. Integra la teoría dentro de la práctica. Las teorías *educativas* se consideran como sistemas de valores, ideas y creencias representadas no tanto en forma proposicional como de práctica. Estas teorías se desarrollan mediante el perfeccionamiento reflexivo de la práctica. El desarrollo de la teoría y el perfeccionamiento de la práctica no se consideran procesos independientes.
5. Supone el diálogo con los compañeros de profesión. En la medida en que los docentes tratan de poner en práctica sus valores profesionales en la acción mediante la investigación-acción, se hacen responsables

de los resultados ante sus compañeros. Esa responsabilidad se expresa en la elaboración de expedientes que documenten los cambios habidos en la práctica y los procesos de deliberación y reflexión que dan lugar a esos cambios.

Elliott (1996, 1998) documenta cómo las propuestas iniciales de evaluación del profesorado no universitario (primaria y secundaria) y universitario que se hicieron, a mediados de los años ochenta, en Inglaterra fueron reconducidas desde la cultura profesional docente y el legado que la investigación-acción aportó después de veinticinco años de experiencia.

En su formulación original, el Libro Blanco sobre la calidad de la enseñanza (Teaching Quality, DES, 1983) planteaba una evaluación como forma de supervisión y control jerárquico del trabajo de los profesores que cumplía funciones de gestión como encontrar motivos fundamentados para efectuar despidos, proporcionar una base racional para la redistribución del personal y el reconocimiento de méritos y la identificación de necesidades de formación. Por su parte, las organizaciones sindicales y asociaciones de profesores defendían que los objetivos primordiales de la evaluación deberían fijarse en el aprendizaje profesional centrado en el aula y en el desarrollo de la carrera docente.

Esta confrontación de posiciones no era sino el reflejo del choque entre las culturas *gerencial* y *profesional*. La cultura gerencial hace referencia a un determinado estilo de gestión de los responsables gubernamentales que, en el caso del profesorado, sustrae la capacidad de decisión de los docentes para controlar su ejercicio y su futuro laborales. Mientras que la profesional se relaciona con el acervo acumulado por los profesores en su ejercicio y desempeños docentes para hacer frente, por sí mismos, a los problemas de la enseñanza desde los procesos de reflexión y deliberación que han de poner en juego en el análisis de su propia práctica.

Ante estos dilemas, Elliott (1996: 120) defiende que las funciones *profesionales* y de *gestión*, o gerencial, sólo pueden reconciliarse mediante el desarrollo en la práctica de un modelo de evaluación del profesorado de dos niveles en el que se reconsidere aspectos tan importantes como quién evalúa a quién, la naturaleza del proceso de evaluación y los elementos que integran el proceso de evaluación. Las características iniciales de todos estos aspectos pueden verse a continuación.

Quién evalúa a quién: un supervisor inmediato del profesor, que puede ser el director u otro profesor experimentado designado por el director.

El proceso de evaluación: ha de ser de naturaleza continua. La evaluación no debe convertirse en una actividad burocrática o en un ejercicio escrito esporádico. Teniendo presente esto, creemos que la frecuencia de los evaluaciones formales que culminan en informes escritos deben variar según la etapa de carrera del profesor.

Los elementos del proceso de evaluación: son la autoevaluación, el diálogo de revisión con el evaluador, la observación a cargo del evaluador, la entrevista de evaluación, las objeciones contra los juicios de los evaluadores, la remisión de informes al director.

En relación con cada uno de estos aspectos, habría que señalar en primer lugar que en una evaluación entre iguales la selección de los evaluadores debería ser realizada y controlada por el grupo de profesores en conjunto y no de manera jerárquica, si realmente se desea respetar lo que significa la autoevaluación y la evaluación entre compañeros. Recuérdese lo que planteábamos, a este respecto, cuando analizábamos algunas propuestas de autoevaluación institucional y nos preguntábamos acerca de ¿en qué consiste el **auto** de la autoevaluación? ante experiencias y realizaciones en las que la autonomía de los centros y profesores se reducía exclusivamente a la decisión de cumplimentar o no los protocolos y/o cuestionarios de recogida de información.

Convendría no olvidar que al individualizarse y jerarquizarse la evaluación, la responsabilidad del profesorado tiene o toma un sentido vertical, la de la dirección, en vez de horizontal hacia los compañeros de profesión. Y, como ya vimos en su momento, ante la evaluación los docentes pueden asumir una responsabilidad moral, ante los alumnos y los padres, y una responsabilidad profesional, ante sí mismo y los compañeros, que tienen unas características totalmente distintas a la responsabilidad contractual, ante los empleadores o responsables políticos. Las dos primeras, en el modelo doble de evaluación, no supone obviar la tercera sino que, por el contrario, se reconoce las funciones legítimas de la evaluación a efectos de la dirección, sobre todo en lo que se refiere a la contribución del profesorado al desarrollo y mejora de la organización en todo lo que tiene que ver con los desafíos institucionales y profesionales de los quehaceres educativos.

Ahora bien, planteada de esta manera parece consecuente que la autoevaluación o evaluación entre iguales, incluidos los procesos de observación en el aula, sean llevados a cabo por un grupo de profesores en vez de por una sola persona, para evitar la deformación de las representaciones de la enseñanza que se construyen a través de la experiencia práctica de la relación con los alumnos en el aula. Porque, por lo general, la individualización de la evaluación tiende siempre a centrarse en los aspectos técnicos de la enseñanza, a un tipo de reflexión aislada sobre cuestiones estructuradas en sentido jerárquico, dejando de lado los aspectos morales, y reforzando su eficacia instrumental en el logro de determinados objetivos especificados a priori. Ello supone aceptar que la enseñanza es básicamente una tarea técnica, cuyo ejercicio profesional se puede valorar en términos de eficacia y consecuencias instrumentales, en vez de una práctica compleja cuyos elementos específicos deben evaluarse en relación con el conjunto de cometidos que, de naturaleza cualitativa, desarrollan los docentes.

Un modelo doble de evaluación, como expresamente afirma Elliott (1996: 132),

«en el que la investigación-acción constituya una faceta del proceso, evitará que la individualización y jerarquización derivadas de la subordinación a la entrevista formal de evaluación deformen la autoevaluación y la evaluación a cargo de los compañeros. Esa primera faceta constituye en sí misma un proceso de desarrollo profesional que facilite que los individuos descubran sus propias necesidades de aprendizaje y determinen la mejor manera de satisfacerlos. Los órganos de dirección tendrían el importante cometido de facilitar y apoyar ese proceso de investigación-acción en el nivel de la organización, por ejemplo, respecto a la provisión del tiempo para la reflexión, la observación de los compañeros y las reuniones de profesores».

En definitiva, el modelo doble de evaluación opta porque sean los profesores quienes lleven a cabo su autoevaluación a través de procesos en los que el diálogo, la deliberación y la observación entre colegas sean la base sobre la que se analice su práctica en las aulas, se detecten sus necesidades de aprendizaje profesional y se contemplen las propuestas de mejora y desarrollo docentes. Como modelo contempla que se consideren no sólo la dimensión técnica de la enseñanza, es decir, determinadas competencias, habilidades y conocimientos del profesor, sino también sus aspectos morales, esto es, los valores educativos que, proyectándose en sus acciones, definen sus identidades como profesional. Considera legítima la función de

evaluación de la dirección de las instituciones educativas, pero apuesta porque se materialice a través de la entrevista de evaluación, una vez que los docentes previamente hayan elaborado sus propios expedientes mediante la autoevaluación, el diálogo y la observación en el aula, y sólo a los efectos conocer la contribución del profesorado al quehacer institucional, así como de promover actuaciones de desarrollo profesional.

Los elementos del proceso de evaluación no deben tomarse de una manera aislada y de forma sucesiva, sino considerando que son etapas interrelacionadas ya que, por ejemplo, la fase inicial de autoevaluación se nutre del diálogo entre los profesores y de las sesiones de observación en las aulas. Y éstas permiten fundamentar los elementos que han de ser analizados en la etapa siguiente, la entrevista de evaluación.

4. EL USO DEL PORTAFOLIO PARA LA AUTOEVALUACIÓN DEL PROFESORADO

Una manera muy distinta de evaluar el rendimiento de los profesores, en relación con algunos otros ya analizados tanto en este como en otros temas, es centrarse en determinados aspectos del trabajo de los mismos. House y Lapan (1989: 62) señalan, como una fuente de evidencia de los desempeños docentes, al portafolio¹². Señalan sobre enfoque:

«... tiene un valor particular para el “profesor como artista”, que tiene la oportunidad de presentar un modelo de rendimiento único. Los profesores pueden empezar ya desde los inicios de su carrera a desarrollar ejemplos para poder ponerlos en su *portafolio*. La mejora y la sustitución de elementos del *portafolio* se convertirá en una parte importante en la tarea de desarrollarse y de superarse como profesor».

En este sentido, Martín-Kniep (2001) destaca que el portafolio sirve tanto para reconocer y certificar la idoneidad del profesorado, como por ser una herramienta de autoevaluación o un mecanismo para evaluar y supervisar docentes. Es decir, en sus usos el portafolio puede jugar un papel

¹² Es bastante frecuente encontrar en los libros y artículos sobre este tema, la utilización del término **portfolio**. En nuestro caso hemos optado por castellanizarlo y utilizar los términos **portafolio** o **portafolios**, tal y como se recoge en el *Diccionario del Español Actual*, de Manuel Seco y otros, pág. 3624, Volumen II (Madrid, Aguilar, Grupo Santillana de Ediciones, SA, 1999, Primera Reimpresión).

fundamental para seleccionar al profesorado y ser un elemento de referencia para abordar su desarrollo y crecimiento profesionales, además de constituirse en una forma de conocer qué está ocurriendo en los centros y cómo trabajan y cuál es la coordinación de los docentes en el interior de las aulas y en el marco de la institución escolar en su conjunto. En cualquiera de los casos, los docentes tienen la posibilidad de hacer uso de su autonomía profesional para decidir la selección de evidencias que pongan de manifiesto su trayectoria profesional.

Bird (1997: 350) comparte las posibilidades que ofrece el portafolio resaltando que es un espacio de trabajo y de desarrollo de los docentes. No obstante, también plantea:

«... con la obsesión que hoy existe con compensar, sancionar o despedir profesorado mediante el empleo de medios principalmente burocráticos, los portafolios tendrían un peso sumativo considerable —a falta de un apoyo formativo eficaz o de una tradición que ayude a los profesores a realizar unos exámenes detallados de su trabajo—. Bajo estas condiciones, cabría esperar que los portafolios pasasen de ser instrumentos de mejora a objetos de exhibición, <<un ejercicio de acumulación de papel>>, y una pérdida de tiempo tanto para profesores como para las personas encargadas de revisarlos».

Una vez advertidas sus posibilidades y limitaciones, sus ventajas e inconvenientes, sobre todo sus dilemas en torno al control versus desarrollo o mejora profesional del profesorado vamos a dejar a un lado sus usos burocrático-administrativos para fijarnos más en detalle en sus potencialidades de autoevaluación del trabajo docente en los planos personal, o de las biografías individuales, y grupal o de relación de intercambio y aprendizaje entre colegas.

Desde esta perspectiva es en la que ha venido trabajando Martín-Kniep (2001: 17) para formar al profesorado en la elaboración de sus propios portafolios, señalando que entre otras características:

«Constituyen una historia en desarrollo. Son fluidos, aunque pueden congelar un momento y un final definidos. Son los museos de nuestro trabajo y nuestro pensamiento, en los que se exhiben nuestros éxitos, experimentos y sueños. Los portafolios son espejos, aunque sean deformantes, de una realidad en evolución. Nos muestran lo que queremos ver y lo que querríamos no ver. Además de ser un reflejo del yo, los portafolios son un espejo profesional».

Siguiendo a nuestra autora, Martín-Kniep (2001: 18-19), los portafolios pueden elaborarse en los siguientes momentos, dinámicas y situaciones:

- En distintos contextos, incluyendo programas internos de formación de las escuelas, grupos o círculos de colegas o como parte de un proceso de evaluación y supervisión.
- Mediante un proceso formal y dirigido, con indicaciones para quien lo realiza sobre lo que debe incluir y anotar, o en forma autodirigida e informal, siendo en este caso su autor el que decide los trabajos que lo integran y las reflexiones que hace sobre ellos.
- Por individuos o por grupos pequeños con capacidad para decidir sobre todos o algunos de los elementos que lo conformaran.
- Para documentar, en el caso de los docentes, lo que han aprendido sobre el currículo, la enseñanza y la evaluación diagnóstica a lo largo del tiempo. Los directivos escolares, por su parte, pueden asimismo documentar lo que han aprendido sobre la enseñanza, el aprendizaje, la dirección y supervisión.
- Cuando los distritos escolares rediseñan sus sistemas de evaluación y supervisión de los docentes, pudiendo el profesorado y los directivos, en estas situaciones, recoger diversas muestras de trabajo de su trayectoria profesional para que sean elementos mediadores de su comunicación con los supervisores.
- En el contexto de una investigación cooperativa, los docentes a través de recopilaciones de sus trabajos pueden explorar, analizar y abordar cuestiones que afectan a sus prácticas.

Klenowski (2005: 36) refuerza la siguiente idea

«[el] portafolios usado para fines de desarrollo profesional, en el contexto de enseñanza y aprendizaje, incluye materiales y ejemplos de trabajos que sirven como pistas para el examen crítico de las prácticas de enseñanza y aprendizaje. Se trata de que los profesores reflexionen y valoren sus propias prácticas de enseñanza y de que sean capaces de dialogar con otros para resaltar la complejidad y la diversidad de estas prácticas de enseñanza y aprendizaje».

Teniendo en cuenta lo anterior, parece claro que el portafolio se constituye en un dispositivo para que cada docente tome conciencia de la incorporación a su práctica educativa de una amplia base de conocimientos

teóricos y experienciales que no son sino el reflejo de cómo ha sido capaz de aprender profesionalmente, de los avances y progresos pero también de las dudas, preguntas e incertidumbres que van surgiendo en sus quehaceres habituales. Por ello, el primer destinatario del portafolio es el propio docente que lo elabora dado que a través de él puede asumir la dirección del desarrollo de su vida profesional.

Por otro lado, el portafolio también es un elemento de interlocución entre pares, de relación, diálogo e intercambio entre docentes que comparten la responsabilidad y el desafío de hacer posible un proyecto educativo. En tanto que medio para la socialización profesional podría, en cierta medida, favorecer la conformación de comunidades de trabajo-aprendizaje de manera que facilitara la creación de una cultura de trabajo en equipo y colaborativa que sea un contrapunto al tan extendido individualismo docente y a la balcanización de los procedimientos y modelos de organización y funcionamiento institucionales.

La dirección o el equipo directivo es asimismo otro de los destinatarios del portafolio de los docentes. Sobre todo cuando desde un modo de proceder participativo, el trabajo conjunto de la dirección, el profesorado y los equipos pedagógicos puede contribuir a la generación de conocimiento, erigiéndose el centro escolar en una fuente de aprendizaje institucional. Los portafolios, en estas dinámicas, abren las puertas de las aulas al análisis de los logros, experiencias innovadoras, necesidades, deficiencias, etc., que es necesario tomar en cuenta para establecer e implementar propuestas de mejora y de desarrollo profesional e institucional.

Existen maneras muy diversas de construir un portafolio. Bird (1997: 338), por ejemplo, propone hasta nueve formas diferentes de elaborarlos, tal y como se puede observar en el Cuadro 9,1, según se combinen, por una parte, los sujetos que participan en estas labores y, por otra, las normas, pautas y prescripciones que pueden seguirse en dichos menesteres.

Cuadro 9.1. Nueve modos de construcción de un portafolios.

PARTICIPACIÓN	NORMAS INFORMALES	PAUTAS FUENTES	PRESCRIPCIONES FORMALES
Entradas realizadas por el profesor.	<i>Entradas electivas:</i> Borradores, hojas, trabajos, tests, resultados de tests que interpretar, informes sobre un proyecto especial.	<i>Entradas dirigidas:</i> presentación escrita de una clase y materiales, una reunión de evaluaciones de alumnos.	<i>Entradas requeridas:</i> solicitud de empleo; cuestionario sobre participación en formación continua
Entradas realizadas por el profesor y otros.	<i>Productos colegiales:</i> Notas sobre un asesor, una clase preparada con un colega, intercambio de notas sobre un problema pedagógico.	<i>Entradas negociadas:</i> Documentación conjunta de prácticas pedagógicas y revisión de productos de los estudiantes, planes de desarrollo.	<i>Entradas censuradas:</i> evaluaciones de rendimiento como construcciones de tests, preguntas, calificaciones.
Entradas realizadas por otros fundamentalmente	<i>Comentario:</i> cartas de recomendación, notas de observación informal, elogios escritos de padres.	<i>Testimonios:</i> informes estructurados o calificación, observaciones, estudios sobre estudiantes o padres, informes sobre entrevistas.	<i>Informes oficiales:</i> diplomas, licenciaturas empleo, informes formación continua.

Fuente. Bird, T. (1997: 338)

Martín-Kniep (2001), por su lado, distingue varias modalidades de portafolios dependiendo del sujeto que los construye, docentes o directivos, de los propósitos por los que se elaboran, del profesor o directivo como personas que aprenden y del docente como responsable del desarrollo curricular y evaluación, así como por el grado de realización que presentan, elaborado, incipiente y rudimentario.

Un portafolios ejemplar sobre el profesor o directivo que aprenden tendría, según Martín-Kniep (2001: 65), las siguientes características:

Generalidades: El portafolio constituye un relato coherente de la historia del docente o directivo como persona permanentemente dedicada a aprender, reflexionar y encontrarle sentido a las cosas. Al revisar el portafolio, el lector llega a conocer al docente o directivo cuyos trabajos y logros

allí muestran, y puede comprender claramente su aprendizaje. El docente o directivo tiene clara conciencia de su destinatario y proporciona suficientes detalles del contexto en el que ha tenido lugar su trabajo o aprendizaje.

Carta al lector, notas diarias y otras piezas de reflexión: El portafolio da claras pruebas de una actitud atenta y reflexiva de parte del autor. (En el caso de los docentes, por ejemplo, las reflexiones revelan nuevas percepciones e interrogantes relacionados con el contenido del curso —formas alternativas de evaluación, autenticidad, teoría del aprendizaje, reflexión de los alumnos, elaboración de la tabla de especificación, investigación-acción—, con la aplicación de nuevos conceptos a las prácticas de enseñanza y evaluación diagnóstica del docente y con las reacciones de los alumnos ante las nuevas prácticas). Las reflexiones incluyen una evaluación de las virtudes y de los aspectos a mejorar del docente o directivo, así como una descripción de sus procesos de aprendizaje. El docente o directivo evalúa explícitamente el grado en que ha cumplido con los objetivos de las evaluaciones diagnósticas y ha establecido metas concretas y realistas para ampliar su aprendizaje, y señala los aspectos específicos a los que debe dar respuesta.

Un portafolio ejemplar del docente como responsable del desarrollo curricular y la evaluación, incluye (Martín-Kniep, 2001: 86-87), los siguientes elementos:

Generalidades: El portafolio cuenta la historia de un docente cuya meta es aunar las prácticas de la enseñanza y la evaluación, y orientarlas hacia resultados de aprendizaje significativos. El portafolio demuestra la excepcional capacidad del docente para aplicar conceptos y destrezas relacionados con el currículo y la evaluación en la preparación de clases, actividades y evaluaciones dirigidas a la obtención de determinados resultados claramente definidos. El docente ha incluido todos los esbozos de las evaluaciones, lo que le permite al lector ver los cambios y las mejoras efectuadas para aplicar principios de diseño e incrementar la coherencia entre el currículo, la enseñanza y la evaluación

Contexto: El portafolio proporciona un panorama claro y cabal del currículo del docente, incluyendo conceptos, temas, destrezas y evaluaciones. Las conexiones entre estos elementos son claros y evidentes. Los resultados en torno a los cuales se ha diseñado el currículo son significativos, están precisamente expresados en función del aprendizaje de los alumnos y su descripción incluye indicadores concretos y observables.

Plan de evaluación (de dos a tres pruebas o un portafolio del alumno):

El plan se combina a la perfección con el ejercicio de la enseñanza por parte del docente.... Las descripciones de las evaluaciones son detalladas y suministran todo el contexto necesario, incluyendo el área de contenidos abordada, los resultados previstos con sus correspondientes indicadores, los objetivos a alcanzar y una reseña detallada de las actividades previas y siguientes a la evaluación...

Diseño del portafolio: El diseño se orienta y pertenece al alumno. Tiene un objetivo y un público reales y documenta el rendimiento, el esfuerzo y el progreso del alumno hacia la obtención de resultados significativos y claramente explicados. El lector ve con claridad cómo se inserta el portafolio en el currículo y qué función cumplen en su elaboración los docentes, alumnos y otras personas...

Descripción de las evaluaciones: Las evaluaciones son verdaderamente auténticas. Son tareas insertas en el currículo, substanciales e integradoras, que requieren que los alumnos utilicen conocimientos previos, apliquen conocimientos de una o más áreas de contenidos, expresen y expliquen en detalle conclusiones, utilicen estrategias metacognitivas, hagan replanteos y correcciones. Las evaluaciones son valoradas por otros destinatarios fuera de la escuela, cumplen un propósito real y son lo bastante flexibles como para permitirles a los alumnos tener opciones y oportunidades de éxito... toman en cuenta diferentes necesidades y orígenes culturales... Incluyen consignas que orientan las reflexiones de los alumnos sobre los productos y los procesos y que pueden consistir en preguntas, listas de verificación o tablas de especificación... El docente ha señalado el grado en que los alumnos son representativos de la diversidad existente en el aula, permitiéndole al lector sacar conclusiones acerca del valor de la evaluación para una amplia gama de niveles de capacidad y conocimientos.

Estándares de desempeño: Los estándares de desempeño para las tareas correspondientes a la evaluación están claros para todos los interesados. El docente y los alumnos los determinaron conjuntamente y los articularon en formas de tablas de especificación, criterios de puntuación y/o ejemplares. Estos estándares sirven de guía a los alumnos para evaluar sus trabajos y fijarse metas para mejorar su desempeño...

5. A MODO DE CONCLUSIÓN

La evaluación entre iguales es una forma de abordar la evaluación del desempeño docente desde la orientación de la mejora y desarrollo profesional del profesorado. Como hemos tenido oportunidad de ver, tanto en el tema de la evaluación de centros como ahora en la del profesorado, la modalidad de autoevaluación puede llevarse a cabo de muy distintas maneras. Tanto desde una perspectiva individual como grupal o en equipo, es una herramienta que puede aportar información y evidencias para conocer las características más relevantes de la docencia y reforzar la capacitación del profesorado. El auténtico desafío de los procesos de autoevaluación reside en la concepción del profesor y de la enseñanza que vertebra la propuesta, así como en la autonomía de los actores educativos para la toma de decisiones de todos los elementos y aspectos que conlleva la misma.

La modalidad de autoevaluación vinculada a la investigación-acción, en sus concepciones de práctica deliberativa y crítica emancipadora, tiene el acierto de propugnar el protagonismo del profesorado en sus procesos de indagación e investigación, no perdiendo de vista en la evaluación docente las dimensiones grupal y colaborativa de sus desempeños y su proyección en el aprendizaje del profesor y en la capacidad de aprender de las mismas instituciones escolares. Quizá, en este sentido, el portafolio es algo más que una técnica o un instrumento, dado que trasciende su función meramente informativa para adentrarse en la documentación de las experiencias y proyectos que jalonan la trayectoria biográfica del profesorado.

BIBLIOGRAFÍA

- AIRASIAN, P.W. y GULLICKSON, A. R. (1999): *Herramientas de autoevaluación del profesorado*. Bilbao, Ediciones Mensajero, 2ª ed.
- BARBER, L.W. (1997): «Autoevaluación» en J. Millman y L. Darling-Hammond, (eds.): *Manual para la evaluación del profesorado*. Madrid, La Muralla, pp. 300-315.
- BIRD, T. (1997): «El portafolios del profesor: un ensayo sobre las posibilidades», en J. Millman y L. Darling-Hammond, (eds): *Manual para la evaluación del profesorado*. Madrid, La Muralla.

- DES (1983): *Teaching Quality*, Cmnd 8836. London, Her Majesty's Stationery Office, Department of Education and Science.
- ELLIOTT, J. et al. (1986): *Investigación-acción en el aula*. Valencia, Generalitat Valenciana.
- ELLIOTT, J. (1996): *El cambio educativo desde la investigación-acción*. Madrid, Ediciones Morata, 2ª ed.
- (1998): «Saber, poder y evaluación de los docentes», en W. Carr, (ed.): *Calidad de la enseñanza e investigación-acción*. Sevilla, Díada Editora, 3ª ed. rev., pp. 167-188.
- GAVARI STARKIE, E. (2009). «El papel del docente y del alumno en la elaboración del Portafolios», en *Revista de Educación*, 349, pp. 451-462
- GRANT y HUEBNER (1998): «The Portfolio Questions: The Power of Self-Directed Inquiry», en N. Lyons, (ed.): *With Portfolio in Hand: Validating the New Teacher Professionalism*. New York, Teachers College Press, pp. 156-71.
- HARGREAVES, A. (1993). «La reforma curricular y el maestro», en *Cuadernos de Pedagogía*, n. 211, febrero.
- HOUSE, E. R. y LAPAN, S. D. (1989): «Teacher appraisal», en H. Simons y J. Elliott, (eds.): *Rethinking appraisal and assessment*. Milton Keynes, Open University Press, pp. 55-63.
- JAMES, M. (1987): «Self-Initiated Self-Evaluation», en P.S. Clift, D. L. Nutall y McCormick (eds.): *Studies in School Self-Evaluation*. Lewes, The Falmer Press-Open University Set Book, pp. 172-189.
- KEMMIS, E.(1986): «Seven Principles for Programme Evaluation in Curriculum Development and Innovation», en E.R. House, (eds.): *New Directions in Educational Evaluation*. Lewes, Falmer Press.
- KLENOWSKI, V. (2005): *Desarrollo del Portafolios para el aprendizaje y la evaluación*. Madrid, Ediciones Narcea.
- LAFFITTE FIGUERAS, R. M. (1994): *Rendimiento de cuentas y evaluación*. Barcelona, PPU.
- LATORRE, A. (2003): *La investigación-acción. Conocer y cambiar la práctica educativa*. Barcelona, Editorial Graó.

- LEWIN, K. (1947): «Group decision and social change» en T. Newcomb y E. Hartley, (eds.): *Readings in Social Psychology*. New York, Henry Holt, pp. 330-344.
- MARTIN-KNIEP, G. O. (2001): *Portfolios del desempeño de maestros, profesores y directivos. La sabiduría de la práctica*. Buenos Aires, Editorial Paidós.
- McKERNAN, J. (1999): *Investigación-acción y curriculum*. Madrid, Edic. Morata.
- NEVO, D. (1997): *Evaluación basada en el centro: un diálogo para la mejora educativa*. Bilbao, Mensajero.
- SOLABARRIETA, J. (1996): «Modelos de autoevaluación del profesor», en F. J. Tejedor y J. L. Rodríguez Diéguez, (eds): *Evaluación educativa II. Evaluación Institucional. Fundamentos teóricos y aplicaciones prácticas*. Salamanca, IUCE, pp.173-181.
- WILSON, J. D. (1992): *Cómo valorar la calidad de la enseñanza*. Madrid, Ediciones Paidós-Ministerio de Educación y Ciencia.

Tema 10

La evaluación de los docentes por los alumnos. Experiencias con el profesorado universitario

Introducción

1. Las contradicciones de la evaluación de los docentes
2. La iniciativa y el control de los procesos evaluativos
3. Los cuestionarios de opinión, modelo hegemónico de evaluación del profesorado universitario
4. Una experiencia alternativa de evaluación de los procesos de enseñanza y aprendizaje en la universidad
 - A. Acciones de información-concienciación
 - B. Trabajo de campo
 - 4.1. Análisis de la experiencia

Bibliografía

TEMA 10
LA EVALUACIÓN DE LOS DOCENTES POR LOS ALUMNOS.
EXPERIENCIAS CON EL PROFESORADO UNIVERSITARIO

Juan Fernández Sierra.
Susana Fernández Larragueta.
Universidad de Almería

INTRODUCCIÓN

En el sistema educativo, tradicionalmente sólo ha habido un ámbito que ha sido evaluado sistemática y constantemente: el alumnado, y, salvo excepciones, con la sola finalidad de sancionar su rendimiento escolar. Esa evaluación es realizada por los profesores, los que, por lo mismo, han ido forjándose a lo largo de su experiencia personal, de la historia de su profesión y de las prácticas sociales en esta cuestión, un concepto controlador de esta actividad. Por lo tanto, es lógico que cuando se habla de evaluación del profesorado *sientan cierta inquietud*, ya que en ello, visto desde esa perspectiva sancionadora, se juegan su prestigio, su autoestima, su anonimato e, incluso, su trabajo. Ahora bien, de los elementos que intervienen en los procesos de enseñanza institucionalizada es sin duda el profesorado el que mayor incidencia tiene pues, como dice Wilson (1992), *la calidad del sistema educativo depende en definitiva de la calidad de las personas que lo desempeñan* (p. 126). Evidentemente los y las docentes son los encargados de introducir, consciente y racionalmente, cambios y matices en sus acciones y en la vida del aula e institución (metodología, modelos de evaluación, tipos de relaciones interpersonales e interprofesionales, etc.) que modulan el clima de trabajo y marcan el ritmo y el cariz de los aprendizajes de los alumnos; no podemos, pues, renunciar a la evaluación del elemento que mayor incidencia ejerce en el proceso de enseñanza.

En esta línea, el presente tema lo dedicamos a la evaluación del profesorado, centrándolo en el universitario, en el que la importancia de la participación e implicación de los y las estudiantes es esencial; lo que ha llevado a menudo a reducir dicha evaluación a una recogida de su opinión mediante un cuestionario, la mayoría de las veces simplificado y simplificador de las relaciones de enseñanza y aprendizaje que se establecen en las instituciones universitarias.

1. LAS CONTRADICCIONES DE LA EVALUACIÓN DE LOS DOCENTES

De la evaluación de los elementos del sistema educativo, como hemos dicho, la del profesorado es la más candente ya que a la complejidad inherente a todo lo educativo, ha de añadirse el hecho de que con ella pueda ponerse en peligro la seguridad psicológica y laboral de algunas personas. Muchos docentes, en principio, se resisten explícita o tácitamente a ser evaluados, por lo que un planteamiento erróneo o un estado de desinformación pueden provocar una oleada de protestas y resistencia activa de éstos y de sus organizaciones laborales, profesionales o sindicales. No obstante, esto se logra disminuir realizándola bajo determinadas condiciones que, en esencia, podrían ser no poner en peligro su prestigio profesional ni su seguridad laboral. Cuestiones relativas a qué datos recoger, quiénes van a obtenerlos y cómo serán empleados, son temas previos de negociación si se quiere solventar el problema más difícil, es decir, conseguir que se acepte la evaluación (Wilson, 1992).

Por otro lado, la evaluación de los profesores y profesoras, según el tipo y forma de planificarla y ejecutarla, puede ser más perjudicial que beneficiosa para los aprendizajes de los estudiantes. Evidentemente, cualquier evaluación condiciona la vida académica y el sentido de los procesos de enseñanza y aprendizaje que en ella se desarrollan y, no es menos evidente que, los docentes, por propia supervivencia laboral y profesional, tenderán a comportarse y actuar de forma tal que le garantice *quedar en buen lugar* ante la evaluación, independientemente de sus convicciones educativas y de la riqueza de los procesos que ello comporte, como expone Montero Mesa (1993: 222),

«Los instrumentos de evaluación [...] tienen implícito un modelo de profesor universitario y pueden configurar las características de la enseñanza que los profesores universitarios realicen por querer “acomodarse” a ese modelo y resultar así positivamente evaluados».

Dado el destacado papel que ellos juegan en los procesos de enseñanza y, consecuentemente, en los tipos de aprendizajes que se potencian en las aulas y en los centros escolares, una actuación forzada y *artificial* de aquellos no supondrá mejoras significativas en la calidad de la enseñanza, por el contrario puede provocar distorsiones y sesgos educativos importantes.

Igualmente, han de tomarse en consideración algunos problemas éticos, entre los que hay que destacar el de la publicidad de las evidencias afloradas mediante la acción evaluadora. Si queremos que la evaluación tenga valor

formativo para todos los implicados en los procesos de enseñanza-aprendizaje, es imprescindible el conocimiento, análisis y debate conjunto de dichas evidencias. Ahora bien, aquí nos encontramos con un dilema de carácter ético difícil de solventar: por un lado, el derecho de los profesionales a su privacidad y, por otro, el derecho público a saber (Sirotnik, 1990); que en nuestro caso podríamos concretarlo en cómo compaginar la preservación de la privacidad del profesor o profesora y la conveniencia de que los estudiantes sean informados e involucrados en el debate y reorientación de algo que les atañe tan directamente como la acción docente de la que son parte esencial, no sólo como un derecho legítimo, sino, y sobre todo, como posibilidad de poder participar racionalmente en el análisis conjunto de las situaciones de enseñanza y aprendizaje que se viven y potencian en las aulas e institución.

Así pues, los problemas fundamentales con los que nos encontramos en la evaluación de los profesores son más de orden pedagógico, ético y político que técnicos. El conciliar el control y la mejora o, lo que es similar, la evaluación formativa y sumativa es técnica y teóricamente posible pero, como han puesto de manifiesto los estudios de Duke y Stiggins (1997) es muy difícil compatibilizar o aunar evaluaciones para la *accountability* y para el desarrollo profesional. Teniendo en cuenta que la evaluación se entiende como la emisión de un juicio de valor sobre una actuación o modo de hacer, siempre hemos de contar con una dimensión política que puede favorecer a unos intereses a costa de otros, sobre todo cuando en la evaluación se vislumbran más objetivos de control que intenciones de mejora (Fernández Sierra, 2002).

Aunque las evaluaciones formativa y sumativa, más que como dos modalidades de evaluación, las entendamos como dos partes de un mismo proceso (Scriven, 1991), sí es cierto que su finalidad difiere, siendo, en el caso concreto de la evaluación del profesorado, la segunda la que despierta más recelos, pues las evaluaciones nacidas con esta finalidad tratan de basarse en criterios pretendidamente objetivos e informar a audiencias externas al grupo de profesores, por lo que pueden resultar más amenazantes personal y profesionalmente para los enseñantes.

Por el contrario, la evaluación formativa, que suele ser llevada a cabo primordialmente por el grupo de iguales, pone menos en peligro la autoestima profesional y los intereses laborales de los profesores, produciendo así

un nivel inferior de angustia, ansiedad y distorsión en la vida personal y profesional de los individuos que son sometidos a evaluación. Ofrece, de esta manera, más posibilidades de sinceridad, autocrítica y reflexión colectiva, favoreciendo el perfeccionamiento de los docentes y de las instituciones educativas, objetivo fundamental, según se desprende de las opiniones mayoritarias de los expertos que debe perseguir la evaluación del profesorado: “El procedimiento de revisión por colegas y sus resultados tienen una gran influencia en la calidad” (Vroeijensti y Acherman, 1991: 106). Claro está que la predisposición a la autosuperación del profesor es condición indispensable para una evaluación propiamente formativa.

2. LA INICIATIVA Y EL CONTROL DE LOS PROCESOS EVALUATIVOS

Una cuestión fundamental por su incidencia en el carácter controlador o de perfeccionamiento profesional de la evaluación del profesorado es la iniciativa, es decir, de quién parte y cómo se pone en marcha el proceso de evaluación. Evidentemente el potencial transformador de la acción evaluadora no es el mismo si la idea de evaluar surge de la administración que si nace del interés de los profesores de un centro educativo. Como dicen Vroeijensti y Acherman (1991), la primera está más centrada en el control y la segunda en la mejora.

Pero no sólo es importante quién da el primer paso para la evaluación, sino también la forma de ejecutarla. Si la iniciativa es externa pero se lleva a la práctica después de un período de información y convencimiento de los profesores, mediante un proceso de negociación abierto y flexible en el que se les da la oportunidad de opinar y participar en el diseño, desarrollo y control de la evaluación, se suavizarán resistencias y se posibilitará la función formativa o de mejora.

Para Elliott (1984), otorgar a los docentes el derecho a tener control sobre el acceso público al conocimiento de su actividad no es sólo una opción ética, sino además un requerimiento metodológico porque la ética es una condición necesaria para comprobar la objetividad de la comprensión de los *externos*. La prueba de que la comprensión de éstos representa bien las actividades de los *internos*, aunque sea parcial, reside en que los últimos la perciben como una representación acertada, justa. La comprensión se desarrolla en el diálogo personal con los *internos*. De ahí la impor-

tancia que MacDonald (1978) atribuye a la negociación como un concepto clave en la evaluación democrática.

Otra alegación en favor del control de los docentes sobre los procesos de evaluación viene del campo del profesionalismo. Si se acepta que la actividad docente es una profesión hay que admitir que los profesores tienen autonomía en sus juicios y en la autorregulación de su práctica, y además que están sometidos al código ético, al análisis colegiado que le impone su gremio para asegurar la suficiente calidad en su actuación (así sucede con médicos, abogados y otras profesiones).

“Conceptualizar al profesor como profesional permite entenderlo como: poseedor de una base de conocimiento suficientemente amplia para desarrollar su tarea, capaz de proporcionar conocimiento sobre su práctica y de buscar los recursos necesarios para mejorarla; con una actitud positiva hacia su desarrollo profesional continuo; con una autonomía en la relación de su tarea, capaz de adecuarla al contexto y de cooperar con otros profesionales; con un código ético y mecanismos de autocritica” (Montero Mesa, 1990: 271).

Sin embargo, este tema hay que abordarlo con sumo cuidado, pues la defensa a ultranza del profesionalismo, como recuerda Strike (1990: 394), puede entrar en conflicto con el control democrático: “La institucionalización del profesionalismo requiere alguna erosión de la autoridad democrática”. Este mismo autor aboga por un profesionalismo que mine la autoridad burocrática, pero que sea compatible con la democracia comunitaria, es decir, la caracterizada por la toma de decisiones utilizando el diálogo entre todas las partes implicadas. La democracia comunitaria —afirma—, “como el profesionalismo, valora la sabiduría de los profesores en aquellas áreas en las que tienen conocimiento, pero a diferencia de él, no considera la pericia como una justificación para defender la autoridad” (íbid: 371).

Lo realmente importante para que la evaluación de los docentes sea un elemento al servicio del perfeccionamiento profesional y de la calidad educativa es que los profesores y la comunidad universitaria acepten, participen y controlen la evaluación, al tiempo que se comprometan tanto en su realización como en la adopción de las medidas necesarias para conseguir el aumento de la pretendida calidad; como expone Bolívar (2008: 58) “la evaluación del desempeño docente y la propia carrera profesional no es sólo una cuestión de gestión de recursos humanos, sino que deben inscribirse, de modo congruente, en una política educativa amplia de mejora”.

3. LOS CUESTIONARIOS DE OPINIÓN, MODELO HEGEMÓNICO DE EVALUACIÓN DEL PROFESORADO UNIVERSITARIO

La evaluación del profesorado es una preocupación generalizada en los países occidentales, aunque no en todos ellos se lleva a cabo sistemáticamente, ni de la misma forma, ni con los mismos fines. La complejidad de la tarea, sus connotaciones políticas, sus implicaciones sindicales y la dimensión ética que comporta, posiblemente, estén en la raíz de su no puesta en práctica generalizada.

Centrándonos en las experiencias de evaluación del profesorado universitario puestas en marcha en el interior de las universidades en nuestro país, podríamos decir que, en general, son más propiamente “sondeos de opinión a los alumnos sobre la calidad docente de sus profesores” (Apodaca, 1992: 75). La mayor parte de estas experiencias de evaluación interna del profesorado han adoptado, desde sus comienzos hasta la actualidad, la técnica del cuestionario para recabar la información, opiniones y valoraciones desde las que enjuiciar la labor docente del profesorado. En este aspecto, se suele señalar que la utilización de estas encuestas han supuesto un gran esfuerzo para identificar las dimensiones de enseñanza efectiva que deben estar presentes, pero también presentan dificultades de validación y de sus valoraciones, al no existir un criterio universal de enseñanza efectiva; igualmente hay que tener en cuenta las limitaciones de este tipo de instrumentos para reflejar fielmente la adecuación entre las valoraciones de los estudiantes y el estado real de la enseñanza, los patrones utilizados en función de los propósitos de la evaluación, la estabilidad con distintos grupos de estudiantes y los problemas relacionados con su validez. Globalmente considerado el cuestionario, no cabe duda, que es un instrumento apropiado para recoger información de grupos o muestras de gran tamaño, como una primera aproximación al conocimiento general del objeto de evaluación. Sin embargo, parece claro que, al igual que ocurre con las fuentes y momentos de la evaluación, su uso debe complementarse con otras técnicas e instrumentos con las que contrastar sus aportaciones.

La utilización de la encuesta o cuestionario presenta, pues, importantes limitaciones, como son:

En primer lugar, por su propia naturaleza como instrumentos cuantificables y cerrados, las escalas y cuestionarios son incapaces de recoger la

variedad y riqueza de los matices que proliferan en las situaciones y acciones docentes. La simplificación y la “horma” que supone cada una de las preguntas, imposibilita respuestas contextualizadas y justas. En este sentido, como apuntan Cisneros-Cohernour y Stake (2010: 219-220) «La efectividad de la práctica docente es medida de manera global, calificando ciertas características percibidas como sinónimos de efectividad. Este paradigma tradicional implica que la docencia es un tratamiento y su evaluación indica la efectividad de este tratamiento»; sin embargo, «la confiabilidad de los resultados de las evaluaciones obtenidos por medio de cuestionarios aplicados a estudiantes (...) describen los puntajes obtenidos de estas evaluaciones como “multidimensionales, confiables, estables y relativamente válidos en contra de una variedad de indicadores de enseñanza efectiva”».

En segundo lugar, estos instrumentos se elaboran partiendo de un modelo de profesor tácito, representado por el tipo y estilo de las preguntas realizadas, que en principio ni tiene por qué ser el único válido ni, posiblemente, el ideal en todas y cada una de las instituciones educativas y aulas exploradas, en este sentido cabe preguntarse qué se considera como *buen profesor* sobre el que construir el modelo ideal de docente al que tender (Tejedor Tejedor, 2009; Fernández y Coppola, 2012), y si esa pregunta es idónea en un contexto social diverso y múltiple, tanto en cuanto a expectativas, pensamiento docente y realidad sociocultural y económica.

En tercer lugar, los estudiantes, igualmente, responden influidos por su propio ideal de profesor y otros factores académicos y extraacadémicos (Fernández Sierra, 1996), pudiendo ser ese modelo más fruto de su historia personal y académica que de una reflexión rigurosa de lo que debe ser la enseñanza y el papel que en ella han de jugar los profesores o profesoras.

En cuarto lugar, este tipo de evaluaciones no dan la oportunidad de contrastar los modelos docentes que subyacen en el pensamiento de los docentes y en el de los estudiantes, por lo que no son aptos para la evolución y superación de viejas creencias y posibles prejuicios, a través del análisis colectivo de los pro y los contra de los juicios previos y de las situaciones reales.

En quinto lugar, no garantizan que la interpretación de los destinatarios se corresponda con las intenciones del elaborador de la prueba ni con los informantes o encuestados, no habiendo oportunidad ni espacio para aclarar las dudas o diferencias al no existir una interacción directa entre ambos.

En sexto lugar, como ocurre con todo instrumento cuantificable, se meten en el mismo «algoritmo» cuestiones diversas de la acción docente (desde la evaluación hasta las relaciones con los alumnos pasando por la supuesta preparación de las clases), lo que, además de aportar información imprecisa, la convierte en poco útil para el profesor interesado en la superación de sus deficiencias.

No obstante, de aquí no debe deducirse que la opinión de los estudiantes es irrelevante para la evaluación de los profesores. Por el contrario, es uno de los factores más importantes a tener en cuenta en todos los niveles educativos, pero especialmente en la enseñanza universitaria. La crítica va en la línea del sesgo informativo que supone recoger datos de un sólo sector de informantes y, sobre todo, la poca o nula oportunidad que docentes y estudiantes tienen de contrastar y discutir sus posiciones y, consecuentemente, avanzar en la búsqueda conjunta de soluciones a las deficiencias detectadas. El *secretismo* al que por diversas, y tal vez fundadas, razones se somete la evaluación del profesorado, tampoco facilita el debate público y el aprovechamiento del esfuerzo evaluativo.

En la comunidad universitaria predomina la idea de que este tipo de trabajos y forma de actuar posee poca o nula capacidad para mejorar las actuaciones docentes y la vida de las aulas. En esta línea, es fácil observar cómo en las universidades en las que se han repetido este tipo de actuaciones durante varios años, muchos estudiantes son reacios a cumplimentar los cuestionarios, quejándose de que *no sirven para nada* y de que ellos no llegan a enterarse nunca de los resultados. Ante esta aparente inutilidad, algunos gestores e, incluso, parte del profesorado, piensan que habrían de hacerse públicos los resultados de las evaluaciones, presuponiendo que de esa forma sí serían eficaces para la mejora de la docencia; y de hecho, si no lo llevan a cabo, no es por falta de convencimiento, sino por el temor a las reacciones y contestación que prevén puede tener esa medida.

La salida, desde nuestro punto de vista, no está en la resolución imposible del dilema o contradicción mencionada, decantándonos por uno de los extremos o premisas, sino que hay que buscarla profundizando en el concepto y la praxis de la acción docente y discente universitarias, entendiéndolas y ejercitándolas como responsabilidad común de profesores y estudiantes que han de potenciar, más que relaciones jerárquicas y transmisoras-receptoras del saber, comunidades de aprendizaje mutuo. Desde esta dimensión, la

privacidad como mezcla de lo personal y lo profesional deja de tener sentido, situándose el análisis de toda acción educativa y, por tanto, evaluadora en el ámbito público del aula, donde las responsabilidades y actuaciones son, inevitablemente, coparticipadas y de dominio público¹³.

Bajo esta concepción es imposible plantear la evaluación del profesorado como una acción aislada, siendo preciso encarar estudios holísticos que no obvien la interconexión entre los diversos elementos y factores que determinan, inciden o condicionan las actuaciones formativas y educativas. Habríamos, pues, de promover evaluaciones del profesorado que, como insiste incluso la OCDE (1991), tengan como punto de mira o intención el facilitarles su desarrollo profesional, a lo que habría que añadir, que además ayuden al/la estudiante a tomar conciencia de su papel y responsabilidad en los procesos de enseñanza y aprendizaje y en el ámbito académico y social en el que se desenvuelven.

4. UNA EXPERIENCIA ALTERNATIVA DE EVALUACIÓN DE LOS PROCESOS DE ENSEÑANZA Y APRENDIZAJE EN LA UNIVERSIDAD

La experiencia que exponemos se ha llevado a cabo en una universidad andaluza. El equipo de gobierno, como el de otras universidades, muestra interés por realizar una evaluación del profesorado similar a las que se habían venido llevando a cabo en muchas universidades españolas, o sea, aplicar un cuestionario para recoger la opinión de los estudiantes sobre sus profesores y profesoras.

En opinión del equipo de evaluación, responsable de estos trabajos, este tipo de actuaciones estaba teniendo escasa incidencia en la mejora de la acción docente del profesorado y de la calidad de la enseñanza universitaria, por lo que propuso avanzar hacia sistemas más globales. Siendo una universidad mediana, pensaba, sería posible experimentar otros modos de acción evaluadora que, además de ser más útiles para la mejora profesio-

¹³ El pretendido secretismo de las opiniones de los estudiantes sobre las acciones de los profesores es una falacia, puesto que los comentarios, creencias más o menos fundamentadas, representaciones, etc., que éstos tienen de cada profesor o profesora es de dominio público en el ámbito del aula (y a menudo fuera de ella). Por lo tanto, el poner sobre la mesa los datos e informaciones recogidas por métodos más rigurosos, lo único que puede suponer es dar seriedad y racionalidad al debate y colaborar en el desmonte de prejuicios e ideas preconcebidas.

nal, fuesen promoviendo una cultura unida al cambio y la innovación. La idea fue aceptada y se acordó elaborar un plan de evaluación que en esencia se sustentaba en cuatro principios:

- a. La evaluación no se ceñiría al profesorado, aunque sí sería el ámbito central de interés.
- b. Todo proceso de evaluación que se pusiera en marcha habría de tener un carácter voluntario; habría de ser solicitado (o al menos aceptado), individualmente cuando se tratase de evaluaciones personalizadas, o por los responsables de los entes que se fueran a evaluar, cuando se tratase de organismos o instituciones.
- c. Las acciones del primer año no serían actuaciones aisladas, puntuales, sino una primera fase de un plan de evaluación global de la Universidad en todas sus dimensiones a lo largo de un quinquenio.

Bajo esa perspectiva se concretó un *plan de evaluación* que contemplaba la actuación en cuatro dimensiones complementarias:

- 1ª. Realizar un estudio de opinión sobre aspectos generales de la vida universitaria (gestión, servicios, docencia, colaboración universidad-sociedad, etc.). Con ello se pretendía obtener una radiografía o panorámica de la satisfacción de los diversos sectores con respecto a diferentes aspectos que afectaban a la vida de la Universidad en sus inicios.
- 2ª. Ofrecer y llevar a cabo paulatinamente evaluaciones de centros, departamentos, organismos de gestión o gobierno, servicios, etc.
- 3ª. Evaluar los procesos de enseñanza y aprendizaje que se ponen en funcionamiento en las aulas. En este apartado se atenderían dos tipos de demandas: las realizadas por los profesores y profesoras individualmente y las que solicitaran los departamentos para sus miembros.
- 4ª. Poner en marcha, paralelamente, un proyecto de metaevaluación del Plan con el fin de valorar y reorientar los procesos de evaluación puestos en escena.

La respuesta de la comunidad universitaria fue altamente positiva, pues, ciñéndonos a las peticiones realizadas por escrito, nos encontramos con que uno de cada cuatro profesores y profesoras solicitaron participar en la experiencia, sometiéndola a evaluación las materias que impartían.

A partir de estos ofrecimientos pusimos en práctica diversos trabajos, entre ellos el de la **evaluación de los procesos de enseñanza y aprendizaje**, que es la que aquí explicamos. La propuesta era llevar a cabo la evaluación de la docencia y el clima de clase en las aulas, empleando diferentes métodos y técnicas de recogida de datos, combinando los cuestionarios con las entrevistas al alumnado y al profesorado, y la triangulación.

La mayor dificultad que preveíamos para llevar a cabo este Plan, dado su carácter voluntario y su pretensión formativa, era la captación de profesores y profesoras dispuestos a participar en esta experimentación, teniendo en cuenta que la implicación de los estudiantes en el debate de los datos obtenidos rompería en parte la privacidad y el secretismo tradicional de este tipo de evaluaciones. Por ello, pusimos en marcha una serie de acciones que, para sistematizar la exposición, podemos agruparlas en dos fases o dimensiones:

A. Acciones de información-concienciación

El proceso más delicado, en el que se juega gran parte del éxito o fracaso de una evaluación —que como ésta, nazca con la intención de ser un proceso de aprendizaje colectivo y de autoaprendizaje que posibilite la comprensión y la mejora del trabajo docente y del ambiente social de aprendizaje que se producen en las aulas—, es el de su presentación, análisis y asunción o aceptación por las personas que de alguna manera se van a ver o sentir como *sujetos evaluados*.

Hay que ser sumamente cuidadosos con este aspecto. En el caso que aquí expongo, lo primero que hicimos fue explicar, debatir y consensuar con las autoridades académicas universitarias la filosofía del *plan*, sus líneas generales de actuación y los focos primordiales de estudio. Alcanzado el acuerdo que otorgaba a la Unidad de Evaluación la independencia de criterio y acción necesarios para desarrollar su trabajo bajo el modelo que MacDonald (1985) llama *Evaluación Democrática*, se abrieron diferentes vías de acción encaminadas a expandir los principios en los que se sustentaban el plan a la comunidad universitaria.

La vía principal en este caso conducía a los profesores y profesoras individualmente y, a partir de ellos, a sus aulas. Se les enviaron diversos documentos sobre lo que nos proponíamos trabajar, se organizaron numerosas

sesiones informativas y de debate en grupos reducidos, se analizaron con ellos los instrumentos en elaboración, se convinieron los términos de la recogida y utilización de los informes, se escucharon sus sugerencias, etc. Seguidamente se les invitó a que solicitaran por escrito someter su docencia y sus clases a evaluación.

Paralelamente, se celebraron reuniones con los representantes del alumnado en las asociaciones estudiantiles y en las Juntas de Centro y, más tarde, con los delegados y subdelegados de los cursos de los profesores y profesoras que habían solicitado la evaluación, a fin de tenerlos informados, explicarles las características, importancia y conveniencia del proceso y pedirles su colaboración activa y consciente, y la de sus compañeros.

Ésta, como hemos dicho, era para nosotros una parte esencial del trabajo, pues de ello depende, no sólo el llevar a cabo la evaluación (o sea captar profesores y profesoras para ello), sino el que los procesos evaluativos puedan desencadenar procesos de aprendizaje profesional y la reorientación de las relaciones y condiciones de enseñanza que se ponen en marcha en las aulas.

B. Trabajo de campo

Una vez concluida la fase anterior que podemos designar como de preparación, entramos en la de evaluación propiamente dicha. Ésta, en esencia, consistió en:

- 1º. Selección de la muestra. Cuando se hizo la propuesta de evaluación al profesorado se preveía que, en el mejor de los casos, durante el primer curso no se llegaría a implicar en este proyecto experimental a más del diez por ciento del mismo; sin embargo, en las dos primeras semanas eran más del doble las solicitudes recibidas, con varias materias diferentes cada uno/a; por lo que era imposible atender todas las demandas, y por esa razón se decidió: a) evaluar una sola de las disciplinas y clases en las que impartían docencia cada uno de ellos; b) la selección se haría de acuerdo a los siguientes criterios: asignaturas troncales u obligatorias, de curso completo o impartidas en el primer cuatrimestre; c) cada profesor aceptaría personalmente la sugerencia o, en caso contrario, se analizarían con él las alternativas oportunas.

La casi totalidad de los profesores y profesoras vieron razonable la propuesta. Hubo algunos de ellos que creyeron conveniente que se les evaluase más de un grupo o materia y, por supuesto, se atendió su demanda. En total nos encontramos con que uno de cada cuatro docentes habían optado por participar en la evaluación en la primera edición. La cantidad fue similar al curso siguiente, en el que se les pidió que no repitiesen los del primero para no sobrecargar de trabajo del equipo y poder así concluir el proceso íntegramente. Además, consideramos que la evaluación del profesorado entendida como un modo más de autoaprendizaje profesional docente no es necesario que se lleve a cabo anualmente.

- 2°. Aplicación de los cuestionarios a los alumnos y alumnas de cada clase. El primero de ellos —de preguntas cerradas y preparado para lector óptico—, estaba compuesto por 41 ítems agrupados en cinco apartados que recogían información sobre: la actuación y actitud personal como alumno en la asignatura que se evaluaba, la actuación y actitud del grupo/clase respecto a la asignatura evaluada, la actitud y actuación del profesor, la materia y sus contenidos, y las condiciones materiales y recursos disponibles. El segundo, en el que se combinaban cuestiones cerradas y abiertas, preguntaba a los estudiantes sobre su historial académico reciente; los aspectos positivos y negativos referidos al profesor/a, materia y a sí mismos; la adecuación de los materiales utilizados en las clases; los tipos de aprendizaje fomentados en las mismas y su grado de acuerdo con ellos, y, finalmente, se les pedían sugerencias para mejorar el funcionamiento de las clases en la asignatura en cuestión.

Los datos se recogían en dos períodos o momentos distintos del curso escolar. Para la evaluación de las asignaturas impartidas en el primer cuatrimestre y las de año completo, se pasaron los cuestionarios a los estudiantes durante la segunda mitad de marzo y las entrevistas se hicieron en el mes de abril (una vez analizados los datos cuantitativos). Las entrevistas a los profesores o profesoras se llevaron a cabo en los meses de mayo y junio. Estos alumnos ya habían completado el ciclo lectivo de dichas asignaturas, es decir, se habían examinado y conocían las notas finales obtenidas en ellas. Para las materias impartidas en el segundo cuatrimestre y alguna de carácter anual, los cuestionarios al alumnado se pasaron en la segunda quincena de mayo, las entrevistas a éstos en la primera de junio y a los profesores al final de ese mes y principios de julio; no obstante, algunos de

ellos, por diversos motivos, fueron entrevistados después de las vacaciones de verano. A diferencia de los anteriores, en esta fase, los estudiantes no habían sido examinados.

- 3º. Entrevistas a una muestra teórica de estudiantes. Una vez analizados los datos de los cuestionarios se realizaban entrevistas a una representación¹⁴ de alumnos y alumnas de cada clase. Durante ellas se profundizaba en aquellos aspectos de la información obtenida que se consideraban más significativos y relevantes para el análisis y mejora de los procesos de enseñanza y aprendizaje de cada materia y profesor en particular.
- 4º. Elaboración y contraste de los informes con los docentes. Después de este proceso, se elaboraron los informes personalizados en los que se plasmaban las interpretaciones y valoraciones que los evaluadores realizaban a partir de los datos cuantitativos y cualitativos recogidos. Se remitía dicho informe a cada docente, junto con las tablas y análisis estadísticos correspondientes, y se les citaba, para unos días después, a una entrevista. En ella se analizaba conjuntamente el informe, se contrastaban opiniones y se recogía información complementaria sobre diversos aspectos de su acción docente, de las actitudes y actuaciones de los alumnos, de las características de la materia, medios y recursos disponibles, etc. Se finalizaba dialogando sobre el proceso de evaluación seguido y solicitando sugerencias para su evolución. Con posterioridad, los evaluadores reelaboraban el informe, lo convertían en definitivo y lo devolvían a cada profesor o profesora.

Sin embargo, nuestras intenciones iban más allá, pretendiendo avanzar dos pasos adelante. El primero consistía en lograr sesiones de triangulación entre profesores, evaluadores y alumnado. Esta propuesta, que estaba prevista en el plan inicial, no se pudo completar por motivos de tiempo, de personal y de posibilidades organizativas, salvo en algún caso puntual; empero seguimos pensando que es una acción imprescindible en un proceso de este tipo. El segundo paso al que aludimos es al análisis y debate de los informes de evaluación conjuntamente profesor y su grupo-clase o una representación de éste. Ésta es una decisión que algunos enseñantes toma-

¹⁴ Se hace un muestreo teórico (delegado/a, subdelegado/a, alumnos/as con diferente nota e implicación en clase), aumentado con algunos estudiantes al azar.

ron y ejecutaron, pero no se generalizó. Esta acción final, además de ofrecer a los profesores y profesoras información valiosa para sus decisiones docentes y ser un escenario excelente de aprendizaje colectivo, imprime seriedad al proceso, demuestra a los estudiantes la voluntad de perfeccionamiento profesional de sus profesores, refuerza la importancia de sus opiniones, los anima a participar y colaborar en futuras actuaciones, evidencia la utilidad de la evaluación y los compromete en las soluciones para mejorar la calidad de sus clases y de la Universidad en general.

4.1. Análisis de la experiencia

El alcance o incidencia del proceso evaluador en el cambio o mejora de las acciones didácticas del profesorado y en los procesos educativos y formativos de los estudiantes es la parte más compleja de la metaevaluación. Podemos presumir que el aprendizaje profesional empieza con la primera propuesta y compromiso del profesor o profesora con el proyecto, ya que el debate y análisis de una forma diferente de enfocar la evaluación ha de producir algún tipo de ruptura epistemológica de su pensamiento práctico docente, pero es más difícil evidenciar el grado, alcance y calidad del previsible cambio.

En esta línea, según manifestaron los enseñantes, las razones que les empujaron a someterse voluntariamente a la evaluación, estaban mayoritariamente relacionadas con la posibilidad que ésta les brindaba para comprender y mejorar su formación profesional en los aspectos didácticos. Un indicador que confirmó este propósito fue la demanda explícita de actividades de formación (sobre todo en cuestiones metodológicas y de evaluación de alumnos) que muchos de ellos realizaron a los evaluadores, después de la lectura de sus informes personalizados. Esto reforzaba la idea o filosofía inicial del plan. No obstante, contamos con poca información de los profesores que no solicitaron la evaluación, puesto que los contactos y las entrevistas se llevaron a cabo con los que lo hicieron. En opinión de estos últimos, hay muchos docentes que no quieren ser evaluados porque prevén resultados negativos, están inmersos en la desidia o se consideran autosuficientes profesionalmente. También aseguraban que muchos no la solicitaron por despiste u olvido y otros porque opinaban que ésta tenía poca trascendencia académica.

Lo que sí estamos en disposición de afirmar es que la experiencia, sin duda, ofreció, a los profesores y profesoras que participaron en la evaluación, oportunidades de aprendizaje profesional, facilitados mediante la realización de las tareas y procesos de discusión y análisis que pusimos en escena, entre los que destacamos por su potencial formativo cuatro:

- 1º. El contraste de las opiniones y valoraciones de los docentes con sus estudiantes sobre diversos aspectos de la vida cotidiana de las aulas universitarias, a veces no coincidentes.

El profesorado resaltaba el importante papel que la evaluación había jugado como agente retroalimentador, informándole sobre cómo percibían sus alumnos y alumnas su actuación y el desarrollo de las materias que impartían. Aseveraban que la opinión del alumnado les había hecho reflexionar sobre el futuro desarrollo de su asignatura, para insistir en los aspectos que su grupo señaló como positivos y modificar, en la medida de lo posible, algunas de las variables que se relacionan con su comportamiento y que pueden propiciar los aspectos que éstos concebían como menos gratos.

Este contraste no siempre ha de significar cambios en los planteamientos docentes, o sólo en éstos, sino que debe conllevar una reorientación de las actuaciones y concepciones de los estudiantes. Por ejemplo, cuando a éstos se les pedía que hicieran sugerencias para mejorar el funcionamiento de la asignatura evaluada, nos encontramos con infinidad de matices, dependiendo, lógicamente, de cada caso concreto; pero también con algunas peticiones que se repetían reiterativamente (menos teoría y más práctica, menos materia, mejores explicaciones, etc.). Era curiosa la ausencia de recomendaciones o comentarios que conllevaran cambios en sus compromisos discentes, individualmente o como grupo. Esto resultaba más llamativo al comprobar que, en general, valoraban altamente la actitud y actuación del profesorado, algo menos las suyas personales y percibían como precarias las del grupo/clase, lo que indica que siguen pensando en la enseñanza como una acción unidireccional donde ellos son el «recipiente a llenar». Esto reafirma la idea de que no debemos ceñirnos a la recogida de opinión sobre el docente, sino que es necesario abrir el foco de estudio y diversificarlo abarcando al cúmulo de interrelaciones que inciden en la vida académica, pues como reitera Santos Guerra. «Los profesores están inmersos en la organización escolar que, a su vez, está inmersa en un sistema educativo que condiciona el desarrollo profesional» (1993: 162).

- 2°. La indagación en los problemas y aspectos concretos que producen el conflicto o el desacuerdo en cada clase y en general (evaluación, metodología, recursos, etc.), lo que permite la toma de decisiones personales y de gestión.

Cuando se les preguntaba a los profesores por su opinión sobre los temas que los alumnos puntuaban más bajo en su asignatura concreta, mantenían generalmente una actitud opuesta a éstos; por ejemplo, no estaban de acuerdo con muchas de las críticas que se hacían al sistema de evaluación que empleaban o con las que se vertían sobre la metodología o con la relación y proporción entre la teoría y la práctica. Las actitudes del profesor en clase ante estos asuntos disonantes son factores que influyen directamente en la marcha de la asignatura, y en este sentido algunos docentes mostraban una postura inamovible en relación al modo en que se desarrollaban las clases, sin plantearse que un cambio en la metodología y contenidos, podría desembocar en dinámicas más motivantes. Claro está que esto no siempre es sencillo, y que uno de los factores limitadores eran las mismas actitudes contradictorias de los estudiantes que se evidenciaban de forma nítida en algunos asuntos como el mencionado de la metodología.

Los profesores universitarios no nos distinguimos precisamente por nuestras innovaciones pedagógicas. Pues bien, pudimos observar cómo los que intentaban dar un enfoque más dinámico a sus clases, introduciendo estrategias de trabajo diferentes a la enseñanza unidireccional o al dictado de apuntes, se encontraban, a menudo, con la resistencia de una parte importante de los estudiantes. Éstos, que reconocían en las entrevistas el mayor valor formativo de modos de enseñanza más implicativos y manifestaban en los cuestionarios sus preferencias por los aprendizajes comprensivos, se contradecían en la práctica presionando al profesorado para que no se saliese del patrón metodológico tradicional. Cuestiones como la *obsesión* por el aprobado a cualquier precio, la presión que sobre ellos ejerce la atomización y proliferación de asignaturas producida por los actuales planes de estudio, junto a otras razones de tipo sociológico (desvalorización del saber, tecnicismo, masificación, *titulitis*, *curriculosis*, etc.), estaban detrás de estas paradojas. La evidencia de estas contradicciones mediante la evaluación realizada, dio la posibilidad de su discusión en las clases de forma argumentada y rigurosa.

- 3°. La indagación en las cuestiones que inciden en la opinión de los estudiantes respecto al profesorado, a las asignaturas y las condiciones materiales de la clases.

En este sentido podemos destacar dos variables que explican las diferencias en sus actitudes para cada una de las categorías o aspectos estudiados —la evaluación, las relaciones, la participación, los contenidos, los recursos y la organización—: la satisfacción con los estudios que cursaban (a mayor grado de satisfacción con ellos, mejores valoraciones y viceversa) y las calificaciones obtenidas (cuanto más altas las notas, mejor opinión manifestaba el alumnado sobre prácticamente todos los aspectos por los que se les preguntaba). Es importante para que el profesorado evalúe adecuadamente las opiniones de los estudiantes, conocer estos factores y circunstancias que, junto a sus actuaciones docentes, inciden de alguna manera en las valoraciones de éstos, pues ello les aporta elementos de juicio para la toma de sus decisiones profesionales y para el debate de éstas con sus alumnos y alumnas.

- 4°. La triangulación de opiniones expertas, es decir, entre el evaluador/a y el profesor/a. Esto, evidentemente, anima el diálogo interprofesional sobre aspectos esenciales de la tarea docente que difícilmente se abordan o potencian en la Universidad.

El informe personalizado que recibieron los profesores y profesoras, fue calificado por ellos y ellas de necesario y oportuno. La mayoría consideró que «hace más amena la interpretación y muestra un análisis de los hechos al margen de los implicados». Muchos enseñantes manifestaron que el informe les brindó la oportunidad de conocer una visión de los datos al margen de la que ellos mismos podían hacer, y que esto les proporcionó mayores posibilidades de confrontar el juicio con el de otras personas con perspectivas diferentes.

Esa lectura profesional objetiva de los datos hecha por los evaluadores, transformados en un informe personalizado, actuaba de punto de contraste y control intersubjetivo de la que realizaba el propio profesor o profesora, convirtiéndose en un elemento de autorreflexión y reflexión compartida y de búsqueda colaborativa de vías y alternativas para ir superando los posibles problemas detectados y, sobre todo, ir avanzando en la mejora de las tareas docentes y discentes. El hecho de contar con didactas en este ejercicio reflexivo, potenciaba su efectividad.

Para concluir decir que, entre las limitaciones hemos de destacar, además del gran esfuerzo y tiempo necesario para poner en marcha procesos de estas características, dos:

En primer lugar, dejó fuera del circuito de mejora y participación a los desinteresados y a los escépticos, aunque a algunos de éstos últimos se podían ir recuperando. No obstante, con respecto a los primeros, no es un problema de evaluación, sino de concienciación profesional, que requieren soluciones y decisiones académicas y políticas de gran calado.

En segundo lugar, las peculiaridades de las organizaciones universitarias, con el gran peso de los intereses personales y de los grupos de presión y poder instalados en nuestro sistema, especialmente en algunas universidades, convierten en misiones casi imposibles los planteamientos y planificaciones de cambio y avance a medio y largo plazo, lo que entra en contradicción directa con propuestas basadas en procedimientos pausados y continuos de aprendizaje profesional docente.

BIBLIOGRAFÍA

- ÁLVAREZ MÉNDEZ, J.M. (2009). *Evaluar para conocer, examinar para excluir*. Madrid, Ediciones Morata.
- APODACA, P. (1992). *Métodos de investigación y diagnóstico en educación: Investigación evaluativa en educación*. I.C.E. de la Universidad del País Vasco.
- BOLÍVAR, A. (2008). Evaluación de la práctica docente. Una revisión desde España, en *Revista Iberoamericana de Evaluación Educativa*, 1 (2), 56-74.
- CARR, W. (Ed.) (1998). *Calidad de la enseñanza e investigación-acción*. Sevilla, Díada Editora.
- CISNEROS-COHERNOUR, E. Y STAKE, R. (2010). La evaluación de la docencia en educación superior: de evaluaciones basadas en opiniones de estudiantes a modelos por competencias, en *Revista Iberoamericana de Evaluación Educativa*, 3 (1), 218-231.
- DEAN, J. (1991). *Professional development in school*. Milton Keynes: Open University.
- DUKE, D. L. Y STIGGINS R. J. (1997). Más allá de la competencia mínima: evaluación para el desarrollo profesional, en J. Millman y Darlinh-Hammond (eds.): *Manual para la evaluación del profesorado*. Madrid, La Muralla.

- ELLIOTT, J. (1984): *Methodology and Ethics*, en C. Adelman, (Ed.): *The Politics and Ethics of Evaluation*. Croom Helm.
- (2009): *La investigación en la acción en educación*. Ediciones Morata, Madrid.
- FERNÁNDEZ SIERRA, J. (1996): «¿Evaluación? No, gracias, calificación», en *Cuadernos de Pedagogía*, 243, 92-97.
- (1999): «Evaluación de la docencia y aprendizaje profesional. Análisis de una experiencia universitaria», en *Revista interuniversitaria de Formación del profesorado*, 34, 87-98.
- (2002): *Evaluación del aprendizaje, evaluación del rendimiento*. Madrid, Akal.
- (2011): *Formar para la economía del conocimiento vs educar para la sociedad del conocimiento*. Málaga, Aljibe.
- FERNÁNDEZ, N. y COPPOLA, N. (2012): «Aportes para la reflexión sobre la evaluación de la función docente universitaria», en *Revista Iberoamericana de Evaluación Educativa*, 5 (1), 106-119.
- GIMENO SACRISTÁN, J. (1993): «El profesorado: ¿Mejora de la calidad o incremento del control?», en *Cuadernos de Pedagogía*, 219: 22-27.
- (2011). *Educación y convivir en la cultura global*. Ediciones Morata, Madrid.
- GIMENO SACRISTÁN, J. y PÉREZ GÓMEZ, A.I (1992). *Comprender y transformar la enseñanza*. Ediciones Morata, Madrid
- MACDONALD, B. (1978): *Democracy and Evaluation*. University of East Anglia (Multicopiado). CARE.
- (1985): «La evaluación y el control de la educación», en J. Gimeno Sacristán y A. Pérez Gómez: *La enseñanza, su teoría y su práctica*. Madrid, Akal, 467-478.
- MILLMAN, J. y DARLINH-HAMMOND (eds.) (1997). *Manual para la evaluación del profesorado*. Madrid, La Muralla.
- MONTERO MESA, L. (1990): «Comportamientos del profesor y resultados del aprendizaje: Análisis de algunas relaciones», en C. Coll, J. Palacios y A. Marchesi, (comps.): *Desarrollo psicológico y educación. Vol.II: Psicología de la Educación*. Madrid, Alianza Editorial, 249-271.
- (1993). El perfil del profesor universitario y su formación inicial. *III Jornadas Nacionales de Didáctica Universitaria. Evaluación y Desarrollo Profesional. Ponencias y Réplicas*. Universidad de Las Palmas de Gran Canaria, 213-227.

- OCDE (1991). *Escuelas y calidad de la enseñanza: Informe internacional*. Barcelona, Paidós-MEC.
- PÉREZ GÓMEZ, A.I. (2013). *Educarse en la era digital*. Ediciones Morata, Madrid
- Revista de Educación* (2008). *Tiempos de cambio universitario en Europa*. número especial.
- RODRÍGUEZ ESPINAR, S. (1993): Experiencias españolas de evaluación de la enseñanza universitaria y nuevas perspectivas. Otro punto de vista. En *Actas de las III Jornadas Nacionales de Didáctica Universitaria «Evaluación y Desarrollo Profesional»*. Las Palmas de Gran Canaria: Servicio de Publicaciones de la Universidad de Las Palmas de Gran Canaria, 111-132.
- SANTOS GUERRA, M.A. (1993). *La evaluación, un proceso de diálogo, comprensión y mejora*. Málaga, Aljibe.
- (2003). *La flecha en la diana. La evaluación como aprendizaje*. Madrid, Narcea.
- SCRIVEN, M. (1991): *Evaluation Thesaurus*. Newbury Park, CA: Sage.
- SIMONS, H. y ELLIOTT, J. (Eds) (1989). *Rethinking appraisal and assessment*. Milton Keynes: Open University Press, 100-118.
- SIROTNIK, K.A. (1990) (Ed.). *Evaluation and Social Justice: Issues in Public Education*. New York: Jossey-Bass Inc, Publishers (New Direction For Program Evaluation, nº 45, Spring).
- STAKE, R.E. (1989). «The Evaluation of Teaching», en H. Simons y J. Elliott, (eds.): *Rethinking Appraisal and Assessment*. Buckingham: Open University, 13-19.
- (2006): *Evaluación comprensiva y evaluación basada en estándares*. Barcelona, Graó.
- STRIKE, K. A. (1990): «The ethics of educational evaluation», en J. Millman y L. Darling-Hammond: *The new handbook of teacher evaluation*. Londres, Sage publications.
- TEJEDOR TEJEDOR, F.J. (2009). «Evaluación del profesorado universitario: enfoque metodológico y algunas aportaciones de la investigación», en *Estudios sobre Educación*, 16, 79-102.
- VROEIENSTI, T.I. Y ACHERMAN, H. (1991): «La evaluación de la calidad basada en el control frente a la evaluación basada en la mejora», en M. de Miguel y otros: *La evaluación de las instituciones universitarias*. Madrid, Consejo de Universidades.
- WILSON (1992). *Cómo valorar la calidad de la enseñanza*. Madrid, Paidós-MEC.



Juan del Rosal, 14
28040 MADRID
Tel. Dirección Editorial: 913 987 521

Los centros escolares y el profesorado se constituyeron, desde mediados de los años setenta del siglo XX, en dos nuevos ámbitos de evaluación. En su emergencia tuvo una influencia decisiva el movimiento de la *accountability*, o rendición de cuentas, que en el contexto anglosajón propició la obligación de que las instituciones educativas informaran periódicamente a la comunidad de sus logros y resultados, haciendo realidad el “derecho a saber”. A su implantación y desarrollo contribuyeron en gran medida las políticas educativas que abordaron, en los años noventa, la reforma de los sistemas educativos de cara a afrontar la mejora de su calidad.

En el contexto de estas nuevas exigencias y condiciones, no cabe duda que la teoría y práctica de la evaluación institucional y de los docentes han ido adquiriendo un protagonismo estelar, como base sobre la que sustentar y legitimar las directrices de las políticas educativas que a nivel nacional se ponen en juego, para fundamentar las decisiones que necesitan tomarse en la conducción de los sistemas educativos, o para proveer a los centros y profesores de herramientas y datos que, además de facilitarles el conocimiento de sus quehaceres, les permitan profundizar en sus desarrollos institucional y profesional.

A través de diez temas se hace un recorrido por los contenidos más relevantes del estado actual de estos dos ámbitos de evaluación. En una primera parte, temas 1 y 2, se analiza la conformación de este campo profesional de estudio e investigación y se clarifican sus bases teórico-conceptuales y los presupuestos de las distintas tradiciones científicas que fundamentan las diferentes propuestas y modelos de evaluación de centros y profesores.

Las partes segunda y tercera están dedicadas monográficamente al estudio de cada uno de los ámbitos. Tanto en los centros, temas del 3 al 6, como en el profesorado, temas del 7 al 10, se empieza con la presentación de las orientaciones y perspectivas desde las que se pueden abordar la evaluación, y se finaliza con las características y experiencias de la evaluación institucional y del profesorado en el contexto universitario.

