

**Las evaluaciones
del sistema
educativo**

Ángel Vázquez
Alonso
*Departamento de
Ciencias de la
Educación, Universidad
de las Islas Baleares*

Educació i Cultura
(2001), 14:
247-274

Las evaluaciones del sistema educativo

Ángel Vázquez Alonso

Departamento de Ciencias de la Educación, Universidad de las Islas Baleares

Introducción

El profesorado siempre ha desempeñado un papel (entre otros muchos) de juez respecto del trabajo y el desempeño del alumnado en el proceso de aprendizaje escolar, de modo que el significado de la evaluación en educación siempre se ha identificado, principalmente, con la evaluación de los aprendizajes del alumnado realizada por el profesorado. El rol del profesor como evaluador resulta obvio, pues el conocimiento del progreso del alumnado es imprescindible para guiar el aprendizaje, pero también resulta necesario para cumplir las exigencias administrativas de calificación, acreditación y certificación de estudios o la emisión de los títulos académicos. En este sentido, se puede decir que durante mucho tiempo y aún hoy, para muchos, evaluación en educación es sinónimo de calificación escolar individual.

Sin embargo, este significado básico y fundamental de la evaluación en educación se encuentra hoy día extraordinariamente ampliado a la evaluación de la práctica educativa, de los programas educativos, del sistema educativo y de la propia administración educativa. Abriendo el horizonte de la evaluación en educación, este trabajo pretende describir los múltiples ámbitos de la evaluación actual en el sistema educativo, y por subrayar esta multiplicidad, se ha elegido el título «evaluaciones», en plural, que pretende ir más allá y superar el sentido meramente unidimensional (personal) que ha tenido (y tiene) la evaluación educativa, cuando se la identifica con la calificación académica individual de los estudiantes.

Como es conocido, en su origen, la evaluación educativa estuvo estrechamente unida a la medición de las características individuales y, por tanto, ligada al desarrollo de la psicometría desde principios de siglo. El racionalismo en la evaluación aplicado por Tyler, propuso a los educadores definir los objetivos y recopilar los datos necesarios para determinar si los resultados eran satisfactorios, aproximación que dominó la práctica de la evaluación educativa a lo largo de las décadas de 1940 y 1950, y que en la actualidad aún tiene importancia. Desde el concepto de evaluación por objetivos de Tyler, hasta el moderno concepto de metaevaluación (evaluación de la evaluación), pasando por la rendición de cuentas, la evaluación continua, formativa, transaccional, iluminativa, de programas, de la calidad, etc., el concepto de evaluación ha ido evolucionando intensamente, haciéndose más complejo y superando la simple idea de evaluación como calificación y examen de alumnos, aunque ésta continua siendo un reto perenne de la enseñanza en el aula.

La necesidad de evaluar programas educativos planificados para el logro de un objetivo determinado fue un primer punto de inflexión que proporcionó un fuerte impulso y creatividad a las actividades sistemáticas de evaluación. La evaluación de programas educativos supuso, de hecho, una nueva etapa en el desarrollo de la evaluación a la cual contribuyeron, sin duda, los importantes recursos económicos destinados a dicha tarea por las administraciones educativas, entre las cuales fue pionera la estadounidense (Carballo, 1990). Se puede decir que en torno a la evaluación de programas se han desarrollado los diversos métodos de evaluación que hoy en día están vigentes en la práctica evaluadora. Como grandes paradigmas de evaluación subyacentes a todos ellos se consideran hoy los enfoques cuantitativos, centrados en un enfoque de sistematización y objetivización de datos, y preocupados por la fiabilidad y la validez de sus resultados estadístico numéricos, y los enfoques cualitativos, basados en la extracción inmediata y personal de los datos mediante el contacto del evaluador con los protagonistas de los hechos evaluados.

Además de estos dos grandes paradigmas, que hoy suelen utilizarse indistinta y conjuntamente, se han propuesto una diversidad de métodos de evaluación que se suelen agrupar en tres grandes grupos: conductistas, humanísticos y holísticos (Castillo y Gento, 1995). Los modelos conductistas encuadran desde la evaluación por objetivos de Tyler hasta el método de evaluación de la figura de Stake, una modernización del anterior basado en la información sobre antecedentes, transacciones y resultados, pasando por la planificación de Cronbach (unidades, tratamiento y operaciones), el método de contexto, entrada, proceso, producto de Stufflebeam o el método de Alkin (necesidades, planificación, instrumentación, progreso y resultados). Los modelos de corte humanístico se caracterizan por su atención al paradigma cualitativo, exclusivo en el modelo de crítica de Eisner (centrado en la interpretación crítica provocada por el evaluador sobre el contexto, los procesos emergentes y sus relaciones mutuas, así como el impacto) y mixto, en los modelos de contraposición de Rolf y Owens (opinión consensuada alcanzada bajo la dirección del evaluador) y de atención al cliente de Scriven (centrados en valorar la satisfacción de las necesidades del cliente). En los modelos holísticos el evaluador promueve la interpretación por los implicados de los datos evaluadores, son cualitativos y desplazan las finalidades de la evaluación desde los resultados hacia una visión global de los efectos producidos; en la propuesta de evaluación respondiente de Stake son las necesidades de los participantes, en la evaluación iluminativa es la comprensión del programa evaluado y en la evaluación holística de McDonald la interpretación de la acción para mejorarla.

Pero el cambio más definitivo del sentido de la evaluación se produciría cuando el foco de atención se desplazó de la valoración de los resultados individuales, para centrarse en las instituciones escolares y los sistemas educativos. Dicha reorientación supuso, en palabras de D. Kallen, «una transformación crucial en los objetivos, los métodos y la finalidad misma de la evaluación» (Kallen, 1990, p. 11). A partir de ese momento, relativamente cercano en el tiempo, los sistemas educativos mismos se convierten en objeto de evaluación.

Una idea del acelerado progreso que ha sufrido la conceptualización de la evaluación puede ejemplificarla la anécdota relatada por el investigador House: «*hace más de veinte años, cuando comencé mi carrera como evaluador, recogí todos los artículos que pude encontrar sobre este tema, los puse en una carpeta y los leí en un mes*» (House, 1992, p. 42).

Hoy día, cualquier búsqueda bibliográfica sobre el tema de evaluación arrojaría como resultado miles de referencias; sólo una severa limitación en el campo de búsqueda

podría hacer manejable el listado para un principiante, aunque otra cosa sería la lectura de algunos artículos de los más duros.

El progreso en el enriquecimiento del concepto de evaluación ha sido tal, que hoy día se puede afirmar, como un lugar común, que la evaluación sistemática ocupa un lugar central en todos los sistemas educativos. Como un ejemplo, la *International Encyclopedia of Educational Evaluation* (Walberg & Haertel, 1990), un manual de unas mil páginas y un centenar de contribuciones independientes, está dividido en grandes secciones entre las cuales se encuentran títulos tan sugestivos como enfoques y estrategias de la evaluación, realización de estudios de evaluación, evaluación del currículo, teoría de la medida, aplicaciones de medidas, exámenes y tests, metodología de la investigación y planificación y política educativa, que dan una idea del desarrollo adquirido por los métodos, las técnicas y los desarrollos en evaluación. Todo estos rasgos sugieren que, desde una perspectiva conceptual y académica, casi todo está prácticamente dicho y escrito sobre cuestiones de evaluación. Sin embargo, como en tantos otros temas escolares y educativos (enseñar, aprender, socializarse, ...etc.), el problema de la evaluación reside en llevar a la práctica las ideas y los modelos teóricos, de modo que la gestión de la evaluación constituye un reto permanente para las políticas educativas de todos los países, y mucho más en un contexto como el actual en los países desarrollados, donde la práctica totalidad de los jóvenes gozan de una escolarización básica hasta la edad laboral, y más allá.

El propósito de este artículo es ofrecer una panorámica sobre la situación actual de la práctica de la evaluación institucional en el marco del sistema educativo global, con referencia expresa a la situación y las experiencias recientes, en nuestro sistema educativo nacional y local.

La demanda social de evaluación

Aunque la delimitación anterior del tema a tratar aquí supone que no están entre los objetivos de este artículo debatir las distintas perspectivas y significados de la evaluación, con los que se supone al lector suficientemente familiarizado, ni entrar en el amplio asunto de la evaluación de los aprendizajes del alumnado, germen de todo el desarrollo evaluativo en educación, la profundización y los desarrollos experimentados por la evaluación requieren una referencia, aunque sea somera, a algunos puntos concretos de esas cuestiones.

Hoy en día, se puede decir que existe una demanda generalizada de evaluación. El crecimiento económico en la década de los 50 y 60 de los sistemas educativos occidentales generó el movimiento de evaluación como rendición de cuentas («*accountability*»), para justificar los gastos educativos a través de la consecución de los objetivos propuestos (eficacia). En su origen, la inspiración limitadamente economicista de la rendición de cuentas respondió a la necesidad de los gobernantes, para justificar socialmente las cuantiosas inversiones educativas. Ante la realidad de la educación como poderoso instrumento de desarrollo social, esta actitud se ha extendido a todos los ámbitos educativos, como un ejercicio de responsabilidad social, política y técnica de los administradores de la educación. Desde esta perspectiva, como un indicador de la creciente exigencia y mejora del sistema educativo, la rendición de cuentas ha producido un desplazamiento progresivo de la teoría y la práctica de la evaluación desde el simple control económico del sistema hacia el control de resultados y el control de los procesos educativos.

Por otro lado, la creciente demanda social de evaluación ha conducido a una situación inflacionaria en el uso del término y en la aplicación de trabajos, de modo que se

produce una multiplicidad de significados e interpretaciones y se corre el riesgo permanente de considerar evaluación a cualquier informe por modesto que sea. Por ello, puede ser conveniente precisar qué es evaluación y de que evaluación hablamos cuando nos referimos al concepto de evaluación institucional.

Los múltiples manuales han propuesto numerosas definiciones del término evaluación, para acomodarse a los contextos concretos y la intensa evolución aludida en los párrafos anteriores. No obstante, aunque con matices diferenciales, todas las definiciones recogen rasgos fundamentales sobre los que existe un consenso tácito, a saber, la evaluación como un proceso sistemático de valorar algo (un programa, una persona, una institución, etc.) con el objetivo de mejorarlo. Esta breve definición contiene ya implícito que se requiere un proceso de recogida y análisis de información relevante sobre un objeto y formular un juicio sobre su adecuación a un patrón o criterio previamente establecido, como base para la toma de decisiones: *«Evaluar es el proceso de obtener información y de su uso para formular juicios que a su vez se utilizarán para tomar decisiones»* (Tenbrink, 1983)

Aunque el propósito de este artículo no es profundizar en estos aspectos conceptuales, puede ser especialmente interesante precisar algunos de los rasgos de la evaluación que han sido desarrollados en los últimos tiempos, tales como la sistematicidad y el valor de una evaluación, por lo que se tratarán sucintamente.

La inflación producida por la demanda social de evaluación ejercida sobre el sistema educativo, aunque globalmente positiva por su generalización y extensión, puede estar produciendo también efectos paralelos espurios. Uno de estos efectos, y tal vez el más conocido, especialmente en ambientes administrativos, es el que lleva a denominar evaluación a cualquier documento informe o memoria que contiene datos, tiempo dedicado, usuarios atendidos, listas de actividades, participantes, resultados o actos realizados. De acuerdo con la definición anterior, para que una memoria de actividades se convierta en una verdadera evaluación se requiere que aporte un juicio de valor, que sea claro y sistemático.

La sistematicidad de una evaluación no es simplemente orden, precisión, fiabilidad o validez, sino que va más allá de meras cuestiones formales para adentrarse en el núcleo de las condiciones básicas que debe satisfacer cualquier evaluación para ser considerada como tal. La preocupación por la validez para determinar qué es y qué no es un informe legítimo de evaluación ha sido sentida desde siempre por los profesionales de la evaluación y sus consecuencias se han plasmado en manuales de estándares evaluativos, que cualquier evaluador debe conocer, aplicar y respetar (Joint Committee, 1988). En un manual como el citado se recogen todos los requerimientos que son exigibles a un informe de evaluación, aunque desgraciadamente, puede resultar imposible que todos los evaluadores los conozcan y los apliquen en sus prácticas, ofreciendo un amplio campo todavía para la formación y el perfeccionamiento profesional a todos los niveles. Para superar este inconveniente, como se verá más adelante, los sistemas educativos han respondido mediante la creación de organismos e instituciones especializadas en evaluación que garanticen la sistematicidad y la validez de las evaluaciones realizadas, especialmente aquellas que se consideran esenciales para tomar las decisiones administrativas y políticas requeridas para mejorar la educación como es el caso del Instituto Nacional de Calidad y Evaluación.

La asignación de valor al objeto evaluado es una cuestión esencial en cualquier evaluación; sin juicio de valor (valoración) no hay evaluación. Ningún documento puede ser considerado evaluador si no contiene una valoración clara, expresa y precisa del objeto

evaluado, y aunque son también necesarias otras condiciones, esta es esencial y básica. Las referencias o criterios para la emisión de la valoración son diversos, de modo que en función de ellos, hoy día se precisan y diferencian los conceptos de valía y mérito.

Cuando las referencias fundamentales de la valoración son las necesidades de partida, se dice que el objeto evaluado tiene (o no) valía. Con frecuencia existen distintos públicos que exigen o necesitan cosas distintas de un centro, programa, servicio, etc. Por ejemplo, las familias necesitan, por lo general, que las escuelas e institutos preparen a sus hijos para continuar los estudios inmediatos superiores, en tanto que, por ejemplo, las familias de colectivos inmigrantes pueden considerar prioritario, sobre todo, que los centros atiendan los retrasos, refuercen la enseñanza del idioma y promuevan la integración y la no discriminación social. Por tanto, a la hora de evaluar un centro, programa o servicio, debe valorarse la medida en que se tienen en cuenta las necesidades existentes en el entorno (altas prestaciones, atención a la diversidad, preparación para afrontar problemáticas especiales, etc.) en función de las distintas expectativas de los usuarios.

Cuando la referencia de la valoración son normas de excelencia basadas en la satisfacción por el objeto de ciertos requerimientos de calidad se habla de mérito. Cuando se evalúe un centro deben considerarse los requerimientos de excelencia (p.e. epistemológicos de las diferentes asignaturas) y, al menos, las más importantes prescripciones técnicas (pedagógicas, psicológicas, didácticas, organizativas) para valorar los resultados del aprendizaje y la actividad docente de los centros, con objeto de determinar el mérito de los mismos.

Stufflebeam y Shinkfield (1987) explican las diferencias entre valía y mérito mediante el siguiente ejemplo:

«Un programa para formar a profesores de Historia, ¿podrá producir profesores que enseñen Historia a otros de un modo efectivo y de manera que se pueda confiar en ellos? En general ¿hará bien lo que se supone que debe hacer? Si lo hace, será meritorio, pero puede que no sea válido. Por ejemplo, si la sociedad ya no necesita más profesores de Historia porque ya tiene más de los que puede emplear, entonces aunque el programa sea muy bueno y quizá sea mejor que otros parecidos continuar con el mismo nivel de expansión del programa quizá no sea la mejor manera de gastar sus escasos recursos».

El mérito es el valor de algo considerado en sí mismo, la valía es el valor de algo con relación a las necesidades que pretende o dice satisfacer (Luján y Puente, 1996). Algo es meritorio si tiene calidad, si hace bien lo que se supone que debe hacer, y tiene valía o es válido si satisface las necesidades para las que fue creado. Las relaciones entre ambos conceptos no son equiparables, pues uno incluye al otro. Así, un aprendizaje, programa, servicio o centro puede ser meritorio, por ejemplo, y al mismo tiempo puede no ser válido; pero para poder ser válido, debe ser meritorio. Desde esta perspectiva, el mérito es una condición necesaria para la validez, pero no es una condición suficiente, ya que el mérito no garantiza siempre la valía.

Otras ideas ingenuas cultivadas en los procesos de evaluación se refieren a las funciones formativa y sumativa de la evaluación. Frecuentemente se tiende a relacionar ambas, exclusivamente, con la temporalidad de la evaluación, de modo que se identifica evaluación sumativa con la evaluación aplicada cuando el programa u objeto evaluado está acabado o completo y evaluación formativa cuando ésta se aplica durante la realización de los programas evaluados. La simple clasificación anterior basada en el momento de la evaluación olvida otra condición importante para decidir una u otra, como son las

finalidades que cumplen en relación al programa evaluado (Chambers, 1994). Una evaluación formativa es capaz de influir el desarrollo del programa u objeto evaluado, de modo que una evaluación se considera formativa si y solo si es usada para mejorar o innovar los procesos, el diseño, la aplicación o la planificación del programa evaluado, con independencia de si se realiza durante o después del programa; por el contrario, una evaluación sumativa es un juicio definitivo que valora los méritos de un programa u objeto, de modo que una evaluación es sumativa si se aplica para fines diferentes a modificar el programa o su aplicación, también con independencia de si se realiza durante o después del programa. Por tanto, la temporalidad no garantiza la función de la evaluación, pues una evaluación realizada durante la ejecución de un programa puede ser sumativa y una evaluación realizada después de concluir el programa puede ser formativa. En general, todas las evaluaciones contienen elementos formativos y sumativos en su planificación, de acuerdo con la función asignada intencionalmente, lo cual complica la distinción o clasificación en uno u otro tipo. Además, está la cuestión de la distancia entre planificación y realización, donde también pueden encontrarse diferencias formativa / sumativa; en la práctica, una evaluación planificada como formativa puede convertirse en sumativa en su realización, si las intenciones originales no se llevan a la práctica fielmente, y viceversa, una evaluación intencional sumativa puede convertirse en formativa en su aplicación práctica.

Los sistemas de evaluación institucionales

La sociedad es cada vez más exigente con el sistema educativo, con su funcionamiento, los logros que alcanza y la calidad del servicio que presta a los usuarios. Por otro lado, las autoridades educativas necesitan conocer la situación real del sistema educativo, más allá de las intenciones o propuestas legislativas. Además, en la última década, muchos países han emprendido reformas que obedecen a diferentes motivaciones y atienden a diferentes estilos y fundamentos psicopedagógicos, lo cual agudiza la necesidad de las autoridades de conocer y diagnosticar los efectos de las reformas. Por ello, existe un movimiento general de los gobiernos en muchos países para crear mecanismos e instrumentos que les permitan dar respuesta a sus necesidades en la toma de decisiones educativas y, al mismo tiempo, ofrecer a la ciudadanía una información fiable y válida sobre el sistema educativo.

Estas preocupaciones y necesidades generales tienen aplicaciones diferentes en los diversos países, que pueden ir desde evaluaciones internas (autoevaluación) hasta evaluaciones externas, abiertas o cerradas, dirigidas o participativas, centralizadas o delegadas, etc. como reflejo de las diferentes realidades sociopolíticas en cada uno de ellos. La muestra ofrecida en el monográfico del número 321 de la *Revista de Educación* (2000) sobre los sistemas de evaluación empleados en algunos países permite una aproximación actual a esta cuestión.

Tras los antecedentes de diversas evaluaciones educativas nacionales en los años 50s, la participación en un estudio de la IEA en 1970 y la realización de una evaluación independiente en 1980, Hungría ha establecido desde 1986 un proceso nacional de evaluación sistemática denominado Monitor (Vari, 2000). Monitor se centra en evaluar la capacidad de aprendizaje de las habilidades culturales básicas, que en Hungría están centradas en la comprensión escrita de la lengua, las matemáticas y las habilidades informáticas, juzgados en función de estándares nacionales o expectativas de aprendizaje. La metodología contempla la inclusión de cuestiones en cadena (las mismas preguntas para

diferentes edades) y cuestiones en puente (las mismas cuestiones aplicadas en ediciones de la evaluación sucesivas en el tiempo). Además se incluyen mediciones de diversos indicadores del entorno social y familiar de los estudiantes.

El gobierno irlandés (Cosgrove, 2000) tiene declarado desde 1995 como objetivo prioritario la eliminación del analfabetismo y con el precedente de la realización de diversos estudios de evaluación nacionales y la participación en estudios internacionales (IEA, IEAP, TIMSS), ha establecido desde 1997 un sistema cíclico y sistemático de supervisión sobre los resultados en determinados cursos (11 y 15 años) y áreas curriculares (irlandés, inglés y matemáticas) del sistema educativo, junto con la participación simultánea en el programa INES de la OCDE.

La evaluación del sistema educativo en México se realiza mediante un plan nacional, desde dos perspectivas básicas y complementarias, una interna, a cargo de los responsables de otorgar el servicio educativo, y otra externa, a cargo de la administración educativa (Martínez, 2000). Este plan pretende ofrecer rigor técnico y fiabilidad para sus resultados, huir de la percepción clásica de la evaluación fiscalizadora y punitiva y trata de arraigar las prácticas evaluadoras en las escuelas (evaluación interna).

El programa sobre indicadores de rendimiento educativo es el nombre que recibe el proyecto de evaluación educativa en Canadá iniciado en 1989, consensuado nacionalmente, para todos sus territorios y provincias (Brackenbury, 2000). El programa se basa en un sistema de indicadores, agrupados en cinco grandes niveles de rendimiento y similares al proyecto PISA de la OCDE, contra el cual se comparan los resultados obtenidos por pruebas específicas de lengua (inglés y francés), matemáticas y ciencias, que se aplican al alumnado de 13 y 16 años desde 1996. Además existen programas específicos de evaluación en cada una de las provincias autónomas.

En el marco de los objetivos de equidad y calidad para el sistema educativo, el gobierno federal de Brasil promueve, desde 1990, un sistema de evaluación de la educación básica (grados 1 a 11), aplicados bianualmente a los grados 4, 8 y 11. El sistema mide entradas, procesos y productos con cuestionarios de rendimiento, hábitos de estudio, demográficos, del profesorado y para la escuela, que se divulgan a nivel nacional y sirven para centrar los debates para la mejora y nuevos enfoques del sistema educativo y consolidar el papel de liderazgo en la educación del gobierno federal (Guimaraes, Gómes y Schulmeyer, 2000).

La reforma educativa de 1994 en Bolivia incluye la puesta en marcha de un sistema de medición y evaluación de la calidad de la educación centralizado, enmarcado en uno de los ejes de la reforma que es la interculturalidad y el bilingüismo y que pretende obtener indicadores de rendimiento escolar y otros factores asociados (Barrera, 2000).

Estos ejemplos muestran bien a las claras la necesidad que sienten las administraciones de evaluar sus sistemas educativos para tomar las mejores decisiones de planificación en orden a la mejora de la calidad o para conseguir el mejor desarrollo posible de sus inversiones.

Los indicadores educativos

Dentro de este contexto de demanda de evaluación y desarrollo acelerado de planes de evaluación de los sistemas educativos se ha planteado la cuestión de la mayor eficacia en el uso de la información suministrada por la evaluación para la toma de decisiones. La idea de manejar un único indicador o un número reducido, aunque sólo sea para describir

aspectos concretos parece descartada, porque puede dejar sin reflejar aspectos importantes del sistema, mientras que un número excesivo puede hacer poco viable la obtención de la información relevante y necesaria, así como su interpretación. Hoy día se trabaja más bien en construir conjuntos de indicadores que permitan comparaciones transversales (entre sistemas educativos diferentes) o longitudinales (entre dos momentos de un mismo sistema), así como realizar un seguimiento fiable de las tendencias evolutivas de los sistemas educativos.

Un indicador es un rasgo observable que describe información relevante de un sistema, tanto cualitativa como cuantitativa. Esto lleva a la cuestión de definir y seleccionar con precisión los indicadores más relevantes de un sistema educativo, es decir, aquellos rasgos capaces de ofrecer la información más significativa sobre el sistema. Para que un indicador pueda ser considerado tal debe cumplir una serie de requisitos de constructo (contenido), de medida y de forma algunos de los cuales han sido sugeridos por Lázaro (1994). Los requisitos de forma son condiciones que permiten una buena definición del indicador y hacen, referencia a propiedades como selección consensuada entre agentes, claridad, concreción, unidad (se refieren a una sola entidad), precisión, brevedad, formalización (desglosable en ítemes más concretos) y dependencia (relaciones definidas con otros). Los requisitos de medida se refieren a las condiciones que permiten su medición de campo válida y fiable, tales como la observabilidad, la aplicabilidad, la accesibilidad, la codificación (cualitativa o numérica), la ponderación (dentro del sistema de indicadores) y la relacionabilidad con otros indicadores (mediante índices). Los requisitos de constructo se refieren al contenido de un indicador, que debe ser reflejo de una propiedad del objeto (relevancia), significativa (significatividad), relacionado con los fines y objetos de la evaluación (teleologicidad), estabilidad, capacidad predictiva y coherencia respecto a todo el sistema.

Para la obtención de indicadores se requieren fuentes de datos fiables y periódicas, de modo que los indicadores sirvan para una mejor comprensión del sistema y, al mismo tiempo, constituyan una base válida para la toma de decisiones educativas. La necesidad de coherencia del sistema de indicadores es una dificultad importante, puesto que hoy en día todavía, la construcción de indicadores no ha sido aún respaldada por un modelo teórico riguroso, de modo que los indicadores actuales son, fundamentalmente, parámetros operativos en un marco de referencia contexto-recursos-escolarización-procesos-resultados.

Como fruto de esta preocupación, algunos países han elaborado indicadores nacionales, de los cuales son buena muestra algunos de los proyectos de evaluación emprendidos por los países citados en párrafos anteriores y otros más antiguos (OFS, 1993; UVM, 1993). Pero en el ámbito internacional, y con el objetivo de permitir una comparación entre diversos países, se ha alumbrado por la OCDE el proyecto internacional de construcción de un sistema de indicadores educativos internacionales (proyecto INES). Este sistema de indicadores se basa en cuatro redes denominadas resultados de los alumnos, educación y mercado de trabajo, funcionamiento de las escuelas y el sistema educativo, y actitudes y expectativas. Dentro de cada una de ellas se definen una serie de indicadores específicos que forman el sistema de indicadores educativos y cuya relación sería muy prolija aquí (OCDE, 1992) pues han tenido una intensa evolución desde esa primera edición hasta la última (OCDE, 1997). Esta última edición presenta en total 41 indicadores agrupados en 5 dimensiones, denominadas indicadores de contexto demográfico, económico y social (3 indicadores), indicadores de recursos económicos y humanos (8), indicadores de escolarización y participación en la enseñanza (7), indicadores de procesos i organización escolar (13) e indicadores de resultados (10).

Así mismo, la Dirección General para la Educación y la Cultura de la Comisión Europea ha elaborado el informe europeo sobre la calidad de la educación (EC, 2000) que presenta 16 indicadores de calidad agrupados en cuatro familias (resultados, éxito y transición, control de la educación y recursos y estructuras). El informe se basa en el tarea del Comité de Trabajo sobre indicadores de calidad y los indicadores propuestos sobre la calidad de la educación en las escuelas están resumidos en la tabla 2. Casi la mitad de estos indicadores se refieren a los resultados de logro en cada una de 7 áreas de aprendizaje, y es de destacar que algunas de las áreas incluidas en la lista resultan bastante novedosas en relación a nuestro currículo, como las tecnologías de la información y la comunicación, el aprender a aprender y el civismo. Al contrario que en España, parece que en Europa no se vive nuestra particular reforma de las humanidades, pues ninguna de las áreas explicitadas está directamente relacionada con la humanidades.

En el plano doméstico, el Instituto Nacional de Calidad y Evaluación (INCE) ha publicado recientemente un sistema estatal de indicadores de la educación (INCE, 2000) que está resumido en la tabla 1. Los indicadores se agrupan en cinco familias, indicadores de contexto (5 indicadores), de recursos (5), de escolarización (7), de procesos y de resultados educativos (7).

El *Consell Superior d'Avaluació del Sistema Educatiu de Catalunya* ha constituido su propio sistema de indicadores, a medio camino entre las cinco dimensiones del programa INES de la OCDE y las del INCE, a saber, indicadores de contexto (5 indicadores), recursos (8), indicadores de procesos y escolarización (11) e indicadores de resultados (6), en total 30 indicadores.

La Fundación de las Regiones Europeas para la Investigación en Educación y Formación (FREREF), una red interregional europea, ha realizado un estudio para identificar las competencias básicas de un estudiante cuando acaba el período de escolaridad obligatoria que podría ser homologable con los estudios de indicadores pero su metodología es radicalmente diferente. El estudio ha consistido en una encuesta de opinión social, basada en un cuestionario formado por un inventario de competencias, donde las personas encuestadas valoraban la importancia de las distintas competencias incluidas en el cuestionario (245 unidades). Las clasificaciones de las distintas competencias entre los distintos ámbitos son discutibles, en algunos casos, pues podrían asignarse a ámbitos diferentes a los que efectivamente se han asignado en el estudio, cuestión hasta cierto punto normal en una lista tan extensa de competencias. El resultado ha sido una clasificación de las competencias sometidas a escrutinio desde la percepción social de su importancia relativa. En general, las competencias del ámbito laboral son las mejor valoradas, seguidas por las de los ámbitos social y lingüístico; en último lugar de importancia se sitúan las competencias de los ámbitos tecnocientífico y matemático (FREREF, 2000).

El carácter inevitablemente reduccionista de los sistemas de indicadores, como muchos otros aspectos de la evaluación, los hace vulnerables a las críticas. De hecho, el aspecto de control que supone toda evaluación sobre las escuelas se traslada a la filosofía de los sistemas de indicadores, como expresión de ideologías de mercado social, que se consideran ajenas a la cultura escolar, y los valores de la práctica educativa (Elliot, 1992). Sin embargo, el movimiento para avanzar en la estandarización de los indicadores de calidad de un sistema educativo es tan fuerte y consolidado, pues goza del apoyo de las autoridades educativas de los países más desarrollados del mundo, con lo que eso supone de aporte de fondos económicos y continuidad de los proyectos, que parece un camino irreversible y sin retorno, por encima de las legítimas críticas que también puedan merecer los resultados.

La evaluación del sistema educativo español

El primer informe serio y global de la educación española, aunque no cabe calificarlo como una evaluación sistemática, fue realizado en 1976 por una comisión evaluadora de expertos, con el objetivo de realizar un seguimiento al proceso de implantación de la Ley General de Educación (LGE) promulgada en 1970. Conducida como un proceso de reflexión y análisis de los indicadores macroeducativos, más que como una evaluación propiamente dicha, este informe ocupó tres volúmenes que no llegaron a publicarse y donde se señalaban los logros (incremento de la escolarización y expansión de la educación pública) y las deficiencias (falta de gratuidad plena de la enseñanza básica y prematura diferenciación del alumnado a los 14 años en dos vías estancas) observadas durante el período de seis años transcurridos de aplicación de la LGE.

Hasta la década de los 90s se puede decir que no hubo ninguna evaluación sistemática del sistema educativo, con la excepción de la denominada «evaluación de la reforma», aunque cabe notar algunos instrumentos valiosos y periódicos que contienen información fiable, pero muy global, de su funcionamiento general. Entre estos instrumentos destacan los informes anuales del Consejo Escolar del estado (desde mediados de los 80s) y las memorias anuales de la Inspección de Educación. También cabe dejar constancia del impulso de investigaciones educativas por organismos como el Centro Nacional para la Investigación y el Desarrollo Educativo (CENIDE) y sus sucesoras, Instituto Nacional de Ciencias de la Educación (INCIE), el Centro de Investigación y Documentación Educativa y el Centro de Investigación, Documentación y Evaluación (CIDE). La evaluación de la reforma fue realizada por el CIDE y su objetivo era valorar una experiencia de reformas que unificaban el ciclo superior de la EGB y las enseñanzas medias y que fueron ensayadas en un número restringido de centros en los últimos años de los 80s. El estudio se centró en el análisis del rendimiento educativo, actitudes y expectativas hacia el estudio de los adolescentes en ese tramo de escolaridad (CIDE, 1988, 1990 y 1992).

También algunas de las comunidades autónomas con competencias educativas plenas realizaron algunos pasos en este sentido de evaluación durante aquella época.

Con todo, estos datos reflejan bien a las claras, por sí mismos, la escasa incidencia de la evaluación sistemática sobre el sistema educativo español, y por otro lado, especialmente, si se comparan con el estado de la cuestión en otros países de nuestro entorno. Tampoco es de extrañar este resultado, en tanto en cuanto, ninguna de las normativas educativas contemplaba la evaluación del sistema educativo como una actividad normativizada. Se puede decir que, prácticamente todos los proyectos relevantes de evaluación del sistema educativo, tienen su origen en las normas básicas donde la evaluación se reconoce como una actividad significativa. Por ello, en primer lugar, se revisan brevemente los contenidos de estas normas para exponer después los desarrollos más importantes a que han dado lugar.

La normativa española sobre evaluación

La normativa básica vigente en nuestro país está contenida, fundamentalmente, en las dos leyes orgánicas que regulan la ordenación general del sistema educativo, la Ley Orgánica 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo (LOGSE) y la Ley Orgánica 9/1995, de 20 de noviembre, de la participación, la evaluación y el gobierno de los centros docentes (LOPEG).

La LOGSE estableció (artículo 62) las líneas generales a seguir en el ámbito de la evaluación. Ésta debería orientarse a la permanente adecuación del sistema educativo a las demandas sociales y a las necesidades educativas y debería aplicarse, sobre los alumnos, los procesos educativos, el profesorado, los centros y sobre la propia Administración teniendo en cuenta en cada caso el tipo del centro.

La LOPEG desarrolla las prescripciones anteriores para cada uno de los niveles y sectores del sistema educativo. Para garantizar un enseñanza de calidad se indica que los poderes públicos establecerán procedimientos para evaluar el sistema educativo, los centros, la labor docente, los cargos directivos y la propia administración educativa. En el caso de la evaluación de los centros docentes se configuran dos modalidades, interna y externa. Los centros evaluarán su propio funcionamiento mediante una evaluación interna, que se llevará a cabo por los órganos de cada centro, al final de cada curso, de acuerdo con lo preceptuado por la Administración educativa de la que dependan. Normalmente, esta evaluación se concreta a través de la elaboración de la preceptiva memoria final de cada curso, que aprueba el Consejo Escolar.

La evaluación externa de los centros será diseñada por la Administración educativa correspondiente mediante planes de evaluación específicos que serán aplicados con periodicidad a los centros docentes sostenidos con fondos públicos y que se llevarán a cabo principalmente a través de la inspección educativa. En la evaluación externa de los centros colaborarán los órganos colegiados y unipersonales de gobierno, así como los distintos sectores de la comunidad educativa (profesores, padres, alumnos y personal de administración y servicios). La evaluación de los centros deberá tener en cuenta el contexto socioeconómico de los mismos y los recursos de que disponen, y se efectuará sobre los procesos y sobre los resultados obtenidos, tanto en lo relativo a organización, gestión y funcionamiento, como al conjunto de las actividades de enseñanza y aprendizaje.

Las Administraciones educativas informarán a la comunidad educativa y harán públicos los criterios y procedimientos que se utilicen para la evaluación de los centros, así como las conclusiones generales que en dichas evaluaciones se obtengan. No obstante, se comunicará al Consejo Escolar las conclusiones de la evaluación correspondiente a su centro. Las Administraciones educativas colaborarán con los centros para resolver los problemas que hubieran sido detectados en la evaluación realizada.

A fin de mejorar la calidad educativa y el trabajo de los profesores, las administraciones educativas elaborarán planes para la valoración de la función pública docente. En la valoración de la función pública docente deberán colaborar con los servicios de inspección los órganos unipersonales de gobierno de los centros y, en los aspectos que específicamente se establezcan, podrán colaborar los miembros de la comunidad educativa que determina la administración correspondiente. En todo caso, se garantizará en este proceso la participación de los profesores.

Paralelamente, se han creado órganos superiores de la administración encargados de realizar una evaluación externa del sistema. Como órgano evaluador nacional la LOGSE diseña el *Instituto Nacional de Calidad y Evaluación* (INCE), que realizará la evaluación general del sistema educativo mediante el desarrollo de las actividades relativas a la evaluación de los alumnos, los procesos educativos, el profesorado, los centros y sobre la propia Administración. El Instituto Nacional de Calidad y Evaluación ofrecerá apoyo a las Administraciones educativas que lo requieran en la elaboración de sus respectivos planes y programas de evaluación. El Gobierno hará públicas periódicamente las conclusiones de interés general de las evaluaciones del sistema educativo efectuadas por el Instituto Nacional de Calidad y Evaluación y dará a conocer los resultados de los indicadores de calidad establecidos.

El Decreto 928/1993 regula el INCE y le atribuye explícitamente una serie de funciones entre las que destacan elaborar sistemas estatales de indicadores educativos para evaluar la eficacia y eficiencia del sistema educativo, diseñar planes de evaluación para las diferentes enseñanzas reguladas en la LOGSE y sus centros, y evaluar la estructura, alcance y resultados de las innovaciones de carácter general en el sistema educativo. Además, cooperar e intercambiar información con las administraciones educativas, proporcionar información al Consejo Escolar del Estado, informar a la sociedad, publicar los resultados de sus estudios y coordinar la participación española en estudios internacionales. El INCE actúa en el nivel estatal del sistema educativo, y por ello su ámbito de actuación son las enseñanzas mínimas del currículo, la elaboración de planes de evaluación y la propuesta de iniciativas y sugerencias. El INCE tiene a su frente un Consejo Rector (decisorio), un director y un consejo científico (consultivo), que concretan sus actividades en planes de actuación donde se contienen los diferentes estudios. Dentro de la evaluación permanente del sistema educativo el INCE ha realizado tareas de evaluación de la educación primaria, secundaria (1997 y 2000) y la elaboración de un sistema de indicadores estatales; además, el INCE promueve actividades continuas en el campo de la construcción de pruebas de rendimiento, la formación de especialistas en educación, y estudios específicos como la educación de las lenguas extranjera y el diagnóstico de la función directiva de los centros docentes.

En el ámbito autonómico el Govern Balear ha creado (Decreto 145/2000) el *Institut d'Avaluació i Qualitat del Sistema Educatiu de les Illes Balears* (IAQSEIB) como órgano de coordinación, de consulta i técnico, encargado de proporcionar información relevante sobre el grado de calidad del sistema educativo balear referida a su grado de eficacia, consecución de objetivos educativos, mejora de la calidad de la enseñanza y colaboración con el INCE y otros órganos similares de otras administraciones. Su organización se basa en un Consejo Rector, un Comité Científico, el director y el secretario técnico.

El plan de evaluación de centros (EVA)

Los reglamentos orgánicos de centros docentes emanados de la LOGSE encomiendan diversas actividades evaluadoras a los propios centros, tales como el análisis del grado de cumplimiento de la Programación General Anual, la evaluación de los Proyectos curriculares, de los procesos de enseñanza y de la práctica docente. Corresponde tanto a los equipos directivos como a los órganos colegiados (consejos escolares y claustros de profesores) dedicar una atención preferente a la planificación y realización de las tareas de evaluación interna. Al comienzo de la década de los 90s y con la finalidad de propiciar la evaluación interna en los centros docentes, el Ministerio de Educación y Ciencia impulsó un plan de evaluación de centros, más tarde denominado Plan EVA. En líneas generales, el Plan EVA se identifica con la definición de evaluación como: « ... *proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y el mérito de las metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados* » (Stufflebeam y Shinkfield, 1987).

El Plan EVA defiende la idea de una evaluación formativa, para ayudar a los centros docentes a superar las dificultades de sus propios procesos de evaluación interna y a apoyar y estimular los que hayan sido iniciados, aunque se trata de un proceso de

evaluación normativa externa, dirigido por un equipo de inspectores y con la participación y audiencia de todos los agentes y sectores implicados en el centro. En coherencia con esa finalidad general, el Plan EVA pretende:

- a) Impulsar o iniciar procesos de evaluación interna de los centros.
- b) Facilitar apoyo técnico a los centros para que puedan desarrollar sus procesos de evaluación interna.
- e) Difundir una cultura evaluadora de los centros docentes.

El Plan EVA integra enfoques metodológicos cualitativos y cuantitativos en un clima de entendimiento y de colaboración entre los evaluadores y los diversos participantes del centro. Debido al papel formativo del Plan EVA, los evaluadores se integran en la dinámica habitual del centro.

El Plan EVA se estructura en cinco módulos, elementos contextuales y personales, proyectos, organización y funcionamiento, procesos didácticos y resultados, diseñados de forma que cada módulo sea autosuficiente, y por tanto, facilite su aplicación independiente. Los aspectos básicos del diseño escolar se encuentran en los módulos 1 y 2 («Elementos contextuales y personales» y «Proyectos»). Los aspectos de planificación se encuentran, en parte, en el citado módulo 2 y en el módulo 3 («Organización y funcionamiento»). Los aspectos de la fase de ejecución se encuentran, en parte, en el módulo 3 y en el módulo 4 («Procesos didácticos»). Finalmente, los aspectos correspondientes al producto se encuentran dentro del módulo 5 («Resultados»).

Los indicadores que constituyen la base de la evaluación se orientan a obtener información sobre el grado en que los centros poseen determinadas cualidades relacionadas con la valía, el mérito y la viabilidad de los ámbitos en que se suelen estructurar las operaciones regulares de un centro docente. No obstante, el Plan EVA al presentar los resultados de la información recopilada mediante fichas de evaluación, construidas sobre la base de escalas descriptivas, permite analizar fácilmente los resultados obtenidos y trazar el perfil de los resultados conseguidos en cada una de las subdimensiones contempladas en el Plan EVA, así como en su totalidad; de este modo, los resultados reflejan los puntos débiles y los puntos fuertes, lo que resulta muy importante para tratar de mejorar los centros evaluados.

Se han elaborado los siguientes tipos de instrumentos:

- Guías para el análisis de la documentación del centro
- Guías para reuniones
- Guías de observación
- Cuestionarios

Cada instrumento consta de tantos apartados como dimensiones hay en las que se ha previsto la utilización de ese instrumento. En ocasiones los apartados se han desglosado en subapartados temáticos para facilitar la utilización del instrumento.

Las Guías elaboradas son múltiples, para analizar los siguientes documentos:

- Proyecto educativo del centro
- Programación General anual
- Proyecto curricular de las diferentes etapas
- Documentación de los Equipos de ciclo (Educación Primaria)

- Documentación de los Departamentos (Educación Secundaria)
- Trabajos del alumnado
- Documento de Organización del Centro (D.O.C.)

Las Guías de entrevistas elaboradas se enumeran a continuación:

Equipo directivo
Representantes del profesorado en el Consejo Escolar
Representantes de los padres y madres en el Consejo Escolar
Comisión de Coordinación Pedagógica
Equipos de ciclo (Educación Primaria)
Departamentos didácticos (Educación Secundaria)
Departamento de Orientación (Educación Secundaria)
Departamento de Actividades extraescolares y complementarias (Educación Secundaria)
Junta de Delegados (Educación Secundaria)

Se han elaborado las siguientes guías para la observación:

Guía de observación de las instalaciones del centro.
Guía para la observación sistemática en el aula.

Además, se han elaborado los siguientes cuestionarios individuales:

- Cuestionario al equipo directivo
- Cuestionario al profesorado
- Cuestionario al alumnado (Educación Secundaria)

La aplicación del Plan EVA a cada centro requiere la elaboración de un Plan de actuación específico, la realización de las actuaciones planificadas y la redacción del Informe final, que se somete a un proceso de consenso y publicidad interna y es la base de los futuros compromisos del centro.

En el Plan EVA no se aplican pruebas externas de rendimiento para evaluar resultados del aprendizaje lo que, evidentemente, constituye uno de sus inconvenientes más importantes, pues renuncia a la descripción e interpretación completa de los productos del centro.

Planes Anuales de Mejora

El cambio de gobierno de 1996 trajo un cambio en la orientación de los procesos de evaluación institucionales promovidos por la administración educativa. El nuevo Ministerio de Educación y Cultura impulsó un movimiento en favor de la calidad en la gestión de los centros docentes, cuyo principal objetivo es contribuir a la mejora de los resultados de nuestro sistema educativo mediante el desarrollo de un conjunto de políticas centradas en la institución escolar como referente fundamental. En aquel contexto, la administración educativa hace llegar a los centros docentes una percepción clara de cuáles son sus expectativas con respecto a la mejora de la calidad; recursos, instrumentos y

competencias para llevar a cabo los procesos de planificación, su desarrollo y su evaluación; métodos para promover la mejora continua; un seguimiento de la evolución de las instituciones escolares en sus programas de evaluación y seguimiento técnico necesarios para el diagnóstico de las áreas de mejora, la concepción del Plan y su evaluación interna. El marco de dicho movimiento integra, con un cierto grado de coherencia, las dimensiones epistemológica, ética, pragmática y metodológica. La definición explícita de dicho marco clarifica los fundamentos, facilita la reflexión sobre su contenido y sobre sus consecuencias, otorga sentido a las acciones tanto individuales como institucionales y orienta el desarrollo de los procesos de mejora.

En octubre de 1996 el Ministerio de Educación y Cultura publicaba los Planes Anuales de Mejora para los centros públicos, donde se articulaban, junto con los elementos metodológicos típicos de la planificación para la mejora, la cooperación entre los centros de enseñanza y la administración educativa, sobre la base de la idea de pacto o compromiso entre instituciones. Los centros docentes que se sumaran a la iniciativa deberían comprometerse a mejorar de un modo ordenado y sistemático, así como a abrirse voluntariamente a la ayuda de la administración y a la realización, por parte de ésta, de una evaluación final externa. De otro lado, la administración educativa, a través de sus Direcciones Provinciales, proporcionaría un apoyo de tipo técnico, mediante el asesoramiento en cualquiera de las fases de desarrollo del Plan Anual de Mejora; de tipo humano, por efecto de una mayor atención a sus esfuerzos; y de tipo económico, en función de las características del Plan y de la calidad de su aplicación. Además, desarrollaría un conjunto de actuaciones de reconocimiento hacia los colegios e institutos que se implicasen en procesos metódicos de mejora (tabla 3).

El análisis de la implantación de planes de mejora muestra un importante crecimiento del número de centros participantes que, en dos años, se ha visto triplicado; lo cual, habida cuenta del carácter voluntario de la implicación y del desigual comportamiento de las provincias constituye un indicador significativo del grado de aceptación de esta nueva orientación.

El diagnóstico previo al diseño de un Plan de Mejora y la identificación de cuáles son los puntos fuertes y los puntos débiles de un centro educativo requiere el uso de una herramienta de autoevaluación, que en este caso es el Modelo Europeo de Gestión de Calidad (MEGC), que combina, de forma ponderada, el interés por las personas con la importancia de los recursos, de los procesos y de los resultados y se beneficia de una invariancia de escala, esto es, puede ser aplicado a organizaciones de cualquier propósito y de cualquier tamaño y también a los propios individuos. Existe una tipología muy diversa de centros educativos públicos que comprende desde Escuelas Unitarias en el medio rural hasta Institutos de Educación Secundaria de más de dos mil alumnos; o desde Escuelas de Educación Infantil hasta Escuelas de Arte o Conservatorios de Música; y el MEGC, por sus características, puede ser aplicado a cualquiera de ellos.

Razones tanto de coherencia como de utilidad llevaron al Departamento a adoptar, de forma experimental, dicho MEGC con el fin de ofrecérselo a los centros como un instrumento sistemático de evaluación interna. A lo largo del curso 1996-97 se efectuó su adaptación a los centros educativos públicos, se editaron los materiales correspondientes (MEC, 1997) y se distribuyeron a todos los centros del ámbito territorial de gestión del Ministerio, el cual creó en 1998 el premio a la calidad en educación, de ámbito estatal, con el objetivo de promover la calidad en la educación a través de la mejora en la calidad de la gestión de los centros y de promover el MEGC como instrumento de autoevaluación para la mejora.

Los proyectos de evaluación transnacionales

Desde hace años, la investigación educativa viene produciendo numerosos estudios específicos de evaluación cuyo nivel va más allá de una simple aula o centro y que evalúan múltiples aspectos específicos del sistema educativo. Los resultados de toda esta ingente evaluación así como la experiencia de diversos organismos especializados (Beaton, 1987) constituyen una base de conocimientos sobre los que se han ido construyendo sistemas de evaluación de mayor envergadura como evaluaciones nacionales y transnacionales.

Simultáneamente con las evaluaciones institucionales en cada país, hace tiempo que existen organizaciones internacionales cuyo objetivo es promover evaluaciones sectoriales que se aplican en los sistemas educativos de diversos países participantes (IEA, 1988; OCDE, 1997, 2000). La Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA) es una asociación de centros de investigación de todo el mundo cuya participación en los sucesivos estudios es decidida libremente por cada uno de ellos. La IEA inició su actividad investigadora en 1964 con el Primer Estudio de Matemáticas, que tuvo su continuidad en el segundo (inicio de los 80s), y recientemente con el Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS); a su vez, a finales de los 70s dirigió el primer estudio de rendimiento escolar en ciencias donde participaron 19 países, continuado a mediados de los 80s con el segundo (IEA, 1988). Estos estudios tienen una fuerte perspectiva comparativista, de modo que uno de los objetivos principales consiste en facilitar las comparaciones entre los participantes en esos procesos de evaluación. Para ofrecer una aproximación a estos proyectos describiremos brevemente dos de los más recientes en los que ha participado nuestro país.

El Tercer Estudio Internacional en Ciencias y Matemáticas (TIMSS) es una evaluación comparativa transnacional de la enseñanza y el aprendizaje escolar en matemáticas y ciencias (Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.L. & Smith, T.A. 1996a, 1996b), dirigida por la Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA). La planificación del TIMSS comenzó en 1989 y abarca tres niveles del sistema educativo, alumnado de 9 años (grados 3 y 4), 13 años (grados 7 y 8) y el año final de la educación secundaria. La toma de datos tuvo lugar en 1994 y 1995, los primeros informes internacionales se entregaron en noviembre de 1996 y junio de 1997 y los informes internacionales adicionales se emitieron en 1998. Los estudiantes participantes de 45 países fueron examinados de ciencias y matemáticas y fueron encuestados, junto con los profesores y los directores de las escuelas, mediante cuestionarios que versan sobre el contexto educativo, antecedentes, actitudes, currículo, experiencia y prácticas en la enseñanza y aprendizaje de cada país. En general, los estudios de evaluación del rendimiento escolar suelen emplear casi exclusivamente preguntas de elección múltiple (de respuesta cerrada); una de las novedades relativas del TIMSS es la incorporación de preguntas de rendimiento abiertas de las que se ofrecen sendas muestras en la tabla 6.

España participó en el nivel 2 (alumnado de 13 años, 7º y 8º grados) con una muestra representativa de 7,596 estudiantes distribuidos en dos mitades aproximadamente iguales entre el grado inferior (7º de EGB, 49.2%) y el grado superior (8º de EGB, 50.8%) y entre chicos (48.7%) y chicas (51.3%). Los estudiantes pertenecen a 150 escuelas públicas y privadas, donde mediante procedimientos aleatorios, se seleccionó una clase en cada curso (séptimo y octavo de EGB). Los aprendizajes se evalúan mediante cuestiones cerradas de elección múltiple y cuestiones de respuesta abierta (corta y extensa) cuyos resultados se operacionalizan en una puntuación estandarizada de rendimiento (media 500

puntos, d.t. 100) aplicando la teoría de respuesta al ítem con el modelo de Rasch generalizado. Los cuestionarios referidos a la escuela, el profesorado y el alumnado se elaboraron mediante un desarrollo de estudios pilotos, enmienda, corrección y traducción por los países participantes y proporcionan los datos correspondientes a los factores contextuales del rendimiento.

La media internacional en el rendimiento de matemáticas es 513 puntos (8º) y 484 puntos (7º); los estudiantes españoles han obtenido una media de 487 (error estándar, 2.0) y 448 (error estándar, 2.2). Por orden de puntuaciones medias de matemáticas, España ocupa el lugar 31 en 8º (sobre 41 países) y el lugar 32 en 7º (sobre 39 países), indicadores claramente por debajo de la media, de modo que la situación de rendimiento comparativo es baja. En ciencias, la media internacional del rendimiento es 516 puntos (8º) y 479 puntos (7º); nuestro país tiene una media de 517 (error estándar, 1.7) y 477 (error estándar, 2.1), siendo el país más próximo a la media global, aunque ligeramente por debajo de ella, de modo que, en relación con todo el conjunto de países participantes, el rendimiento en ciencias se puede considerar intermedio (Vázquez, 2000).

Las puntuaciones porcentuales en las distintas áreas de matemáticas son más altas en el área Representación de datos y probabilidad, mientras el área de Álgebra se encuentra en torno a la media, siendo las áreas más bajas Fracciones y Medida. Las puntuaciones de los chicos son mayores que las puntuaciones de las chicas en todas las áreas (con excepción de Álgebra), aunque las diferencias solo son estadísticamente significativas en Medida para el curso 8º.

Las diferencias de rendimiento porcentual entre las distintas áreas de contenidos de ciencias, no son importantes; Geología está situada por encima de la media, mientras Física (7º) y Biología (8º) se encuentran por debajo de la media. Los chicos tienen puntuaciones significativamente superiores a las chicas en Física, Geología y Química, y las diferencias no son significativas en Medio ambiente naturaleza de la ciencia; en Biología, la mayoría de los países no tienen diferencias de género, y sin embargo, España es una excepción pues, en ambos cursos, los chicos tienen puntuaciones significativamente mayores que las chicas.

Los cuestionarios de contexto permiten relacionar el rendimiento con las variables del ambiente escolar, social y familiar, mediante variables categóricas, a las que se ha aplicado un análisis de relaciones monótonas entre el rendimiento y las categorías determinadas por las variables, y variables continuas, a las que se han aplicado análisis de correlaciones.

Así, se han encontrado relaciones positivas entre el rendimiento y el número de libros en casa, el nivel de estudios completados por el padre, la posesión de calculadora, ordenador, pupitre, diccionario, enciclopedias, instrumentos de laboratorio, programas educativos y sitio para estudiar, las expectativas de la madre respecto a la importancia de ir bien en la escuela en ciencias, matemáticas y lengua, de estar entre los primeros de la clase y de tener tiempo para divertirse; la lectura de libros o revistas, la imagen positiva de matemáticas y ciencias, las atribuciones a la suerte, la necesidad personal de ir bien en ciencias y matemáticas motivada por acceder a los estudios deseados (extrínseca) o estar contento consigo mismo (intrínseca); variables del profesorado como la mayor edad, el género femenino, el aprecio y comprensión de los alumnos como esencial para enseñar, la frecuencia de uso de utilización de guías o libros del profesor; variables del centro como la zona geográfica donde se ubica es el centro, tener establecido por escrito su propio contenido curricular, ofertar clases de recuperación de ciencias y matemáticas, y actividades de profundización en matemáticas para estudiantes avanzados (minoritarias), el número de

ordenadores disponibles en un centro, el número de chicos matriculados (tamaño del centro) y la proporción de estudiantes con escolarización en preescolar.

También cabe destacar algunas variables relacionadas negativamente con el rendimiento como las metodologías didácticas de trabajar problemas o proyectos en grupos pequeños y las preguntas del profesor sobre lo que los alumnos saben del tema nuevo, la proporción de estudiantes de un centro que se encuentran en situaciones desfavorecidas (bajo estatus económico, padres sin estudios, problemas de aprendizaje, de salud o de alimentación). Otras variables cuya relación con el rendimiento es negativa son la capacidad de pensar creativamente y la mayor familiaridad del profesorado con las guías de orientaciones pedagógicas (en el caso de matemáticas) y para ciencias, entender conceptos, principios y estrategias o la dedicación del director del centro a las tareas de administración interna.

Un resultado notable encontrado es la relación positiva y significativa entre el rendimiento y el número de alumnos por clase (ratio), que contradice una creencia muy extendida en la enseñanza (mejor rendimiento en las clases más pequeñas). En esta misma línea, también cabe notar la ausencia de relación empírica con el rendimiento de muchas variables, tradicionalmente consideradas factores importantes del aprendizaje, como por ejemplo, entre otras, la frecuencia de reuniones del profesorado, las metodologías de agrupamientos de alumnos, las actividades de aprendizaje o las técnicas de evaluación (Vázquez, 2000).

Los proyectos de la Organización para la Cooperación y el Desarrollo Económicos (OCDE)

El Proyecto Internacional para la producción de Indicadores de Resultados Educativos de los Alumnos (PISA) de la OCDE supone una expresión de voluntad y un compromiso de los países miembros de la OCDE y de los países participantes para radiografiar en un cuadro internacional común los resultados y niveles alcanzados por los estudiantes de los diferentes sistemas educativos. El fundamento conceptual del desarrollo de las pruebas y cuestionarios administrados a los estudiantes que incluye los dominios de evaluación, los objetivos de la evaluación y las condiciones de los contextos donde se realizan, ha sido sistematizado en una publicación previa (OCDE, 2000).

El PISA es un esfuerzo conjunto de colaboración de los países participantes y dirigido por sus gobiernos, partiendo de sus intereses comunes en el área de la política educativa. El programa investiga las competencias adquiridas por los estudiantes de 15 años en lectura, matemáticas y ciencias que pretende ser duradero en años sucesivos (2000, 2003, 2006). No obstante, la materia dominante en el 2000 fue lectura, en 2003 serán las matemáticas y el 2006 la cultura científica. La evaluación esta planificada para suministrar diversos tipos de indicadores, tales como perfil general y competencias de los estudiantes, indicadores contextuales sobre las relaciones entre competencias y diversas variables sociales, demográficas, económicas y educativas, e indicadores de tendencias a partir de la comparación de los sucesivos estudios en el tiempo. A diferencia del TIMSS, que es un estudio basado en el currículo común consensuado por los participantes, el PISA enfoca la evaluación sobre los conocimientos, destrezas y competencias que se consideran más relevantes para preparar a los jóvenes para la vida.

El proyecto PISA se ha aplicado durante el año 2000 en 32 países y sus resultados están a punto de aparecer publicados en el momento de escribir estas líneas. El PISA aplicó

en nuestro país sus pruebas de rendimiento escolar a 6,840 alumnos de final de secundaria en 171 centros escolares diferentes y cuestionarios de contexto a los estudiantes y a los directores de los centros escolares participantes.

Conclusión a modo de prospectiva

Al margen de los estudios de investigación en evaluación, mucho más autónomos, concretos y limitados en amplitud, el panorama de la evaluación educativa en los próximos años aparece lleno de alicientes y retos. Por un lado, la línea de los macroestudios institucionales, tanto nacionales como internacionales, parece el instrumento más poderoso de los gobiernos como suministrador de los datos de base para tomar las decisiones educativas y, además, tiene garantizada su pervivencia, por la seguridad del aporte de los recursos necesarios para su realización. En este tema, sin duda, asistiremos en los próximos años a verdaderos aluviones de datos cuya valoración se tomará como justificación para los cambios y las decisiones institucionales, legales y reales, sobre el sistema educativo. Todo ello tendrá un impacto social intenso, contribuirán a producir debates públicos sobre la educación, cuyo alcance, en estos momentos, aunque podemos barruntar, tal vez, no seremos capaces de imaginar por la intervención de los medios de comunicación. En el anverso, el exceso de datos puede plantear también el problema de su digestión; que esta no se vuelva pesada o imposible, especialmente cuando estos datos sean abiertamente contradictorios o conduzcan a resultados altamente polémicos, será un objetivo importante. Por otro lado, las microevaluaciones desarrolladas al nivel de los centros escolares, tales como la evaluación de los proyectos educativos, la evaluación de la práctica docente, la evaluación del centro, la evaluación para la calidad o la mejora, etc. bien en la forma de evaluación externa o de autoevaluación constituyen los desafíos más cercanos, a pie de obra de la escuela. Aunque cabe la posibilidad que este nivel se pueda desarrollar externamente, con poca o nula participación del profesorado, no parece que sea esta la posibilidad más verosímil, por la imagen negativa que tendría, y cabe pensar que los centros escolares tienen aquí un reto nuevo, hasta ahora rechazado. Las razones para este rechazo son variadas, pero el aspecto de control que subyace siempre en toda evaluación, en este caso un control directo del trabajo docente, junto con la falta de preparación del profesorado para realizar esta evaluación, y en general, la falta de una cultura de evaluación en el pensamiento profesional del profesorado, estarían entre los factores más importantes del rechazo. Por ello, salvo porque nuevos factores en el escenario educativo incentiven de forma poderosa esta actividad, cabe esperar la misma reluctancia del profesorado en la gestión de la autoevaluación, como ya ha ocurrido en años pasados. Si esto es así, la consecuencia más importante sería que la innovación educativa de abajo arriba, las propuestas desde el profesorado y las aulas hacia el sistema global, estarían prácticamente yuguladas, e incluso aplastadas, si se tiene en cuenta el previsible alud de continuas reformas de arriba hacia abajo, nacidas de la toma de decisiones producidas por las macroevaluaciones.

Como reflexión prospectiva final, cabe augurar la potencial importancia práctica que pueda cobrar un aspecto de la evaluación hasta ahora sólo presente en la investigación: la metaevaluación (de la Orden, Martínez, 1992). La metaevaluación, la evaluación de la evaluación, la evaluación de los programas de evaluación mismos, que incluye también la evaluación de los evaluadores, es decir, la evaluación de la administración educativa, es el último reto en evaluación que quizá se desarrolle en el futuro próximo. La metaevaluación

supone un sistema de garantía y responsabilidad ante la sociedad por la gestión de los sistemas de evaluación aplicados y su desarrollo necesita las normas generales de evaluación ya citadas (Joint Committee, 1988) que desempeñarán, sin duda, un papel crucial en el futuro.

INDICADORES DE ESCOLARIZACIÓN

- E1. Escolarización en cada etapa educativa
- E2. Escolarización y población
 - E2.1. Escolarización en las edades de 0 a 29 años
 - E2.2. Esperanza de vida escolar a los 6 años
- E3. Evolución de las tasas de escolarización en las edades de los niveles no obligatorios
 - E3.1. Educación infantil
 - E3.2. Educación secundaria postobligatoria
 - E3.3. Educación superior universitaria
- E4. Acceso a la educación superior universitaria
- E5. Atención al alumnado con necesidades educativas especiales

INDICADORES DE PROCESOS EDUCATIVOS

- P1. Tareas directivas
 - P1.1. Tareas directivas en centros de EP
 - P1.2. Tareas directivas en centros con ESO
- P2. Participación de los padres en el centro
 - P2.1. Participación de los padres en centros de EP
 - P2.2. Pertenencia a asociaciones de padres de alumnos
- P3. Trabajo en equipo de los profesores
 - P3.1. Trabajo en equipo de los profesores en centros de EP
 - P3.2. Trabajo en equipo de los profesores en ESO
- P4. Estilo docente del profesor
 - P4.1. Estilo docente del profesor de EP
 - P4.2. Estilo docente del profesor de ESO
- P5. Actividades del alumno fuera del horario escolar
 - P5.1. Trabajos escolares en casa
 - P5.2. Actividades extraescolares
- P6. Formación permanente del profesorado
 - P6.1. Formación permanente del profesorado de EP
 - P6.2. Formación permanente del profesorado de ESO
- P7. Relaciones en el aula y en el centro
 - P7.1. Relaciones en los centros de EP
 - P7.2. Relaciones en los centros con alumnos en edades de ESO

INDICADORES DE CONTEXTO

- C1. Proporción de población en edad escolarizable
- C2. PIB por habitante
- C3. Relación de la población con la actividad económica
- C4. Nivel de estudios de la población adulta
- C5. Expectativas de nivel máximo de estudios

ESO: Educación Secundaria Obligatoria. EP: Educación Primaria. PIB: Producto Interior Bruto

Tabla 1. Sistema estatal de indicadores de la educación 2000 del Instituto Nacional de Calidad y Evaluación (INCE, 2000).

INDICADORES DE RECURSOS

- Rc1. Gasto total en educación con relación al PIB
- Rc2. Gasto público en educación
- Rc3. Gasto en educación por alumno
- Rc4. Proporción de población activa empleada como profesorado
- Rc5. Alumnos por grupo y por profesor
- Rc5.1. Alumnos por grupo educativo
- Rc5.2. Alumnos por profesor

INDICADORES DE RESULTADOS EDUCATIVOS

- Rs1. Resultados a los 12 años
- Rs1.1. Resultados en Lengua castellana
- Rs1.2. Resultados en Matemáticas
- Rs2. Resultados a los 16 años
- Rs2.1. Resultados en comprensión lectora
- Rs2.2. Resultados en reglas lingüísticas y Literatura
- Rs2.3. Resultados en Matemáticas
- Rs3. Adquisición de actitudes y valores
- Rs3.1. Manifestación de conductas en los alumnos
- Rs3.2. Importancia de la educación de actitudes y valores en las edades de ESO
- Rs4. Idoneidad en la edad del alumnado de educación obligatoria
- Rs5. Tasas de graduación
- Rs6. Educación y situación laboral de los padres y educación de los hijos
- Rs6.1. Educación de los padres y educación de los hijos
- Rs6.2. Situación laboral de los padres y educación de los hijos
- Rs7. Tasa de actividad y de desempleo según nivel educativo
- Rs7.1. Tasa de actividad según nivel educativo
- Rs7.2. Tasa de desempleo según nivel educativo

ESO: Educación Secundaria Obligatoria. EP: Educación Primaria. PIB: Producto Interior Bruto

Tabla 1. Sistema estatal de indicadores de la educación 2000 del Instituto Nacional de Calidad y Evaluación (INCE, 2000).

Indicadores de logro

- 1. Matemáticas
- 2. Lectura
- 3. Ciencias
- 4. Tecnologías de la comunicación y de la información (ICT)
- 5. Lenguas extranjeras
- 6. Aprender a aprender
- 7. Civismo

Indicadores de éxito y transición

- 1. Tasa de abandonos
- 2. Finalización de educación secundaria superior
- 3. Participación en la educación terciaria

Indicadores de control educativo

- 1. Evaluación y conducción de la educación escolar
- 2. Participación de los padres

Indicadores de recursos y estructuras

- 1. Formación del profesorado
- 2. Participación en educación preescolar
- 3. Número de estudiantes por ordenador
- 4. Gasto educativo por alumno

Tabla 2. Grupos de indicadores de calidad de la educación escolar propuestos por la Comisión Europea.

QUÉ ES UN PLAN ANUAL DE MEJORA

- Un instrumento para aprender como organización.
- Una herramienta para mejorar la gestión educativa.
- Un medio para elevar la calidad de los centros docentes.
- Un compromiso entre el centro y la administración.

CUÁLES SON SUS PRINCIPALES CARACTERÍSTICAS

Debe estar precedido de un diagnóstico explícito de la situación de partida con relación al área o áreas prioritarias sobre las que se centrará el Plan.

La identificación de las áreas de mejora ha de ser objetiva y debe apoyarse en hechos o en resultados antes que en juicios subjetivos o meras apariencias.

Los objetivos de mejora han de ser realistas, concretos, evaluables y alcanzables en un curso escolar, sin perjuicio de que puedan integrarse como parte de un plan de mejora de carácter plurianual.

Debe explicitar los objetivos, los procedimientos y actuaciones previstos, las personas responsables de su ejecución, los recursos y apoyos necesarios, un calendario para su cumplimiento y un plan para su evaluación.

Debe implicar a las personas, desde una orientación participativa y bajo el impulso asociado a un liderazgo efectivo por parte de la Dirección.

QUÉ SE PUEDE CONSEGUIR CON EL PLAN ANUAL DE MEJORA

- El aprendizaje de procedimientos para la mejora continua.
- El apoyo y asesoramiento externo de la administración.
- Recursos adicionales para financiar nuevos planes de mejora.
- El reconocimiento interno, externo y de la administración educativa por la voluntad de mejorar y por los resultados obtenidos.
- El impulso de los procesos y de los resultados que son importantes para el centro docente.
- La satisfacción compartida de formar parte de un equipo humano comprometido que convierte los desafíos en oportunidades.

QUÉ PERSONAS O INSTITUCIONES INTERVIENEN EN EL PLAN ANUAL DE MEJORA

- El equipo directivo del centro que asume su liderazgo y se implica en la puesta en marcha del Plan.
- El equipo de mejora del centro, formado al efecto, que colabora muy directamente en la concepción, el desarrollo y la aplicación del Plan.
- El Claustro de profesores y el Consejo Escolar que otorgan su criterio, su apoyo y su aprobación.
- La Inspección Educativa que media entre el centro y la administración, colabora con el centro en el seguimiento del Plan y efectúa su evaluación externa.
- La Unidad de Programas Educativos que proporciona el apoyo y el asesora
- El Director Provincial que se implica personalmente en los procesos de mejora, cumple con los centros los compromisos adquiridos y garantiza la correcta difusión y aplicación de las instrucciones recibidas.
- La Dirección General de Centros Educativos, que impulsa globalmente el proceso desde el convencimiento de su incidencia, a medio plazo, en la mejora de nuestro sistema educativo.

QUÉ CLASES DE AYUDA SE PUEDEN OBTENER DE LA ADMINISTRACIÓN

- De tipo técnico, mediante el asesoramiento en cualquiera de las fases de desarrollo del Plan Anual de Mejora.
- De tipo económico, en función de las características del Plan y de la calidad de su aplicación.
- De tipo humano, por efecto de un mayor compromiso de los diferentes niveles de la administración con los centros dispuestos a mejorar de un modo ordenado y sistemático.

QUÉ FORMAS DE RECONOCIMIENTO ESTÁN PREVISTAS

- Reconocimiento personal, por parte del Director Provincial, en representación de la administración, a los directivos de todos los centros que desarrollaron satisfactoriamente el Plan.
- Publicación en el BOMECE de la relación de centros cuyos Planes Anuales de Mejora hayan sido evaluados positivamente.
- Aportación de una asignación suplementaria de un millón de pesetas para los cincuenta mejores Planes, en un acto de entrega formal organizado por la Dirección General de Centros Educativos.
- Difusión de los mejores Planes y de sus prácticas de buena gestión.

Tabla 3. Características de los Planes Anuales de Mejora.

Ficha de evaluación nº 12		RESULTADOS	
1	El profesorado no está satisfecho con el funcionamiento del Centro	0 1 2 3 4 5	El profesorado está muy satisfecho con el funcionamiento del Centro.
2	El alumnado no está satisfecho con el funcionamiento del Centro	0 1 2 3 4 5	El alumnado está muy satisfecho con el funcionamiento del Centro.
3	Los padres de los alumnos no están satisfechos con el funcionamiento del Centro	0 1 2 3 4 5	Los padres de los alumnos están muy satisfechos con el funcionamiento del Centro.
4	El nivel de desarrollo de actitudes cívicas por parte del alumnado es insatisfactorio	0 1 2 3 4 5	El nivel de desarrollo de actitudes cívicas por parte del alumnado es muy satisfactorio.
5	El porcentaje de alumnos con evaluación positiva al finalizar el pasado curso es muy poco satisfactorio	0 1 2 3 4 5	El porcentaje de alumnos con evaluación positiva al finalizar el pasado curso es muy satisfactorio.
6	El porcentaje de alumnos que promocionaron al finalizar el pasado curso es muy poco satisfactorio	0 1 2 3 4 5	El porcentaje de alumnos que promocionaron al finalizar el pasado curso es muy satisfactorio.
7	Las tasas de alumnos que abandonaron los estudios durante el pasado curso es muy elevada	0 1 2 3 4 5	La tasa de alumnos que abandonaron los estudios durante el pasado curso es muy reducida.
8	El índice de alumnos repetidores es muy elevado	0 1 2 3 4 5	El índice de alumnos repetidores es muy bajo.
9	Las tasas de absentismo del alumnado durante el presente curso son muy altas	0 1 2 3 4 5	Las tasas de absentismo del alumnado durante el presente curso son muy bajas.
10	Los resultados de los alumnos del Centro en las PAAU* en el año anterior fueron insatisfactorios	0 1 2 3 4 5	Los resultados de los alumnos del Centro en las PAAU* en el año anterior fueron insatisfactorios.
11	Los índices de progresión absoluta del alumnado del Centro son muy bajos	0 1 2 3 4 5	Los índices de progresión absoluta del alumnado del Centro son muy altos.

Tabla 4. Ficha de evaluación estandarizada del plan EVA referida a los resultados.

Análisis del Proyecto Curricular de Etapa	
PROYECTOS	
03.02.	Adecuación entre los objetivos del Proyecto curricular de etapa y las necesidades y posibilidades de los alumnos y alumnas del Centro.
03.03.	Adecuación de la distribución de los objetivos, contenidos y criterios de evaluación de las distintas áreas en el Proyecto curricular de etapa.
03.04.	Adecuación de las orientaciones para incorporar los temas transversales.
03.05.	Explicitación y adecuación de la programación de las actividades complementarias y extraescolares.
03.06.	Adecuación de las decisiones de carácter general sobre metodología didáctica recogidas en el Proyecto curricular de etapa.
03.07.	Adecuación de los criterios y procedimientos fijados para realizar las adaptaciones curriculares para los alumnos con necesidades educativas especiales, si los hay.
03.08.	Idoneidad de los materiales y recursos didácticos que se van a utilizar, incluidos los libros para uso de los alumnos y alumnas.
03.09.	Adecuación del plan de acción tutorial y de la organización de la orientación educativa que figuran en el Proyecto curricular de etapa.
03.10.	Adecuación de los criterios generales fijados para la evaluación de los aprendizajes y promoción de los alumnos.
03.11.	Adecuación de los criterios fijados para evaluar los procesos de enseñanza y la práctica docente de los maestros.
03.12.	Utilidad del Proyecto curricular de etapa como marco de referencia para decidir sobre objetivos y contenidos, metodología y criterios de evaluación.

Tabla 5. Cuestionario con los indicadores para el análisis del Proyecto Curricular de Etapa.

Matemáticas. Tema: Representación de Datos, Análisis y Probabilidad. Destreza: Resolver Problemas.

V2. Los anuncios siguientes aparecieron en un periódico de un país cuya moneda es el zed.

EDIFICIO A

Se alquilan oficinas
8595 metros cuadrados
475 zeds al mes
100120 metros cuadrados
800 zeds al mes

EDIFICIO B

Se alquilan oficinas
35260 metros cuadrados
90 zeds por metro cuadrado al año

Si una empresa está interesada en alquilar durante un año una oficina de 110 metros cuadrados ¿en qué edificio de oficinas, A o B, debe alquilar la oficina para conseguir el precio más bajo? Razona la respuesta.

Ciencias. Tema: Geología. Destreza: Teorizar, Analizar y Resolver Problemas.

El Sol es mayor que la Luna, pero parece que es del mismo tamaño si lo miramos desde la Tierra.
¿ A qué se debe esto ?

Tabla 6. Ejemplo de las cuestiones abiertas propuestas a los estudiantes de la población 2 en el TIMSS.

Bibliografía

- ÁLVAREZ, R. (2000). La evaluación del sistema educativo cubano. Una experiencia. *Revista de Educación*, 321, 4246.
- ALVARO PAGE, M. Y OTROS (1990). *Hacia un modelo causal del rendimiento académico*. Madrid: Centro de Investigación, Documentación y Evaluación.
- BARRERA, S. (2000). Estrategia para la viabilidad de un sistema de evaluación de la calidad de la educación: caso Bolivia. *Revista de Educación*, 321, 2334.
- BEATON, A. E. (1987). *The NAEP Technical Report*. Princeton, NJ: Educational Testing Service University of Princeton.
- BEATON, A. E., MULLIS, I. V. S., MARTIN, M. O., GONZALEZ, E. J., KELLY, D. L. & SMITH, T. A. (1996a). *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science study (TIMSS)*. Chestnut Hill, MA: Boston College.
- BEATON, A. E., MULLIS, I. V. S., MARTIN, M. O., GONZALEZ, E. J., KELLY, D. L. & SMITH, T. A. (1996b). *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science study (TIMSS)*. Chestnut Hill, MA: Boston College.
- BRACKENBURY, J. (2000). Canadá: informe de evaluación nacional. *Revista de Educación*, 321, 4760.
- CASTILLO, S. Y GENTO, S. (1995). Modelos de evaluación de programas educativos. En A. Medina y L.M. Villar (1995), *Evaluación de programas educativos, centros y profesores*, (pp. 2569), Madrid, Universitas.
- CARBALLO, R. (1990). Evolución del concepto de evaluación; desarrollo de los modelos de evaluación de programas. *Bordón*, 42, 423431.
- CHAMBERS, F. (1994). Removing confusion about formative and sumative evaluation: purpose versus time. *Evaluation and Program Planning*, 17, 912.
- CIDE, Centro de Investigación, Documentación y Evaluación (1988). *Evaluación externa de la reforma experimental de las enseñanzas medias (I)*. Madrid: MEC-CIDE.
- CIDE, Centro de Investigación, Documentación y Evaluación (1990). *Evaluación externa de la reforma experimental de las enseñanzas medias (II)*. Madrid: MEC-CIDE.
- CIDE, Centro de Investigación, Documentación y Evaluación (1992). *Evaluación externa de la reforma experimental de las enseñanzas medias (III)*. Madrid: MEC-CIDE.
- COSGROVE, J. (2000). El sistema de evaluación de resultados educativos: Irlanda. *Revista de Educación*, 321, 2334.
- DE LA ORDEN, A. Y MARTINEZ, M. J. (1991). Metaevaluación educativa. *Bordón*, 43, 517527.
- EC, European Commission (2000). European Report on quality of school education. Sixteen quality indicators. Directorate General for Education and Culture.
- ELLIOT, J.(1992). ¿Son los indicadores de rendimiento indicadores de la calidad educativa? *Cuadernos de Pedagogía*, 206, 5660.
- GUIMARAES, M. H., GOMES, M. I. Y SCHULMEYER, M. A. (2000). El sistema de evaluación de la educación básica de Brasil. *Revista de Educación*, 321, 6180.
- HOUSE, E. R. (1992). Tendencias en evaluación. *Revista de Educación*, 299, 4355.
- IEA, International Association for the Evaluation of Educational Achievement (1988). *Science Achievement in Seventeen Countries. A Preliminary Report*. Oxford: Pergamon Press.
- INCE, Instituto Nacional de Calidad y Evaluación (1998). *Diagnóstico general del sistema educativo. Avance de resultados*. Madrid: Ministerio de Educación y Cultura.

- INCE, Instituto Nacional de Calidad y Evaluación (2000). *Sistema estatal de indicadores de la educación 2000: síntesis*. Madrid: Centro de Publicaciones del Ministerio de Educación y Deporte.
- Joint Committee on Standards for Educational Evaluation (1988). *Normas de evaluación para proyectos y material educativo*. México: Trillas.
- KALLEN, D. (1990). *Nouveaux paradigmes éducatifs et nouvelles politiques d'évaluation*. París: OECDE.
- LAZARO, A. (1994). Indicadores para evaluar centros educativos. *Organización y Gestión Educativa*, 3, 2128.
- LUJAN, J. Y PUENTE, J. (1996). *Evaluación de centros docentes. El Plan Eva*. Madrid: Ministerio de Educación y Ciencia.
- MARTINEZ, F. F. (2000). El sistema de evaluación educativa de México. *Revista de Educación*, 321, 3540.
- MEC, Ministerio de Educación y Ciencia. (1997). *Modelo Europeo de Gestión de Calidad*. Madrid: MEC-Argentaria.
- OCDE (1992). *Education at a Glance. OECD Indicators*. París: OCDE.
- OCDE (2000). *PISA. La medida de los conocimientos y destrezas de los alumnos. Un nuevo marco para la evaluación*. Madrid: MEC-DINCE.
- OFS (1993). *Les indicateurs de l'enseignement en Suisse*. Berna: Office Fédéral de la Statistique.
- PÉREZ, R., LOPEZ, F., PERALTA, M. D., MUNICIO, P. (2000). *Hacia una educación de calidad. Gestión, instrumentos y evaluación*. Madrid. Narcea.
- PUENTE, J. (2000). La evaluación del sistema educativo español. *Revista de Educación*, 321, 8196.
- STUFFLEBEAM, D. L. Y SHINKFIELD, A. J. (1987). *Evaluación sistemática. Guía teórica y práctica*. Madrid: Paidós, MEC.
- TENBRINK, T. D. (1983). *Evaluación. Guía práctica para profesores*. Madrid: Narcea.
- UVM, UNDERVISNINGSMINISTERIET (1993). *Facts and figures. Education Indicators. Denmark*. Copenhagen: Danish Ministry of Education and Research.
- VARI, P. (2000). La evaluación educativa Monitor y el sistema de evaluación de Hía. *Revista de Educación*, 321, 922.
- VAZQUEZ, A. (2000). *Análisis de los datos del tercer estudio internacional de matemáticas y ciencias (TIMSS) desde la perspectiva del sistema educativo español*. Memoria final de investigación. Palma de Mallorca: Autor.
- WALBERG, H. J. Y HAERTEL, G. D. (1990). *The international encyclopedia of educational evaluation*. Oxford. Pergamon Press.