



VARIABLES PREDICTORAS DE LA CALIDAD DE LA EDUCACIÓN SECUNDARIA

EQUIPO REDES¹

RESUMEN. En las dos últimas décadas se han desarrollado dos tradiciones diferentes para conocer el funcionamiento de los centros e impulsar su cambio: las investigaciones sobre la eficacia de las escuelas y el movimiento para su mejora. Sin embargo, recientemente se han desarrollado nuevos proyectos que consideran necesario conectar ambos paradigmas y generar una dinámica de apoyo mutuo. Una de estas iniciativas es la Red de Evaluación de Centros que tiene dos objetivos principales: la cooperación con los centros para conocer mejor su funcionamiento, lo que está más relacionado con el movimiento de mejora, y la investigación sobre los factores responsables de la calidad de la enseñanza, lo que se incluye dentro de los estudios sobre la eficacia de las escuelas. El artículo desarrolla cada una de estas finalidades y aporta los resultados iniciales obtenidos en la investigación. En la última parte se detallan las aportaciones de la red de evaluación para impulsar la confluencia entre el estudio sobre la calidad de los centros y las iniciativas para su mejora.

LOS ESTUDIOS SOBRE LA EFICACIA Y LA MEJORA DE LAS ESCUELAS

En las dos últimas décadas se han desarrollado con fuerza dos potentes movimientos relacionados con el conocimiento y el cambio de las escuelas: las investigaciones sobre su eficacia y los proyectos para su desarrollo y mejora. Ambos han tenido una indudable influencia en la conceptualización del funcionamiento de las escuelas y en el análisis de los factores más responsables de su calidad o mejora. Sin embargo, los dos enfo-

ques han tenido un desarrollo paralelo durante mucho tiempo y ha sido sólo en los últimos años cuando se ha producido un intento claro de aproximación y confluencia (Reynolds et al., 1996; Stoll y Fink, 1996). De forma intuitiva podría pensarse que los estudios sobre la eficacia de las escuelas conducirían a impulsar programas de mejora basados en los resultados obtenidos. No ha sido así durante muchos años, debido a sus diferencias en los objetivos, en los criterios que utilizan para evaluar la calidad de los centros y en la metodología empleada.

(1) El equipo Redes está dirigido por Álvaro Marchesi y está constituido por Leoncio Fernández, Elena Martín, Rosario Martínez Arias, Isabel Monguilot, Amparo Moreno, Enrique Roca, Vicente Rivière y Alejandro Tiana. Las tareas de coordinación son realizadas por Eva M.ª Pérez y Ricardo Lucena.

Por lo que se refiere a los objetivos que pretenden, los estudios sobre la eficacia de las escuelas entran claramente dentro del campo de la investigación científica por lo que su finalidad principal es ampliar el conocimiento disponible sobre el funcionamiento de las escuelas y los factores que en él influyen. Además, los resultados que se obtienen en estas investigaciones no se pueden trasladar directamente a la práctica. Es preciso también disponer de modelos de cambio educativo que faciliten la transición de la teoría a la práctica. Por el contrario, los programas para mejorar la escuela tienen un objetivo más aplicado y pretenden principalmente comprobar la utilidad de determinadas estrategias de cambio. El conocimiento que se genera es menos riguroso por lo que sus conclusiones son más limitadas y con menor poder de generalización.

La segunda diferencia se sitúa en la selección de los indicadores de calidad. Los estudios sobre la eficacia de las escuelas se concentran prioritariamente en los resultados que obtienen los alumnos e incluso seleccionan dentro de ellos los más directamente vinculados a la dimensión cognitiva en detrimento de los objetivos sociales, afectivos o morales. El foco de su investigación es determinar qué dimensiones de los procesos de centro y de aula están más directamente relacionadas con los resultados de los alumnos. Por el contrario, los proyectos de mejora de las escuelas abordan las condiciones y los procesos del cambio y tienen una visión más amplia y sistémica de lo que es la realidad educativa.

Finalmente, existen diferencias importantes en la metodología utilizada. Los investigadores en el campo de las escuelas eficaces tienen una orientación principalmente cuantitativa y correlacional mientras que los estudiosos en el ámbito de la mejora de las escuelas prefieren una orientación cualitativa para investigar de forma

más completa y profunda algunas escuelas individuales.

Como ya se ha apuntado, existen intentos recientes para aproximar ambos paradigmas y generar una dinámica de apoyo mutuo. Los estudios sobre la eficacia de la escuela manifiestan un mayor interés por conocer los procesos de cambio, ampliar los métodos de estudio y utilizar todo tipo de conocimientos disponibles. A su vez, la tradición de la mejora de la escuela está incorporando el conocimiento base generado en la investigación sobre las escuelas eficaces y está intentando comprobar de forma rigurosa la eficacia de las estrategias de cambio utilizadas y los resultados obtenidos. Este proceso de convergencia entre ambos paradigmas puede recibir un fuerte impulso a través de programas de evaluación que mantengan dos objetivos prioritarios: investigar los factores responsables de la calidad de la enseñanza y utilizar el conocimiento que se obtenga para que las escuelas que están participando en la investigación puedan mejorar su funcionamiento. Estas son las finalidades de la Red de Evaluación de Centros (REDES) que se está desarrollando en varias comunidades autónomas y en la que se basa este artículo.

LA RED DE EVALUACIÓN DE CENTROS

LAS CARACTERÍSTICAS DEL PROYECTO DE EVALUACIÓN

La red de evaluación de escuelas se constituye con la incorporación voluntaria de centros de educación secundaria y la aplicación a cada uno de ellos de las mismas pruebas y cuestionarios elaborados por un equipo externo. La duración inicial es de cuatro años con el fin de realizar un seguimiento de los mismos alumnos durante toda la etapa y constatar los

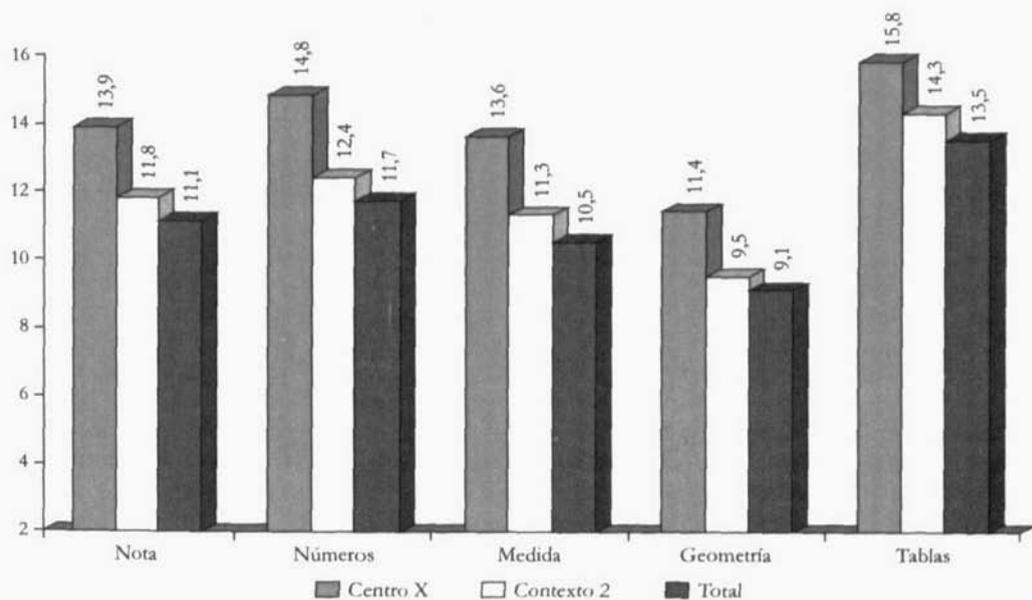
cambios que se producen a lo largo del tiempo. Se trata de crear un modelo de evaluación que tenga el atractivo suficiente para compensar el esfuerzo que se realiza, que amplíe la información que poseen las escuelas sobre los resultados que obtienen y que les ayude a impulsar programas de mejora y de cambio. Las características de la información que se obtiene son las siguientes:

- **Contextualizada.** Se controla el rendimiento inicial de los alumnos cuando empiezan la educación secundaria a los 12 años así como el nivel socioeconómico de cada uno de ellos. La información que se proporciona a los centros y que se utiliza para la investigación tiene siempre en cuenta estos datos iniciales.

- **Comparada.** Cada escuela recibe los resultados que obtiene en todas las dimensiones estudiadas, en comparación con la media de los resultados obtenidos por las escuelas de su mismo contexto social y por la totalidad de las escuelas participantes en la red. De esta forma puede interpretarlos mejor. El gráfico I es un ejemplo de cómo se proporciona la información a un centro en la prueba de Matemáticas. Cuando el número de escuelas participantes de una comunidad autónoma es superior a treinta, las comparaciones son dobles: primero en relación con los centros de la comunidad autónoma (por contexto sociocultural y con la totalidad de las escuelas) y después incluyendo a todas las escuelas que están en la red.

GRÁFICO I

Modelo de gráfico que se proporciona a un centro de los resultados que obtienen sus alumnos en la prueba de Matemáticas de primero de secundaria



- **Confidencial.** La información de cada centro sólo es conocida por el propio centro. A ninguno se le informa sobre qué otras escuelas están participando en la red y sólo conoce, además de sus propios datos, el número de centros que están en su mismo contexto sociocultural y el número total de los que participan en la red.
- **Objetiva.** La información que se proporciona procede de pruebas estandarizadas y de los cuestionarios que se han aplicado en el centro, todos los cuales han sido previamente analizados y alcanzan un alto nivel de fiabilidad y validez. Esta información se completa con las opiniones y valoraciones recogidas en el estudio cualitativo.
- **Interpretada por los propios centros y profesores.** La información que se proporciona debe ser analizada y valorada por cada centro. Son ellos los que pueden comprender mejor el alcance de los resultados obtenidos y pueden explicarlos con mayor seguridad ya que conocen la historia de la escuela, su cultura, los problemas que se han vivido y las decisiones que se han adoptado.
- **Amplia y convergente.** La información que se obtiene no se refiere solamente a los resultados académicos de los alumnos sino que recoge también sus actitudes y estrategias de aprendizaje, los procesos de centro y de aula y las valoraciones de padres, profesores y alumnos. De esta forma es posible presentar los puntos de vista de los distintos sectores de la comunidad educativa, que no siempre son coincidentes.
- **Formativa.** La función de la evaluación es colaborar con las escuelas para que se conozcan mejor y puedan diseñar estrategias de cambio. La función de control o supervisión debe ser desarrollada por otros agentes educativos.
- **Continuada.** La propuesta que se hace a las escuelas es que se mantengan en el programa durante cuatro años seguidos. De esta forma hay un seguimiento longitudinal de los mismos alumnos durante este tiempo y es posible estudiar el cambio, al tener al menos tres medidas de los mismos sujetos.
- **Capaz de comprender los factores que influyen en la calidad de la enseñanza.** Finalmente, la evaluación, en la medida en que se aplique a un número representativo de escuelas, puede proporcionar una información más científicamente fundada de la influencia relativa de los diferentes niveles y dimensiones que constituyen el modelo de evaluación.

NIVELES Y DIMENSIONES

El modelo de evaluación está organizado en cuatro niveles (véase cuadro I): entrada, contexto, procesos y resultados. En cada uno de ellos se establecen varias dimensiones, de acuerdo con lo que se recoge en el mismo cuadro. El nivel de entrada incluye los resultados iniciales de los alumnos al incorporarse a la educación secundaria, lo que se considera fundamental para analizar los cambios que se producen a lo largo del tiempo y determinar con mayor fiabilidad el valor añadido de la escuela (Gray, Jesson y Nicholas Sime, 1990; Sammons, Thomas, Mortimore, Owen y Pennell, 1994; Thomas y Mortimore, 1995). El nivel de contexto se refiere al contexto sociocultural y al tamaño del centro. Todos los estudios ponen de relieve la influencia del contexto sociocultural en los

resultados que obtienen los alumnos aunque su valor predictivo es menor si se han controlado también los resultados iniciales de los alumnos. El tamaño del centro se ha incluido ya que varios estudios (ver Raywid, 1997 para una breve revisión) sostienen que las escuelas pequeñas tienen más posibilidades de mejorar su funcionamiento y los resultados de sus alumnos.

El nivel de los procesos incluye dos ámbitos: los de organización y funcionamiento del centro y los más específicos del aula. Por lo que se refiere a los procesos institucionales (Hopkins, 1994), se incluye el conocimiento del funcionamiento del centro, la participación, el clima de trabajo, la valoración del equipo directivo y del trabajo en el seminario didáctico, las relaciones con los alumnos y las expectativas sobre el centro escolar. En relación con los procesos de aula se ha tenido en cuenta especialmente la planificación de la enseñanza, la innovación en la evaluación de los alumnos, la amplitud de los contenidos y la capacidad para mantener un ambiente de trabajo ordenado. También se ha considerado un factor de buena enseñanza la preocupación por enseñar adecuadamente a la diversidad de los alumnos (Slavin, 1996, Hopkins, West y Ainscow, 1996).

El estudio piloto, realizado para comprobar la validez y fiabilidad de los cuestionarios utilizados, puso de manifiesto la mayor debilidad del cuestionario que respondían los profesores sobre su actividad docente en el aula. Mientras que los datos obtenidos sobre los procesos de la escuela no se ven apenas afectados por la deseabilidad social, no sucede lo mismo con los referidos a los procesos en el aula. No es extraño que esto suceda así ya que los profesores no tienen problemas en ser críticos con lo que sucede en su escuela pero se muestran más condescendientes con su práctica docente y tienden a dar su opi-

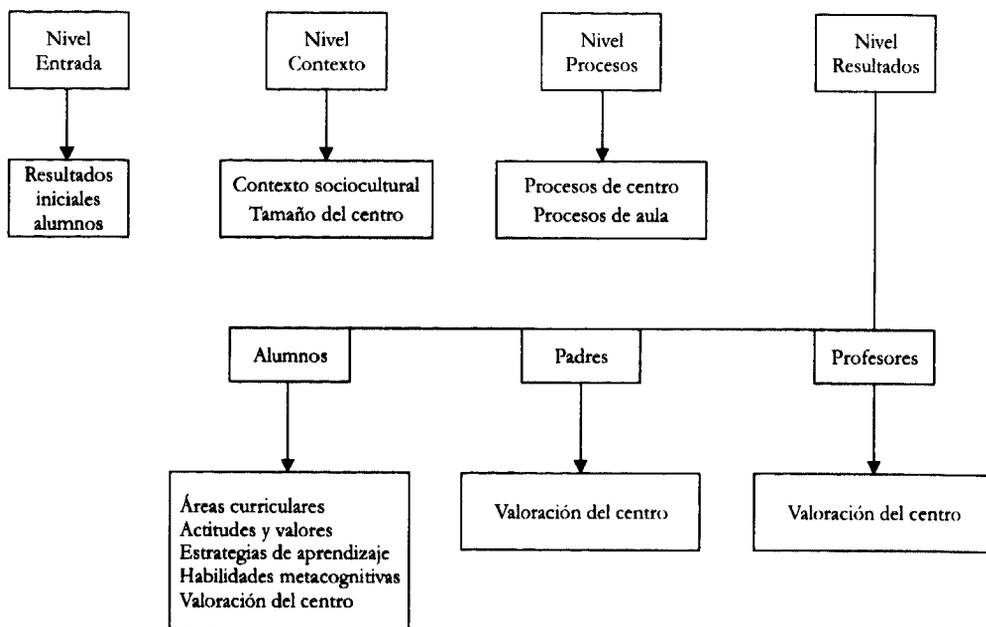
nión pensando más en el modelo ideal de enseñanza que en el real. Para corregir en lo posible esta situación, se incluyó una metodología cualitativa basada en entrevistas con el equipo directivo y con varios grupos de profesores de la escuela. Este segundo tipo de información se utiliza principalmente para contrastar y enriquecer los datos cuantitativos y para darles a las escuelas un mayor número de claves para su interpretación.

El nivel de resultados se centra principalmente en los alumnos, pero incluye también la opinión de los padres y de los profesores. En relación con los alumnos, se evalúan los aprendizajes en las principales áreas curriculares (Matemáticas, Lengua, Ciencias sociales, Ciencias experimentales y Tecnología) así como sus habilidades metacognitivas, estrategias de aprendizaje y actitudes. Los alumnos completan también un cuestionario en el que manifiestan su opinión sobre el funcionamiento del centro, sobre las calificaciones que obtienen, sobre la preparación recibida, y sobre sus profesores, tutores y compañeros. Los padres también expresan su opinión a través de un cuestionario sobre el funcionamiento del centro, la atención que reciben, la disciplina que hay en la escuela, las calificaciones de sus hijos y las actividades extraescolares. La valoración de los profesores se resume en su satisfacción por el funcionamiento del centro y por las condiciones en las que realizan su trabajo.

Las pruebas curriculares mantienen un equilibrio entre los conceptos y los procedimientos. La necesidad de una corrección rápida y fiable ha exigido que la mayoría de las preguntas de las pruebas sean de tipo cerrado con opciones múltiples. Esta opción, que es necesaria pero limitada, no debe sustituir nunca a la evaluación continua que el profesor realiza con sus alumnos.

CUADRO I

Modelo de cuatro niveles para la evaluación de las escuelas



LOS CENTROS QUE PARTICIPAN EN LA RED

La participación en la red de evaluación tiene que proceder de la decisión voluntaria de los centros. Este hecho, que no plantea ningún problema desde el punto de vista de la función de ayuda a los centros, supone una mayor dificultad cuando se pretende estudiar los factores responsables de la calidad de la enseñanza debido a la no aleatoriedad de la muestra. Sin embargo, se ha hecho un esfuerzo en la mayoría de las comunidades autónomas cuyos centros están participando en la red para que su distribución por contexto sociocultural, titularidad (pública, concertada y privada) y modelo lingüístico, en el caso de que existan varios, sea representativa de la población de referencia. Normalmente esto es posible cuando el número de centros de una comunidad autónoma es superior a

treinta. La red de evaluación comenzó en Madrid con 31 centros y después de dos años de aplicación se ha extendido a otras comunidades autónomas. En el curso 1999-00, el número de centros participantes es de 146. Hay que tener en cuenta también que la red se va ampliando cada año por lo que las posibilidades de estudio y de análisis son cada vez mayores.

LAS DOS FUNCIONES BÁSICAS DEL PROYECTO REDES

La Red de Evaluación de Centros se constituye con dos funciones principales. En primer lugar, colaborar con los centros participantes para que conozcan mejor su funcionamiento y puedan impulsar programas de cambio con mayor seguridad. En segundo lugar, investigar los factores

responsables de la calidad de la enseñanza. Ambas funciones se desarrollan de forma coordinada, lo que contribuye a su mutuo enriquecimiento. Las páginas siguientes analizan cada una de estas funciones y describen con mayor detenimiento los primeros resultados de la investigación realizada.

UNA RED PARA CONOCER Y MEJORAR EL FUNCIONAMIENTO DE CADA ESCUELA

El primer objetivo del programa de evaluación es colaborar con los centros participantes para que conozcan su funcionamiento y sus resultados y puedan elaborar estrategias de cambio. Este objetivo se lleva a la práctica a través de las cinco etapas siguientes.

PRIMERA ETAPA. COMPROMISO DEL CENTRO

El equipo directivo recibe información sobre las características del programa de evaluación y se le pide que decida sobre su incorporación a la red. A partir de este momento se producen una serie de consultas a los órganos de gobierno del centro que concluyen en una decisión positiva o negativa. El compromiso inicial supone la voluntad de mantenerse durante cuatro años en el programa aunque el centro puede desvincularse del mismo en cualquier momento.

SEGUNDA ETAPA. OBTENCIÓN DE LA INFORMACIÓN

Las pruebas y cuestionarios se aplican en tres momentos distintos en cada curso escolar: en octubre, los alumnos de 1.º de la ESO realizan las pruebas de rendimiento para conocer su nivel inicial y el cuestionario de contexto sociocultural.

Este último también se aplica a los alumnos de los demás cursos para determinar el contexto del conjunto del centro. En febrero, los profesores completan los cuestionarios de procesos de centros y de aula y los padres, el que se refiere a la valoración del centro. En abril, los alumnos de 2.º y 4.º de la ESO responden a las pruebas y cuestionarios sobre las áreas curriculares, habilidades metacognitivas, estrategias de aprendizaje, actitudes y valoración del centro.

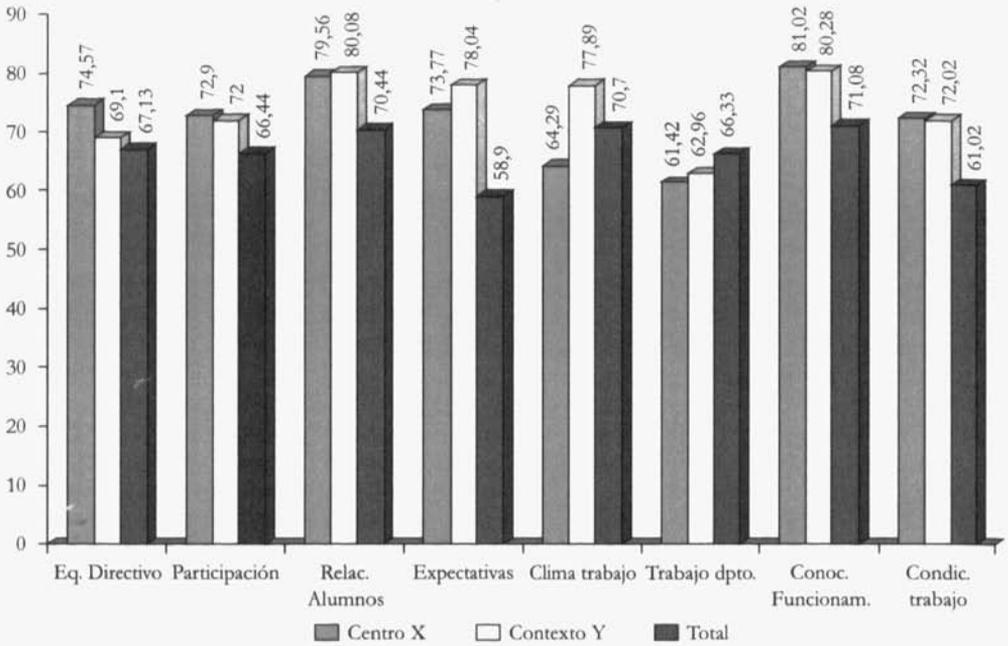
TERCERA ETAPA. DEVOLUCIÓN DE LA INFORMACIÓN A CADA CENTRO

Los centros reciben los resultados de las pruebas antes de que haya pasado un mes desde la aplicación de las mismas. Los datos de las pruebas de rendimiento incluyen las puntuaciones de cada uno de los alumnos, la media de cada clase y la media del curso. Este último dato se proporciona en comparación con la media de los centros de su mismo contexto sociocultural y en comparación con la totalidad de los centros. El gráfico I es un ejemplo. Los resultados de las valoraciones de padres, profesores y alumnos sobre el centro se presentan también de forma comparativa. El gráfico II muestra las opiniones de los profesores sobre los procesos de su centro.

A partir del segundo año, los centros también reciben información sobre las conclusiones que se obtienen del análisis del conjunto de los datos. La comunicación de estos resultados, junto con las tendencias generales que se observan, es bien aceptada por los equipos directivos de los centros, que valoran la investigación como algo propio y próximo a su realidad. Aquí se encuentra, como luego se expondrá con mayor detenimiento, uno de los puntos de encuentro entre los estudios sobre la eficacia de las escuelas y los programas orientados a su mejora.

GRÁFICO II

Resultados obtenidos por un centro en el cuestionario de Procesos de Centro en comparación con los de su mismo contexto sociocultural y con la totalidad de los centros



CUARTA ETAPA. INTERPRETACIÓN DE LA INFORMACIÓN

Una vez recibida la información, el equipo directivo de cada centro, la comisión de coordinación pedagógica y los seminarios didácticos la analizan, discuten e interpretan. En la mayoría de los casos, la información que se proporciona, tanto si es positiva como negativa, no sorprende a los equipos docentes. En algunos casos, sin embargo, las opiniones expresadas crean desconcierto: equipos directivos con una gran dedicación que no han sido valorados positivamente por los profesores; centros con buenos resultados académicos que no reciben una opinión positiva de los padres; equipos competentes de profesores que no son apreciados por sus alumnos. A pesar del malestar que esta in-

formación provoca, los centros intentan encontrar las razones de estos hechos y proponer cambios para mejorar.

También los datos generales que se obtienen del análisis estadístico de todos los datos recogidos ayuda al centro a valorar los suyos propios. El valor predictivo de los resultados iniciales, la influencia del contexto sociocultural o del tamaño del centro, junto con la influencia del resto de las variables estudiadas, ofrece un marco más amplio y contrastado para que cada centro interprete sus datos específicos.

QUINTA ETAPA. PROYECTO DE CAMBIO Y EVALUACIÓN POSTERIOR DE SUS CONSECUENCIAS

A partir del análisis e interpretación de los datos proporcionados por la evaluación,

los centros pueden, y habitualmente lo hacen, adoptar decisiones que traten de resolver algunos de los problemas detectados o mejorar el funcionamiento del centro. Las pruebas iniciales de los alumnos al comienzo de la ESO son, al mismo tiempo, una evaluación final de la educación primaria. Los centros que imparten ambas etapas, que son los concertados y privados, hacen uso de esta información para analizar el funcionamiento de la primaria y proponer los cambios que se consideren oportunos. Pero también algunos institutos de secundaria han contactado con los colegios de primaria, de donde proceden sus alumnos, y han comentado con sus equipos directivos los resultados obtenidos por sus alumnos.

Las valoraciones de los profesores, padres y alumnos tienen normalmente un gran impacto y conducen a propuestas de cambio: mayor información y atención a los padres, cambios de los tutores el curso siguiente, normas más estrictas para mantener el orden en el centro, al manifestar los alumnos la escasa disciplina existente, mayor información a los profesores por parte del equipo directivo o cambios en la organización del centro. En los años siguientes, el centro puede comprobar el efecto de las iniciativas que ha adoptado, a través de la valoración que cada sector de la comunidad educativa vuelve a expresar.

UNA RED PARA INVESTIGAR LOS FACTORES RESPONSABLES DE LA CALIDAD DE LA ENSEÑANZA

El objetivo de colaborar con cada centro para conocer y mejorar su funcionamiento se completa con la investigación de los factores más directamente responsables de la eficacia de las escuelas. La participación en la red de evaluación de un número importante de centros, la aplicación de pruebas comunes a todos ellos y el seguimiento de las mismas cohortes de alumnos a lo

largo del tiempo, permite aplicar modelos estadísticos de niveles múltiples e investigar la influencia de las distintas variables en los resultados que se obtienen. La utilización, al mismo tiempo, del análisis cualitativo sobre los procesos de centro y de aula, completa la información obtenida en el estudio cuantitativo. Este objetivo entra de lleno en la lógica investigadora por lo que su exposición requiere la descripción de la metodología utilizada y la discusión de los resultados obtenidos.

A continuación se presentan algunos resultados iniciales, correspondientes a los cursos 1997-98 y 1998-99, obtenidos con alumnos que habían ingresado en los centros en septiembre de 1997 para cursar 1.º de ESO y que fueron evaluados en el último trimestre del curso 1998-99, a finales de su 2.º curso de ESO. En todos los resultados que se presentan, las variables analizadas son del nivel de alumno. En trabajos posteriores se tratarán los efectos de variables de los centros sobre los resultados de los alumnos, desde la perspectiva del valor añadido, y utilizando procedimientos multinivel para el análisis de los datos. En este artículo únicamente se consideran dos variables de contexto utilizadas para la formación de los grupos: el tamaño y el contexto sociocultural.

MÉTODO

SUJETOS

Los sujetos participantes en el estudio están escolarizados en los 32 centros que formaban parte de la Red de Evaluación de Educación Secundaria (REDES) en los cursos 1997-98 y 1998-99. El número total de alumnos de 2.º fue de $n = 2.330$. No obstante, los tamaños muestrales efectivos utilizados en los diferentes análisis de datos presentados difieren considerablemente de esta cifra, variando, además, en función de las variables analizadas. Las

principales diferencias se encuentran en los rendimientos académicos, debido a que un considerable número de alumnos no realizó las pruebas de rendimientos académicos del nivel inicial (octubre de 1997, cuando comenzaron 1.º de ESO). En otras variables, las diferencias se deben a la ausencia de algunos alumnos los días y horas en las que se realizaron las pruebas cumplimentadas en 2.º de ESO, protocolos incompletos, etc.

También se obtuvieron datos en 2.º de ESO de los padres de estos alumnos, para evaluar los aspectos relacionados con su satisfacción con los centros. El número de padres que respondieron a los cuestionarios fue $n = 1.217$. La muestra efectiva utilizada en los análisis presentados fue menor ($n = 1.053$) debido fundamentalmente a la presencia de cuestionarios incompletos o mal cumplimentados.

Dado que la participación de los 31 centros en el proyecto REDES es voluntaria, las muestras de sujetos tienen carácter incidental.

PROCEDIMIENTO

Las pruebas de los alumnos fueron aplicadas de forma colectiva en las aulas. Las de resultados académicos de 1.º de ESO y el cuestionario de contexto sociocultural, en octubre de 1997 y las restantes en abril de 1999. Los cuestionarios de los padres fueron aplicados en abril de 1999, enviados a sus casas por medio de los hijos quienes posteriormente, los devolvían al centro.

DESCRIPCIÓN DE LAS PRUEBAS

Todas las pruebas utilizadas en la evaluación fueron desarrolladas por el equipo y probadas en un estudio piloto, en el curso 1996-97, con grupos de alumnos que no participaron posteriormente en la evaluación. A continuación se presenta una breve

descripción de las características técnicas de las pruebas.

MEDIDA DEL CONTEXTO SOCIOCULTURAL

El contexto sociocultural de los alumnos fue evaluado mediante un cuestionario formado por 20 elementos, que hacían referencia a los siguientes aspectos: nivel de estudios de los padres, profesión, expectativas educativas para sus hijos y ciertas posesiones de la familia. De estos últimos elementos fueron eliminados la mayor parte por no mostrar variabilidad. Los alumnos y sus padres respondieron al cuestionario al comienzo del curso 1997-1998. Para todos los ítems se calcularon las correlaciones entre las respuestas de los padres y de los hijos, mediante el coeficiente V de Cramer, mostrando correlaciones elevadas, siempre superiores a 0,65. Aunque la tasa de respuestas de los padres fue elevada (aproximadamente el 70%), debido a la mayor tasa de respuesta de los alumnos y a la correlación mencionada, se utilizaron para el estudio las respuestas de los alumnos. Se realizó un análisis factorial con los elementos del cuestionario que proporcionó un único factor. Se obtuvieron las puntuaciones factoriales para cada sujeto por medio del procedimiento de regresión. Las puntuaciones factoriales fueron transformadas linealmente a una escala con media 100 y desviación típica 15.

MEDIDA DE LOS RESULTADOS EDUCATIVOS EN EL NIVEL INICIAL

Todos los alumnos fueron evaluados al comienzo de 1.º de ESO, para examinar sus niveles de rendimiento inicial en las áreas de Matemáticas, Ciencias sociales y Lengua castellana. Todas ellas fueron desarrolladas a partir del currículo oficial del MEC, en vigor para la ESO. Las pruebas

de área tienen un formato de elección múltiple, con cuatro alternativas. Para su confección se utilizó el modelo psicométrico de la *Teoría de la Respuesta al Ítem*, con 3 parámetros. En todos los casos se examinó la unidimensionalidad por medio del análisis factorial de las matrices de correlaciones tetracóricas, utilizando el programa TESTFACT (Wilson, Wood, Kandola y Gibbons, 1991). En todas las pruebas, fue encontrado un primer factor con autovalor asociado al menos 5 veces mayor que el segundo, por lo que se consideraron unidimensionales y adecuadas para el modelo propuesto de Teoría de Respuesta al Ítem, según el criterio de Lord (1980). Este modelo psicométrico permite la equiparación vertical de las puntuaciones obtenidas en pruebas de niveles escolares diferentes. Para obtener esta equiparación, se dispuso de un conjunto de ítems comunes o de anclaje entre las pruebas de primero y segundo y las de segundo y cuarto. El análisis de los ítems y su ajuste al modelo fue realizado por medio del programa BILOG-3 (Mislevy y Bock, 1990), eliminándose de las pruebas los ítems con pobres ajustes. La equiparación o conversión de las puntuaciones a una escala común a primero y segundo, que es la que se utilizó en este estudio, fue realizada con el programa BILOG-MG (Zimowski, Muraki, Mislevy y Bock, 1996). Todas las puntuaciones fueron convertidas a una escala con media 100 y desviación típica 15.

HABILIDADES METACOGNITIVAS Y ESTRATEGIAS DE APRENDIZAJE

Además de los resultados académicos, los alumnos fueron evaluados en otros aspectos cognitivos: habilidades metacognitivas y estrategias de aprendizaje, utilizando cuestionarios construidos para la investigación.

Habilidades metacognitivas

Intentan informar en qué grado los estudiantes reflexionan sobre sus propios procesos de aprendizaje y solución de problemas, enfrentados a contenidos lo más semejantes posibles a los que se encuentran en las aulas. Esta prueba fue dividida en cuatro bloques teóricos: conciencia del propio conocimiento, metacompreensión, que demanda encontrar contradicciones o incoherencias en la información presentada; verificación de resultados o valoración de las soluciones dadas a los problemas, y conciencia de las estrategias utilizadas, es decir, conciencia de cómo han llegado a determinada solución. Todas las puntuaciones están en la misma escala de 1 a 20.

Estrategias de aprendizaje

Las estrategias o modos de pensar, estudiar y resolver problemas de los alumnos fueron evaluados mediante un cuestionario con formato de respuesta en escala tipo Likert con cinco grados. Las respuestas de una muestra de aproximadamente 2.000 alumnos fueron sometidas a análisis de componentes principales, con rotación varimax, obteniéndose cinco factores: dominio estratégico, o capacidad para tratar la información, diferenciando lo fundamental de lo accesorio; elaboración y organización de la información por medio de resúmenes y esquemas; reflexión sobre el propio trabajo, en cuanto al grado de corrección y las razones para explicarlo; establecimiento de relaciones entre lo que ya sabe y los nuevos contenidos y entre diversas áreas de estudio y aprendizaje memorístico, repetitivo y pasivo. Los coeficientes de fiabilidad como consistencia interna de los cinco factores, obtenidos mediante el coeficiente alpha de Cronbach fueron: 0,86, 0,85, 0,81, 0,68 y 0,65, respectivamente. Todas las puntuaciones fueron convertidas a la misma escala, con valores comprendidos entre 1 y 100.

OPINIÓN-SATISFACCIÓN DE LOS ALUMNOS CON EL CENTRO

Se utilizó un cuestionario en formato de escala tipo Likert con 5 grados. El análisis factorial de los ítems, mediante análisis de componentes principales con rotación varimax, puso de relieve 7 factores: orden en el centro, relaciones con los tutores, relaciones con los profesores, valoración de la preparación recibida, opinión sobre las calificaciones obtenidas, actividades del centro y relaciones con los compañeros.

Las puntuaciones en cada uno de los factores fueron convertidas a una escala con valores comprendidos entre 1 y 100.

Estos cuestionarios fueron recogidos de forma anónima, por lo que no se pueden relacionar con otros datos de los alumnos.

OPINIÓN-SATISFACCIÓN DE LOS PADRES CON EL CENTRO

Un cuestionario de características similares fue contestado por los padres en las mismas fechas, dando lugar a 7 factores: funcionamiento del centro y preparación de los alumnos, atención y comunicación con los padres, participación de los padres, disciplina, actividades del centro, valoración de las calificaciones y opinión sobre el centro.

VARIABLES DE CLASIFICACIÓN

Como ya se ha señalado, los sujetos fueron clasificados según *contexto* y *tamaño* del centro. La puntuación de contexto de los centros se obtuvo mediante la media de los alumnos en el cuestionario ya descrito. Se dividieron los centros en cuatro contextos socioculturales mediante la división por los cuartiles (percentiles 25, 50 y 75). El contexto 1 está formado por estudiantes procedentes de contextos altos, el 2 por sujetos de contextos medio-altos,

el 3, sujetos de contextos medio-bajos, y el 4 por sujetos de contexto sociocultural bajo. Para la clasificación según el tamaño, se consideró el número de unidades por curso de la ESO. Los centros con una o dos unidades por curso fueron considerados centros pequeños, mientras que los que tenían tres o más unidades, se clasificaron como centros grandes

RESULTADOS

LA PREDICCIÓN DEL RENDIMIENTO ACADÉMICO EN 2.º DE ESO A PARTIR DE VARIABLES DE LOS ALUMNOS

Se consideró interesante examinar la influencia de diversas variables de los alumnos sobre sus resultados académicos en 2.º de ESO para cada una de las tres materias. Las variables predictoras utilizadas fueron: rendimiento obtenido en el nivel inicial de la misma materia, puntuación en contexto sociocultural, estrategias de aprendizaje, habilidades metacognitivas y género (introducida como variable ficticia, representando con el valor 0 a los varones y con el 1 a las mujeres). Para establecer qué variables son predictoras del rendimiento y su importancia relativa, se obtuvieron las ecuaciones de regresión lineal múltiple para cada materia, utilizando el procedimiento de pasos sucesivos (*stepwise*). Un resumen de los principales resultados se presenta en la tabla I.

Una observación de la tabla muestra un patrón muy similar para las tres materias. Una vez más se confirman los resultados habituales en los estudios de predicción del rendimiento, en este caso con tests objetivos equiparados, y no con notas escolares. El mejor predictor del rendimiento es el rendimiento anterior en la misma materia, como puede observarse en la tabla. La variable género únicamente es un predictor significativo, con bajo peso, en Matemáticas, aunque a diferencia de

TABLA I

Resultados de los análisis de regresión lineal múltiple (stepwise) en las tres materias

Materias 2.º de ESO	Variables Predictoras	Pesos estandarizados
Matemáticas N = 1184 $F_{7, 1176} = 171,46$ *** R = 0,711 $R^2_{ajustado} = 0,502$ Error típico = 9,64	Matemáticas 1.º	0,491 ***
	Verificación de los resultados	0,221 ***
	Contexto sociocultural	0,090 ***
	Dominio estratégico	0,073 ***
	Conciencia del propio conocimiento	0,081 ***
	Metacomprensión	0,083 ***
	Aprendizaje memorístico	-0,054 *
	Género	0,048 *
Lengua castellana N = 1205 $F_{6, 1198} = 167,58$ *** R = 0,676 $R^2_{ajustado} = 0,454$ Error típico = 9,26	Lengua castellana 1.º	0,469 ***
	Verificación de los resultados	0,140 ***
	Aprendizaje memorístico	-0,127 ***
	Metacomprensión	0,082 ***
	Conciencia de las estrategias	0,078 ***
	Dominio estratégico	0,065**
Ciencias Sociales N = 1094 $F_{7, 1086} = 102,69$ *** R = 0,631 $R^2_{ajustado} = 0,394$ Error típico = 12,33	Ciencias Sociales 1.º	0,391 ***
	Verificación de los resultados	0,157 ***
	Metacomprensión	0,145 ***
	Contexto sociocultural	0,110 ***
	Dominio estratégico	0,112 ***
	Conciencia del propio conocimiento	0,098 ***
	Conciencia de las estrategias utilizadas	0,050 *

Nota: *: $p < .05$; **: $p < .01$; ***: $p < .001$.

otros estudios, el peso positivo indica un mejor rendimiento de las mujeres. Las habilidades metacognitivas se manifiestan como predictoras del rendimiento, especialmente la verificación de los resultados y la metacomprensión, que aparecen en las tres materias; la conciencia del propio conocimiento es importante en Matemáticas y Ciencias sociales, pero no en Lengua. En cuanto a la percepción de las estrategias de aprendizaje utilizadas,

el dominio estratégico aparece con peso significativo en las tres materias; el aprendizaje memorístico es un predictor negativo en Matemáticas y Lengua, aunque no parece afectar significativamente en las Ciencias sociales. Finalmente, el contexto sociocultural del alumno tiene un peso significativo en la predicción del rendimiento en Matemáticas y Ciencias sociales, pero no en Lengua castellana.

DIFERENCIAS EN RESULTADOS ACADÉMICOS EN RELACIÓN CON EL CONTEXTO Y EL TAMAÑO DEL CENTRO

En primer lugar se analizaron las diferencias en resultados académicos en las tres materias evaluadas en 2.º curso de ESO: Matemáticas, Lengua castellana, y Ciencias sociales, en relación con los factores contexto (cuatro niveles) y tamaño del centro (2 niveles), por medio del *análisis de covarianza*, utilizando como covariable para cada una de las materias la medida correspondiente del nivel inicial (1.º de ESO), equiparada con la de 2.º de ESO, por procedimientos de teoría de la respuesta al ítem, como ya se ha señalado anteriormente. Como es frecuente en los estudios no experimentales, existe una confusión entre los efectos principales de tamaño y contexto (el test χ^2 reveló una asociación significativa), no obstante, el efecto de interacción, un objetivo importante en nuestro estudio, no se ve afectado

por dicha asociación en los diseños con dos factores (Jaccard, 1998, p. 85). En los casos en los que se encontraron diferencias relacionadas con el contexto, se analizaron las diferencias, a posteriori, por medio del estadístico de Games-Howell, que no requiere del supuesto de igualdad de varianzas entre los grupos, dado que el supuesto en general no se mantuvo, como puso de relieve el test de Levene ($p < .05$). Para el examen de las diferencias relacionadas con el efecto de interacción, se realizó el contraste de efectos principales simples, dentro de cada uno de los tamaños de centro. En la tabla II se presentan las medias ajustadas (después de la eliminación del nivel inicial de 1.º de ESO) y los resultados más relevantes del análisis de covarianza. Dado que los tamaños muestrales efectivos fueron distintos en cada una de las materias, se presentan también entre paréntesis el número de casos en cada uno de los grupos, para las tres mate-

TABLA II

Resultados del análisis de covarianza contexto \times tamaño, eliminando el efecto de las puntuaciones obtenidas en 1.º de ESO, en Matemáticas, Lengua castellana y Ciencias sociales

Materias	Centros grandes				Centros pequeños			
	C1 (216)	C2 (443)	C3 (386)	C4 (228)	C1 (95)	C2 (115)	C3 (156)	C4 (43)
Matemáticas 2.º ESO	105,21	102,74	100,64	97,85	104,21	105,91	101,38	98,16
$R^2 = 0,463$ ($p < .001$), F (Contexto) _{3, 1673} = 19,31, $p < .001$; F (tamaño) _{1, 1673} = 1,571, ns.; F (Tamaño \times Contexto) _{3, 1673 - 2.18. ns}								
	C1 (212)	C2 (466)	C3 (387)	C4 (265)	C1 (95)	C3 (115)	C3 (150)	C4 (40)
Lengua castellana 1.º ESO	107,02	106,64	104,96	102,57	105,10	104,98	104,58	103,77
$R^2 = 0,398$, $p < .001$; F (Contexto) _{3, 1721} = 2,97, $p < .05$; F (tamaño) _{1, 1721} = 1,15, ns.; F (Tamaño \times Contexto) _{3, 1721} = 2,98, ns)								
	C1 (229)	C2 (393)	C3 (378)	C4 (170)	C1 (58)	C2 (117)	C3 (149)	C4 (43)
Ciencias sociales 1.º ESO	111,53	105,03	103,69	98,08	112,74	109,58	99,04	101,79
$R^2_{ajustado} = 0,390$, $p < .001$; F (Contexto) _{3, 1319} = 40,86, $p < .001$; F (tamaño) _{1, 1319} = 2,02, ns.; F (Tamaño \times Contexto) _{3, 1319} = 9,73, $p < .001$								

En la tabla puede observarse el importante efecto del nivel inicial que tenían los alumnos en 1.º de ESO, sobre los resultados de 2.º, reflejado en los valores de R^2 , en todas las materias. Cuando se eliminan sus efectos, se mantienen diferencias en los resultados, entre los grupos. No se encontraron diferencias relacionadas con el tamaño del centro en ninguna de las materias. Por el contrario, sí hubo importantes diferencias relacionadas con el contexto. El efecto de interacción tamaño \times contexto fue estadísticamente significativo solamente en el caso de las ciencias sociales. Los resultados de las pruebas a posteriori fueron los siguientes:

- Matemáticas. Examinadas las diferencias entre contextos, que se mantienen similares en los dos tipos de tamaño de centro, las diferencias significativas ($p < .05$) encontradas se manifiestan en que los centros de los contextos 1 y 2 muestran resultados superiores a los de los contextos 3 y 4, no encontrándose diferencias entre 1 y 2, ni entre los contextos 3 y 4.
- Lengua castellana. Los resultados en Lengua muestran un patrón algo diferente del anterior en los contrastes a posteriori entre contextos. La única diferencia estadísticamente significativa entre las medias corregidas fue la encontrada entre los contextos 1 y 4, siendo superior el rendimiento de los primeros, manteniéndose esta diferencia en los dos tamaños de centros.
- Ciencias sociales. Como en las otras dos materias, se encontraron efectos relacionados con el contexto, aunque en este caso, también fue significativa la interacción tamaño \times contexto. En las comparaciones entre contextos, se encontraron diferencias significativas entre el contexto 1 y los restantes; también el contexto 2 muestra resultados significativamente superiores a los de los contextos 3 y 4,

no encontrándose diferencias significativas entre estos últimos. En el análisis de los efectos principales simples, se encontró que en los centros grandes se muestra el mismo patrón anterior, mientras que en los pequeños, existen diferencias entre los contextos 1 y 2, considerados globalmente y los contextos 3 y 4.

En el gráfico III se muestra la representación gráfica de las medias corregidas (eliminado el efecto del nivel inicial). G y P hacen referencia a centros grandes y pequeños, respectivamente.

SATISFACCIÓN DE LOS ALUMNOS CON LOS DIVERSOS FACTORES DEL CENTRO

También se examinaron las diferencias entre los grupos definidos por contexto (4 niveles) \times tamaño (2 niveles) en los factores de satisfacción de los alumnos: orden y disciplina en el centro, relaciones con los tutores, relaciones con los profesores, valoración de la preparación recibida, opinión sobre las calificaciones obtenidas, actividades del centro, relaciones con los compañeros.

En la tabla III se presentan las medias y desviaciones típicas de cada uno de los ocho grupos.

En la tabla IV se presentan las medias y desviaciones típicas, entre paréntesis, de los 8 grupos en las siete variables de satisfacción, así como los totales según tamaño de centro y contexto.

Con relación al tamaño de los centros, no se encontraron diferencias significativas en los aspectos relacionados con valoración de la preparación recibida, satisfacción con las calificaciones, ni en relaciones con los compañeros. En las variables orden y disciplina, relaciones con los tutores, relaciones con los profesores y valoración de las actividades del centro, se encontraron diferencias significativas similares, mostrando siempre los centros pequeños medias superiores a las de los centros grandes.

GRÁFICO III

Medias corregidas según tamaño de centro y contexto

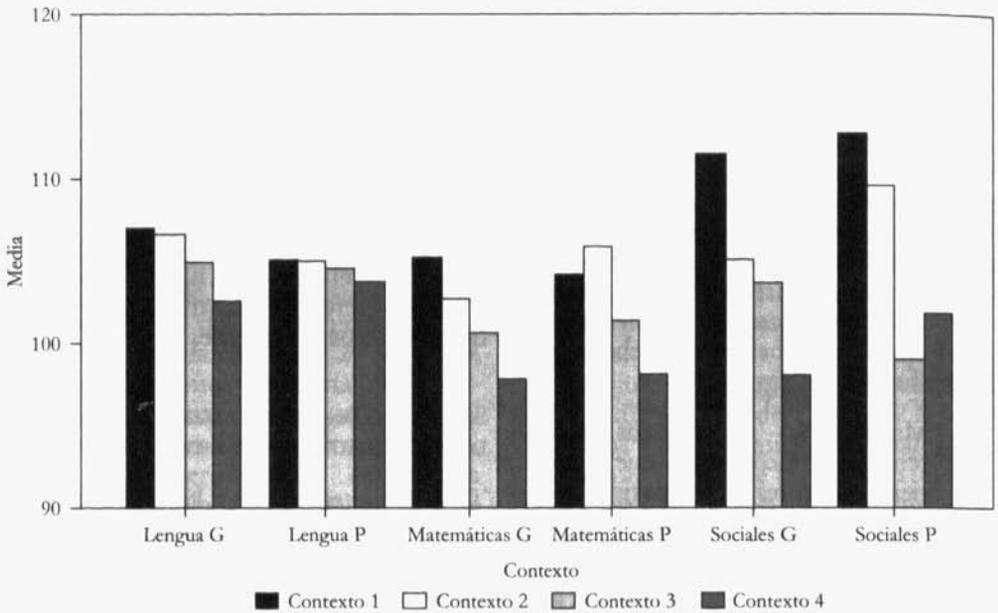


TABLA III

Resultados del análisis de varianza tamaño \times contexto en los factores de satisfacción de los alumnos

Orden y disciplina	F (tamaño) _{1, 2132} = 16,37, p < .001; F (contexto) _{3, 2132} = 4,572, p < .01; F (contexto * tamaño) _{3, 2132} = 19,32, p < .001.
Tutores	F (tamaño) _{1, 2132} = 3,87, p < .05; F (contexto) _{3, 2132} = 4,35, p < .01; F (contexto * tamaño) _{3, 2132} = 18,95, p < .001.
Profesores	F (tamaño) _{1, 2132} = 5,21, p < .05; F (contexto) _{3, 2132} = 5,33, p < .01; F (contexto * tamaño) _{3, 2132} = 19,04, p < .001.
Preparación recibida	F (tamaño) _{1, 2132} = 0,73, ns; F (contexto) _{3, 2132} = 4,72, p < .01; F (contexto * tamaño) _{3, 2132} = 20,66, p < .001.
Calificaciones	F (tamaño) _{1, 2132} = 0,37, ns; F (contexto) _{3, 2132} = 2,62, p < .05; F (contexto * tamaño) _{3, 2132} = 0,78, ns.
Actividades del centro	F (tamaño) _{1, 2132} = 34,91, p < .001; F (contexto) _{3, 2132} = 3,44, p < .05; F (contexto * tamaño) _{3, 2132} = 36,06, p < .001.
Relaciones con compañeros	F (tamaño) _{1, 2132} = 2,06, ns; F (contexto) _{3, 2132} = 8,70, p < .001; F (contexto * tamaño) _{3, 2132} = 4,42, p < .01.

TABLA IV

Medias y desviaciones típicas de los factores de satisfacción del alumno en los ocho grupos

Factores de satisfacción	Grandes (n = 1797) Pequeños (n = 433)	C1 (n = 372)	C2 (n = 577)	C3 (n = 665)	C4 (n = 526)	Total (n = 2140)
Orden y disciplina	Grandes Pequeños	72,20 (13,11) 65,10 (13,81)	62,70 (15,38) 66,67 (13,40)	63,48 (15,11) 68,48 (12,37)	60,82 (15,58) 72,92 (16,62)	63,77 (15,46) 67,43 (13,78)
Total según contextos		69,74 (13,76)	63,59 (15,04)	64,44 (14,75)	61,93 (16,04)	64,51 (15,20)
Relaciones con tutores	Grandes Pequeños	73,13 (13,92) 65,19 (16,89)	66,84 (17,85) 64,68 (17,94)	66,14 (16,81) 72,90 (13,51)	63,64 (15,95) 74,56 (18,27)	66,62 (16,72) 68,34 (16,94)
Total según contextos		70,38 (15,46)	66,36 (17,88)	67,43 (16,44)	64,64 (16,46)	66,97 (16,77)
Relaciones con profesores	Grandes Pequeños	74,13 (12,94) 66,89 (13,42)	66,47 (13,84) 66,87 (13,33)	66,93 (13,26) 71,25 (14,02)	64,11 (12,66) 73,71 (14,46)	67,04 (13,55) 68,92 (13,88)
Total según contextos		74,13 (12,94)	66,47 (13,85)	66,93 (13,25)	64,11 (12,66)	67,42 (13,64)
Preparación recibida	Grandes Pequeños	65,71 (10,40) 56,86 (14,06)	57,82 (12,06) 60,90 (11,98)	58,12 (11,93) 59,86 (10,27)	54,92 (12,06) 59,70 (14,84)	58,23 (12,25) 59,26 (12,59)
Total según contextos		62,64 (12,51)	58,51 (12,10)	58,45 (11,65)	55,36 (12,40)	58,44 (12,32)
Calificaciones	Grandes Pequeños	63,73 (19,03) 63,57 (17,88)	62,49 (17,99) 61,48 (16,59)	60,66 (19,26) 60,48 (14,17)	57,67 (17,98) 61,59 (17,00)	60,74 (18,66) 61,82 (16,36)
Total según contextos		63,68 (18,62)	62,26 (17,68)	60,63 (18,39)	61,82 (16,36)	60,96 (18,22)
Actividades del centro	Grandes Pequeños	55,99 (11,75) 63,57 (13,24)	48,56 (13,49) 61,48 (11,60)	47,63 (12,30) 60,48 (11,52)	45,31 (12,01) 61,59 (13,63)	48,41 (12,90) 52,32 (12,78)
Total según contextos		53,42 (12,77)	49,32 (13,16)	48,71 (12,35)	46,73 (12,95)	49,21 (12,97)
Relación con compañeros	Grandes Pequeños	79,29 (12,54) 75,58 (13,98)	75,25 (14,36) 78,73 (12,22)	72,19 (15,36) 75,34 (13,94)	70,88 (16,21) 72,79 (15,65)	73,62 (15,23) 76,14 (13,75)
Total según contextos		77,94 (13,15)	76,03 (13,98)	72,79 (15,66)	71,05 (16,16)	74,13 (14,97)

Por lo que se refiere al contexto, se encontraron diferencias significativas en los siete factores de satisfacción. Los contrastes a posteriori fueron realizados con el estadístico de Games-Howell, estableciendo el nivel de significación en $p < .05$. En la valoración del orden y disciplina del centro, los centros del contexto 1 fueron superiores a los restantes, no encontrán-

dose diferencias significativas entre éstos. En cuanto a las relaciones con los tutores y con los profesores, se encontró el mismo patrón, mostrándose los centros del contexto 1 superiores a los restantes, y los centros del contexto 3 superiores a los del contexto 4. En la valoración de la preparación recibida, los contextos 1, 2, y 3 superan al contexto 4, pero el contexto 1 también

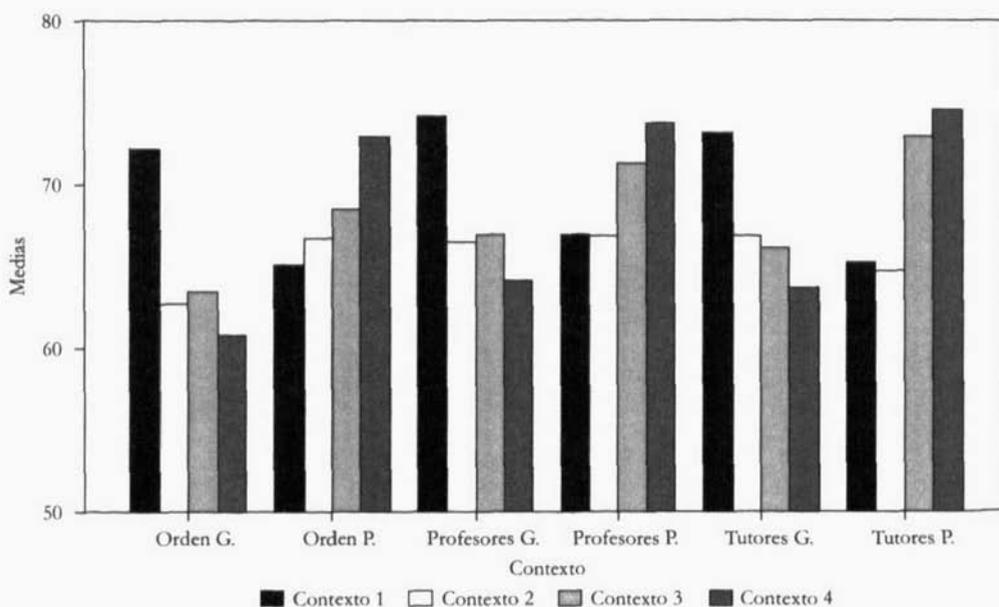
supera a los 2 y 3. Por lo que se refiere a la valoración de las calificaciones del centro, los centros de los contextos 1 y 2 superan a los del contexto 4. Las actividades realizadas en el centro también reciben valoraciones diferentes según los contextos, mostrando valoraciones más altas los de los contextos 1, 2 y 3, que los del contexto 4; también el contexto 1 es superior a los otros 3. Finalmente, en la satisfacción con las relaciones entre compañeros, los centros de los contextos 1 y 2 superan a los de los contextos 3 y 4.

No obstante, algunos de los resultados anteriores relativos a los efectos principales deben reexaminarse a la luz de algunos importantes efectos de interacción encontrados, que se encuentran en todas las variables excepto en la valoración de las calificaciones recibidas. Seguramente son éstos los resultados más relevantes de la investigación. Los efectos de interacción fue-

ron examinados por medio de los *efectos principales simples* (Kirk, 1995, pp. 377-381), examinando las diferencias dentro de cada uno de los tamaños, con un nivel de significación fijado de $p < .05$. En todos los casos el patrón de resultados es básicamente similar, mostrándose en los centros grandes valoraciones más altas en los centros del contexto 1, y más bajas en los centros del contexto 4. Por el contrario, en los centros pequeños, se invierte el patrón de valoración, mostrando en general los centros del contexto 4 la mayor satisfacción en las dimensiones evaluadas. Estos resultados se aprecian bien en la figura 4, en la que se presentan gráficamente los resultados para las tres dimensiones en las que se observa de forma más clara este efecto: valoración del orden y la disciplina, relaciones con los tutores y relaciones con los profesores. En la figura, G hace referencia a los centros grandes y P, a los pequeños.

GRÁFICO IV

Satisfacción de los alumnos según contexto y tamaño del centro



**RESULTADOS RELATIVOS A LA
SATISFACCIÓN DE LOS PADRES**

También se examinaron las diferencias entre los grupos definidos por contexto (4 niveles) \times tamaño (2 niveles) en los factores de satisfacción de los padres: orden y disciplina en el centro, relaciones con los tutores, relaciones con los profesores, valoración de la preparación recibida, opinión sobre las calificaciones obtenidas, actividades del centro, relaciones con los compañeros.

En la tabla V se presentan las medias y desviaciones típicas de cada uno de los ocho grupos.

En la tabla VI se presentan las medias y desviaciones típicas, entre paréntesis, de los 8 grupos en las siete variables de satisfacción, así como los totales según tamaño de centro y contexto.

Con relación al tamaño de los centros, se encontraron diferencias significativas en los siete factores, siendo superior-

res siempre las valoraciones de los padres de los centros pequeños.

Por lo que se refiere al contexto, se encontraron diferencias significativas en todos los factores, excepto en la satisfacción con las calificaciones de los alumnos. Los contrastes a posteriori fueron realizados con el estadístico de Games-Howell, estableciendo el nivel de significación en $p < .05$. En los seis factores en los que se mostraron diferencias significativas, los padres de los centros del contexto 1 mostraron puntuaciones medias superiores a las de los otros tres contextos, excepto en participación, en que únicamente superan a los de los contextos 2 y 4. En participación y en valoración del orden y disciplina del centro, los centros del contexto 3 superan a los del contexto 2. Finalmente, en opinión sobre el centro, también se encontró una superioridad de los centros del contexto 3 frente a los del contexto 4.

TABLA V

*Resultados del análisis de varianza tamaño \times contexto
en los factores de satisfacción de los padres*

Funcionamiento del centro	F (tamaño) _{1, 1045} = 17,63, $p < .001$; F (contexto) _{3, 1045} = 4,482, $p < .05$; F (contexto * tamaño) _{3, 1045} = 6,60, $p < .01$.
Atención recibida	F (tamaño) _{1, 1045} = 19,31, $p < .001$; F (contexto) _{3, 1045} = 9,19, $p < .01$; F (contexto * tamaño) _{3, 1045} = 4,97, $p < .05$.
Participación en el centro	F (tamaño) _{1, 1045} = 33,09, $p < .001$; F (contexto) _{3, 1045} = 11,50, $p < .01$; F (contexto * tamaño) _{3, 1045} = 12,82, $p < .001$.
Orden y disciplina	F (tamaño) _{1, 1045} = 12,73, $p < .001$; F (contexto) _{3, 1045} = 8,58, $p < .01$; F (contexto * tamaño) _{3, 1045} = 4,60, $p < .01$.
Actividades del centro	F (tamaño) _{1, 1045} = 31,02, $p < .001$; F (contexto) _{3, 1045} = 5,70, $p < .001$; F (contexto * tamaño) _{3, 1045} = 13,89, $p < .001$.
Calificaciones recibidas por los alumnos	F (tamaño) _{1, 1045} = 20,36, $p < .001$; F (contexto) _{3, 1045} = 2,55, ns; F (contexto * tamaño) _{3, 1045} = 1,21, ns.
Opinión sobre el centro	F (tamaño) _{1, 1045} = 21,50, $p < .001$; F (contexto) _{3, 1045} = 5,93, $p < .001$; F (contexto * tamaño) _{3, 1045} = 9,39, $p < .001$.

TABLA VI
*Medias y desviaciones típicas
de los factores de satisfacción de los padres en los ocho grupos*

Factores de satisfacción	Grandes (n = 801) Pequeños (n = 252)	C1 (n = 372)	C2 (n = 577)	C3 (n = 665)	C4 (n = 526)	Total (n = 1053)
Funcionamiento del centro	Grandes Pequeños	79,03 (10,18) 76,02 (11,43)	69,95 (13,01) 74,79 (12,42)	71,22 (13,49) 74,82 (13,07)	68,45 (13,82) 80,22 (10,10)	70,70 (13,42) 75,88 (12,20)
Total según contextos		77,52 (10,88)	70,91 (13,02)	72,09 (13,46)	70,40 (13,96)	71,92 (13,31)
Atención recibida	Grandes Pequeños	77,64 (13,77) 79,18 (13,97)	71,20 (16,18) 71,84 (16,42)	72,14 (17,14) 78,02 (14,82)	72,65 (15,51) 85,62 (10,17)	72,32 (16,15) 77,50 (15,17)
Total según contextos		78,44 (13,84)	71,33 (16,21)	73,49 (16,80)	74,74 (15,52)	73,57 (16,15)
Participación en el centro	Grandes Pequeños	66,41 (13,56) 64,92 (14,88)	58,41 (16,84) 60,93 (15,89)	63,20 (14,94) 69,13 (15,66)	57,78 (15,79) 78,36 (13,14)	60,56 (15,96) 67,07 (16,17)
Total según contextos		65,66 (14,19)	58,95 (16,66)	61,26 (17,19)	61,26 (17,19)	62,12 (16,24)
Orden y disciplina	Grandes Pequeños	77,69 (9,75) 75,15 (12,94)	67,07 (14,53) 70,62 (12,82)	70,15 (14,46) 74,13 (14,31)	68,79 (16,15) 79,62 (12,40)	69,34 (14,83) 74,20 (13,55)
Total según contextos		76,49 (11,40)	67,80 (14,25)	71,05 (14,50)	70,65 (16,08)	70,48 (14,68)
Actividades del centro	Grandes Pequeños	70,04 (12,95) 63,39 (13,51)	56,23 (19,53) 64,56 (16,41)	58,81 (17,43) 65,27 (18,32)	55,49 (18,07) 77,36 (12,70)	57,96 (18,39) 66,27 (16,53)
Total según contextos		66,72 (13,60)	58,02 (19,19)	60,24 (17,81)	58,82 (19,94)	59,91 (18,39)
Calificaciones de los alumnos	Grandes Pequeños	62,58 (12,97) 70,45 (10,40)	61,26 (12,79) 63,82 (14,19)	61,64 (13,06) 64,91 (14,02)	61,08 (13,55) 65,87 (15,61)	61,45 (13,06) 65,99 (13,73)
Total según contextos		66,48 (12,37)	61,80 (13,12)	62,40 (13,34)	61,86 (13,99)	62,54 (13,36)
Opinion sobre el centro	Grandes Pequeños	77,34 (13,89) 72,35 (18,15)	63,49 (18,05) 69,48 (18,94)	66,52 (20,04) 72,12 (19,21)	58,89 (18,33) 78,13 (19,76)	64,43 (19,10) 72,22 (19,06)
Total según contextos		74,81 (16,32)	64,76 (18,39)	67,82 (19,97)	61,92 (19,80)	66,30 (19,37)

No obstante, igual que en el caso de la satisfacción de los alumnos, la mayor parte de los anteriores factores relativos a los efectos principales deben reexaminarse, a la luz de los importantes efectos de interacción encontrados, que se encuentran

en todas las variables excepto en la valoración de las calificaciones recibidas. Los resultados son muy similares a los encontrados en los alumnos. Como en el caso anterior, los efectos de interacción fueron examinados por medio de los *efectos*

principales simples, examinando las diferencias dentro de cada uno de los tamaños, con un nivel de significación fijado de $p < .05$. En todos los casos, el patrón de resultados es básicamente similar, mostrándose, en los centros grandes, valoraciones más altas en los centros del contexto 1, que en las de los centros de los restantes contextos. Por el contrario, en los centros pequeños, se invierte el patrón de valoración, mostrando, en general, los centros del contexto 4, la mayor satisfacción en las dimensiones evaluadas.

Los anteriores resultados se aprecian bien en la figura V, en la que se presentan gráficamente los resultados para tres de las dimensiones en las que se observa claramente este efecto: satisfacción con la atención recibida, valoración de la participación en el centro y opinión sobre el centro. En el gráfico, G hace referencia a los centros grandes y P, a los pequeños.

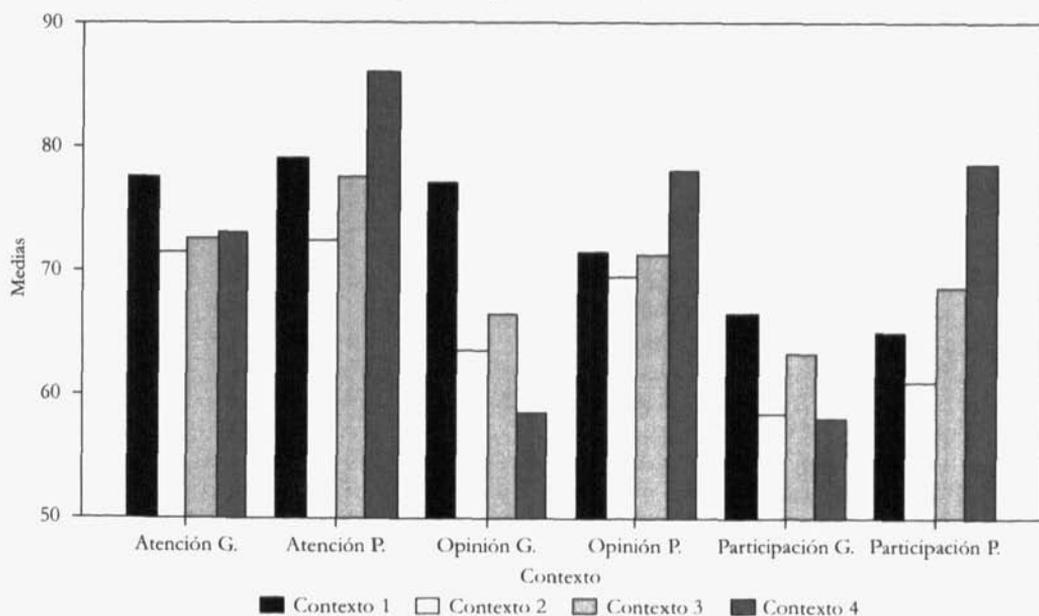
DISCUSIÓN

Hay que tener en cuenta que los resultados presentados anteriormente son sólo una parte de los obtenidos ya que no se han incluido los efectos de las variables de los procesos de centro y de aula, ni se han utilizado procedimientos de niveles múltiples para el análisis de los datos. Estas limitaciones no impiden, sin embargo, que los datos recogidos presenten unas tendencias sólidas y sugestivas.

El primer bloque de datos analiza las variables de los alumnos que predicen los resultados que obtienen en 2.º de la ESO. El mejor predictor en todas las áreas es el rendimiento anterior en esas mismas áreas, lo que pone de relieve la importancia de controlar esta variable cuando se trata de estudiar la calidad de los centros docentes. La ausencia de este control y el análisis directo de los resultados finales de

GRÁFICO V

Satisfacción de los padres según contexto y tamaño del centro



los alumnos ofrece una imagen distorsionada del esfuerzo que cada centro realiza para mejorar los aprendizajes de los alumnos. Este dato refuerza la importancia de los modelos de valor añadido que toman medidas repetidas de los resultados de los mismos alumnos. El modelo de evaluación de REDES comparte este planteamiento.

En ocasiones y ante la ausencia de los resultados iniciales de los alumnos, se han utilizado medidas de su contexto sociocultural para controlar su influencia y ofrecer comparaciones más equilibradas a partir de los residuales obtenidos. Los datos obtenidos ponen de relieve que este método es mejor que la comparación de las puntuaciones directas pero claramente insuficiente si no se tienen en cuenta las puntuaciones iniciales. El contexto sociocultural tiene un peso significativo para predecir el rendimiento pero está muy alejado del que tienen los resultados iniciales. Un dato sorprendente es la ausencia de relación entre el contexto sociocultural y los resultados que obtienen los alumnos en Lengua ya que hay una amplia evidencia empírica de esta relación. Posiblemente el tipo de prueba, basada en la comprensión y en el análisis de diferentes tipos de textos frente a otras pruebas más orientadas hacia la expresión y hacia la estructura del texto, sea el responsable de este hecho. En todo caso, hay que esperar a nuevas medidas longitudinales para confirmarlo.

La influencia de las habilidades metacognitivas de los alumnos, especialmente la capacidad de verificar los resultados y la metacompreensión, en los resultados que obtienen en las tres áreas curriculares, pone de relieve su importancia en el aprendizaje. El desarrollo metacognitivo no debe considerarse, a la luz de estos datos, una actividad marginal en el proceso de enseñanza sino que debe constituir uno de sus ejes prioritarios. La incorporación explícita de este objetivo, en la enseñanza de las distin-

tas áreas curriculares, contribuye a mejorar sus aprendizajes. Igualmente existe una relación positiva entre la conciencia de los alumnos de que disponen de un amplio dominio de estrategias para aprender y los resultados de sus aprendizajes en las tres áreas estudiadas. Por el contrario, la negativa relación entre la utilización de los alumnos de la memorización como estrategia de aprendizaje y los resultados que obtienen, pone de manifiesto la insuficiencia de esta estrategia y la necesidad de orientar a los alumnos hacia aquellas actividades relacionadas con los aprendizajes eficaces.

Finalmente, conviene destacar que el género tiene un escaso valor predictivo. Sólo se encuentra una relación débil en el área de Matemáticas y a favor de las mujeres. El estudio realizado por el INCE (1998) sobre la educación secundaria obligatoria (ESO) muestra una tendencia diferente. A los 14 años, las mujeres son mejores en diversos aspectos lingüísticos y peores en Matemáticas y en Ciencias sociales. Frente a estos datos, la opinión mayoritaria de los profesores es que las mujeres se desenvuelven mejor que los hombres en la ESO, lo que estaría más en consonancia con lo datos obtenidos en la presente investigación. Sin embargo, será necesario completarla para comprobar el efecto del género en el conjunto de la etapa.

Un segundo grupo de datos se refiere a los resultados académicos que obtienen los alumnos al final de 2.º de la ESO, una vez controlado el efecto del nivel inicial de los alumnos, al comenzar a estudiar esta etapa educativa. En las tres áreas estudiadas se encuentra un significativo efecto de la variable sociocultural. En Matemáticas y Ciencias sociales, las diferencias están entre los dos primeros contextos frente a los dos últimos, mientras que en Lengua, las diferencias son menores, sólo relevantes entre el contexto 1 y el 4.

El tamaño del centro no influye en los resultados de los alumnos en ninguna de las tres materias estudiadas. Sin embargo,

se encuentra una interacción entre tamaño y contexto en el área de Ciencias sociales: los alumnos de los centros pequeños, del contexto más bajo, obtienen mejores resultados que los de los centros grandes. Este dato va a manifestarse con mucha mayor amplitud en las otras dos variables estudiadas relacionadas con los resultados de un centro: la valoración de los alumnos y de los padres.

Los datos obtenidos sobre la valoración de los alumnos ofrecen sugerencias interesantes. Por una parte, existen diferencias significativas en los siete factores estudiados en relación con el contexto socioeconómico. Los centros situados en los contextos más altos tienen valoraciones más positivas. Estas diferencias también se manifiestan en el tamaño: los centros pequeños tienen puntuaciones más altas en cuatro dimensiones: orden, relación con los tutores, relación con los profesores y actividades del centro. Este dato es importante para impulsar programas de mejora de los centros: la organización del centro en unidades más reducidas puede tener un efecto mucho más beneficioso que cualquier otra iniciativa.

Sin embargo, el resultado más relevante encontrado es la interacción entre contexto y tamaño en las dimensiones de orden, valoración de los profesores y valoración de los tutores, tres dimensiones de indudable importancia en el ámbito educativo. Los alumnos de los contextos más bajos (3 y 4), que estudian en centros pequeños, ofrecen las valoraciones más altas en estas tres dimensiones. Sólo los centros grandes del contexto alto se aproximan a ellas, pero en este caso hay que explicarlo por la presencia, en este último grupo, de un gran centro muy bien valorado por sus alumnos. ¿Cómo se puede explicar este tipo de resultados? Se pueden apuntar dos interpretaciones. En primer lugar, los centros pequeños permiten una relación más estrecha entre los profesores y sus alumnos, lo que favorece la convivencia y

la aceptación de las normas establecidas. ¿Pero por qué hay diferencias tan claras, entre los contextos superiores e inferiores, a favor de estos últimos, en los centros pequeños? Aquí está la segunda interpretación. En este tipo de centros existe un profesorado más motivado y comprometido con sus alumnos. A su vez, los alumnos, muchos de los cuales viven en situaciones sociales y familiares difíciles, valoran más el afecto y la dedicación de sus profesores. El análisis de los procesos de centro y el seguimiento de los alumnos puede ayudar a confirmar estas interpretaciones en el futuro. Conviene destacar, finalmente, que la mayor satisfacción de los alumnos en estos centros sólo está asociada a mejores rendimientos académicos en Ciencias sociales, pero no en las otras áreas estudiadas.

La valoración de los padres sigue una pauta similar. Existen diferencias significativas en relación con el contexto sociocultural de los alumnos en la mayoría de los factores, si bien la diferencia más acusada se encuentra entre el contexto 1 y el resto. En el caso del tamaño de los centros, hay diferencias significativas en todos los factores, a favor de los centros pequeños. Finalmente, también se encuentra un importante efecto de interacción contexto-tamaño. Los padres de los centros pequeños, del contexto sociocultural más bajo, dan las valoraciones más altas en el funcionamiento del centro, en la atención recibida, en la participación, en el orden, en las actividades del centro y en la opinión sobre el centro. Sólo en las calificaciones que reciben los alumnos son superados por los padres de centros pequeños, de contexto sociocultural alto. Las explicaciones a este hecho son similares a las anteriormente expuestas. Los padres de estos centros se encuentran bien acogidos por profesores motivados, perciben que sus hijos están contentos y se encuentran muy satisfechos de que sus hijos puedan progresar académicamente mucho más de lo que ellos pudieron.

CONCLUSIÓN: LAS POSIBILIDADES DE UNA RED DE EVALUACIÓN

En síntesis, el planteamiento de la red de evaluación sienta las bases para superar muchas de las insuficiencias que existen, tanto en las investigaciones sobre la eficacia como en los programas para mejorar las escuelas (Reynolds y Stoll, 1996). En relación con los estudios sobre la eficacia, las principales aportaciones pueden resumirse en las siguientes:

- Se incluyen estudios de casos a través de una metodología cualitativa para completar la información cuantitativa.
- Se tienen en cuenta las condiciones específicas de las escuelas y sus objetivos y procesos internos.
- Se adopta una perspectiva dinámica para estudiar el cambio en las escuelas.
- No sólo se obtienen datos sobre los resultados, sino que además se estudian los factores más directamente responsables de los mismos.
- Se amplía el concepto de resultados, tanto por tener en cuenta otras dimensiones de los alumnos (al incorporar el estudio de sus actitudes, habilidades metacognitivas y estrategias de aprendizaje), como por incluir la valoración de los padres, de los alumnos y de los profesores sobre el funcionamiento del centro.

A su vez, la red de evaluación contribuye a enriquecer los proyectos de mejora con las siguientes propuestas:

- Se amplía el conocimiento disponible con nuevos datos y explicaciones que proceden de la investigación realizada. La importancia del tamaño y su interacción con el contexto sociocultural del centro, una de las conclusiones de la inves-

tigación realizada, es una información muy relevante sobre las estrategias de cambio.

- Los centros, al sentirse participantes, relacionan más fácilmente las conclusiones obtenidas en la investigación con su realidad.
- Se aportan datos y relaciones entre los distintos niveles educativos que ayudan a seleccionar la estrategia de cambio más adecuada.
- Se evalúa el efecto de los cambios a lo largo del tiempo.
- Se enriquece el conocimiento de las condiciones que influyen en el funcionamiento de las escuelas y se presta una mayor atención a las valoraciones que padres, profesores y alumnos realizan sobre el centro.

Para seguir avanzando en el campo de la evaluación de centros, hacen falta combinaciones de estrategias no reduccionistas, que aporten información amplia y contextualizada y contrasten los resultados obtenidos. La red de evaluación que se ha impulsado, por la amplitud de sus objetivos, por los criterios que utiliza y por la complementariedad de sus metodologías, puede contribuir a este objetivo.

Los efectos que está produciendo el desarrollo de la red de evaluación manifiestan sus enormes posibilidades para contribuir a la mejora de las escuelas y para estudiar, a partir de los mismos centros participantes, los factores responsables de su eficacia o ineficacia. La tradición de las escuelas eficaces y el movimiento para la mejora de las escuelas, diferentes en sus objetivos y en su metodología, encuentran un lugar común en esta red de evaluación. No quiere eso decir que se fundan en un único proyecto sino que en él se desarrollan de manera diferenciada pero interconectada.

La investigación sobre la eficacia de las escuelas que participan en la red y sobre los factores responsables de la misma no solo permite aumentar la comprensión

del funcionamiento de las escuelas sino que posibilita a cada uno de los centros participantes un doble objetivo. Por una parte, conocer su situación analizándola en relación con otros centros y comprobar su evolución a lo largo del tiempo. Por otra, introducir posibles mejoras a partir de la información obtenida y hacer un seguimiento de las consecuencias de estos cambios en el funcionamiento del centro.

BIBLIOGRAFÍA

- BRYK, A. S. y RAUDENBUSH, S. W.: *Hierarchical linear models*. Thousand Oaks, CA, Sage, 1992.
- CREEMERS, B.: *The Effective Classroom*. London, Cassell, 1994.
- GOLDSTEIN, H.: *Multilevel statistical models*. London, Arnold. 2.ª ed., 1995.
- GRAY, J.; JESSON, D. y SIME, N.: «Estimation differences in the examination performances of secondary schools in six LEAs: a multi-level approach to school effectiveness», en *Oxford Review of Education*, 16, 2 (1990), pp. 137-158.
- HOPKINS, D.: «Les indicateurs de processus au service de l'amélioration de l'école», en OCDE, *Évaluer l'enseignement. De l'utilité des indicateurs internationaux*. París, OCDE, 1994.
- HOPKINS, D.; WEST, M. y AINSCOW, M.: *Improving the Quality of Education for All. Progress and Challenge*. London, David Fulton, 1996.
- JACCARD, J.: *Interaction effects in factorial analysis of variance*. Thousand Oaks, CA, Sage, 1998.
- KIRK, R. E.: *Experimental Design: procedures for the behavioral sciences*. Pacific Grove, CA, Brooks/Cole Pub, 3.ª ed., 1995.
- LORD, F. M.: *Applications of item response theory to practical testing problems*. Hillsdale, NJ, Lawrence Erlbaum, 1980.
- MISLEVY, R. J. y BOCK, D. R.: *BILOG: Item analysis and test scoring with binary logistic models*. Chicago, SI-Scientific Software International, 1990.
- RAYWID, M.: «Small Schools: A reform that works», en *Educational Leadership*, 55, 4 (1997), pp. 34-39.
- REYNOLDS, D., et al. (eds.): *Making Good Schools. Linking School Effectiveness and School Improvement*. London, Routledge.
- REYNOLDS, D. y STOLL, L.: «Merging school effectiveness and school improvement: the knowledge bases», en D. REYNOLDS, et al. (eds.), *Making Good Schools. Linking School Effectiveness and School Improvement*. London, Routledge, 1996.
- SAMMONS, P., et al.: *Assessing School effectiveness. Developing measures to put school performance in context*. Londres, Institute of Education y OFSTED, 1994.
- SLAVIN, R.: *Education for all*. Lisse, Swets & Zeitlinger, 1996.
- STOLL, L. y FINK, D.: *Changing our Schools*. Bristol, Open University Press, 1996.
- THOMAS, S. y MORTIMORE, P.: «Comparison of value added models for secondary school effectiveness», *Paper presented at the Annual Conference of the British Educational Research Association*. Bath, 1995.
- WILSON, D. T., et al.: *TESTFACT: Test scoring, item statistics, and item factor analysis*. Chicago, SSI-Scientific Software International, 1991.
- ZIMOWSKI, M. F., et al.: *BILOG-MG: Multiple-Group IRT analysis and test maintenance for binary items*. Chicago, SSI-Scientific Software International, 1996.