



Para citar este artículo, le recomendamos el siguiente formato:

Lukas, J. F. y Santiago, K. M. (2004). Evaluación de centros escolares de educación secundaria del País Vasco. *Revista Electrónica de Investigación Educativa*, Vol. 6 (2). Consultado el día de mes de año en: <http://redie.uabc.mx/vol6no2/contenido-lukas.html>

Revista Electrónica de Investigación Educativa

Vol. 6, No. 2, 2004

Evaluación de centros escolares de educación secundaria del País Vasco

Evaluation of Secondary Education Schools in the Basque Autonomous Community

José Francisco Lukas Mujika (*)
plplumuj@sf.ehu.es

Karlos Santiago Etxeberria (*)
plpsaetc@sf.ehu.es

* Departamento de Métodos de Investigación y Diagnóstico en Educación
Facultad de Filosofía y Ciencias de la Educación
Universidad del País Vasco

Avenida de Tolosa, 70
San Sebastián - 20.018
Gipuzkoa, España

(Recibido: 30 de diciembre de 2003; aceptado para su publicación: 10 de marzo de 2004)

Resumen

El objetivo del presente artículo es presentar el modelo de evaluación de centros escolares de educación secundaria que se aplica en la Comunidad Autónoma Vasca (España), el cual que tiene como propósito ofrecer información relevante para instaurar procesos de cambio y mejora institucional en cada una de las escuelas participantes. Se

analizan, por una parte, las bases teóricas que subyacen detrás de él y, por otra, los aspectos metodológicos y técnicos más relevantes de esta experiencia. Junto con la descripción del modelo de evaluación se presenta una descripción detallada de las dimensiones analizadas, así como de los procesos de recogida y construcción de las pruebas y cuestionarios que se utilizan, la situación actual y las perspectivas de cara al futuro.

Palabras clave: Evaluación de centros escolares, educación secundaria, evaluación formativa.

Abstract

The objective of this paper is to present the school evaluation model that is being applied in different centres of the Basque Autonomous Community. It analyses, on the one hand, the theoretical bases that underlie it and, on the other, the more relevant methodological and technical aspects of this experience. The evaluation attempts to offer relevant information to set up processes of change and improvement in each one of the participant centers. Along with the description of the evaluation model that is being applied, the reader will find in this article a detailed account of the analysed dimensions and the processes of collection and construction of the tests and questionnaires that are being used, as well as that of the present situation and the perspective facing the immediate future.

Key words: School's evaluation, secondary education, formative evaluation.

Necesidad de la evaluación de centros

La evaluación de centros escolares es un ámbito de evaluación importante en la actualidad. No obstante, hasta hace unas décadas éstos no suscitaban el interés que provocan hoy, y ello ha sido debido a diversas razones. En primer lugar, debemos citar los famosos informes de Coleman *et al.* (1966) y Jencks *et al.* (1972) que demostraron que los antecedentes socioculturales y económicos de los alumnos condicionaban su rendimiento posterior. Sin embargo, no es menos cierto que en investigaciones posteriores se descubrió que alumnos con antecedentes similares que asistían a centros educativos distintos obtenían un rendimiento diferente. Es decir, se descubrió otro factor que podía explicar parte de las diferencias entre los alumnos y al que se denominó *efecto escuela*. Sin embargo, hay otras razones que han contribuido a que los administradores, educadores e investigadores centren su atención en la escuela (Pérez Juste, 1999; Tiana, 1999; Lukas, Santiago, Munarriz, Moyano y Sedano, 2000), como:

- La necesidad de información que muestra la sociedad tanto respecto al sistema educativo en general como respecto a cada centro educativo en particular.
- La idea de que el centro educativo contribuirá a la mejora cualitativa de la educación.

- El aumento de la autonomía de los centros que lleva como contrapartida la rendición de cuentas.

La evaluación de centros educativos es una tarea muy compleja, sobre todo en la práctica, debido a razones técnicas, presiones ambientales, falta de evaluadores cualificados y dificultades para conseguir la participación y colaboración necesarias (Escudero, 1997). Evaluar un centro educativo supone responder a tres cuestiones fundamentales, a saber: ¿qué queremos evaluar?, ¿cómo lo vamos a hacer? y ¿para qué vamos a evaluar? (De Miguel, 1997).

A continuación se expone el modelo de evaluación externa de centros de educación secundaria obligatoria (ESO), denominado REDES, que ha sido puesto en funcionamiento en el País Vasco.

Implantación del modelo REDES en el País Vasco

En 1996 un grupo de profesores de educación secundaria y de universidad de la Comunidad de Madrid decidieron crear una red de evaluación externa de centros educativos de la ESO, como una aportación que se podía hacer a los mismos y desde una perspectiva diferente a la que hasta aquel momento se había desarrollado en España (Marchesi, 2002). Esta decisión se tomó con base en tres razones:

- 1) La inexistencia de tradición de evaluación de centros en España.
- 2) La creencia de que era positivo ofrecer un nuevo modelo de evaluación que completara los que se estaban utilizando y ofreciera alternativas viables, rigurosas y atractivas para la comunidad educativa. Los modelos de evaluación puestos en práctica tenían un carácter formativo y conjugaban la evaluación externa con la interna; características muy positivas, pero apenas utilizaban pruebas e instrumentos cuantitativos y rechazaban la posibilidad de ofrecer datos contextualizados y comparativos. El nuevo modelo intentaba superar esas limitaciones sin caer en una evaluación reduccionista y jerarquizadora de los centros.
- 3) La consideración de que una evaluación equilibrada podía ser un poderoso instrumento de cambio y de mejora.

En 1999 el equipo evaluador con financiación de la Fundación Santa María creó el Instituto de Evaluación y Asesoramiento Educativo (IDEA), organizado de manera descentralizada y con varias sedes en distintas comunidades autónomas, entre ellas la del País Vasco.

En el curso 1999-2000 inició la evaluación en el País Vasco y desde entonces participan 55 centros de enseñanza secundaria obligatoria (para jóvenes de 12 a 16 años). Como se ha señalado previamente, está promovida y financiada por el Departamento de Educación del Gobierno Vasco y es el profesorado de la Universidad del País Vasco el encargado de llevarla a la práctica.

En el convenio firmado entre el IDEA y la Universidad del País Vasco se contempló un plan de evaluación para cuatro cursos consecutivos. Al mismo tiempo en otras comunidades autónomas del Estado español (Cataluña, Madrid, Andalucía, etc.) se está llevando a la práctica el mismo plan de evaluación. Esto dará la oportunidad de tener puntuaciones de referencia de diferentes comunidades autónomas.

Objetivo de la evaluación

El objetivo es que cada uno de los centros participantes obtenga información acerca de su funcionamiento y de sus resultados, que les posibilite tomar las decisiones que consideren oportunas para mejorar su propio desempeño y el rendimiento de su alumnado. Más concretamente, se pretende que cada uno de los centros conozca, de la manera más amplia posible, cuál es su situación en los distintos aspectos que se evalúan, cuál es la evolución de su alumnado teniendo como referencia los indicadores de los centros de su mismo contexto sociocultural, y que conozca sus propios cambios. A fin de cuentas, la clave de esta evaluación es ayudar a que en los centros se den procesos internos para su mejora. Para ello, deben conocer su funcionamiento con un cierto nivel de rigor.

Por lo tanto, no se pretende controlar, seleccionar ni ordenar a los centros educativos. No se trata tampoco de que los alumnos pasen una serie de pruebas de rendimiento para comprobar su preparación con una función acreditativa. La finalidad, como ya se señaló, es mucho más amplia y global: pretende colaborar en el diagnóstico de las escuelas para detectar con ellas sus puntos fuertes y débiles, y para que, desde un conocimiento más preciso de su situación, sean capaces de diseñar proyectos de mejora y cambio.

Características de la evaluación

El modelo de evaluación de los centros se configura en una estructura compartida, y con él se pretende ofrecer beneficios para compensar el esfuerzo realizado. Por ello, se aboga por un modelo en el que la información se devuelva de manera ágil y rápida, amplio y equilibrado, capaz de reflejar la realidad del centro educativo y de ofrecer información contextualizada para evitar comparaciones inadecuadas; que amplíe la información que tienen los centros sobre su realidad y, por último, que les ayude a impulsar programas de cambio y mejora (Marchesi, 2002).

Las características del modelo de evaluación REDES son las siguientes:

a) Evaluación externa

Indudablemente y en el sentido de que los profesionales que elaboran y aplican las pruebas son agentes ajenos a los propios centros, nos encontramos ante una

evaluación externa con las ventajas y limitaciones que lleva consigo. Esta evaluación no pretende suplantar la que el profesorado lleva a cabo en su centro correspondiente. Sin embargo, puede utilizarse como complementaria, en el sentido de que ofrece un conjunto de datos objetivos, cuya interpretación y valoración debe realizar el personal del propio centro con base en sus propios criterios. De alguna manera, esta evaluación se puede plantear como ayuda a la del propio centro. Es el personal de éste el que mejor puede entender la importancia de los resultados obtenidos y puede explicarlos con mayor rigor, ya que son los maestros quienes conocen la historia y cultura del centro, los problemas actuales o pasados y las decisiones tomadas. La combinación de las dos evaluaciones puede redundar en un conocimiento más completo de la realidad educativa de cada caso. Como menciona Tiana (1999), y huyendo de clasificaciones dicotómico-antagónicas, podríamos hablar de una *evaluación externa* con carácter *formativo* (Marchesi y Martín, 1999b).

b) Información contextualizada

Es obvio que no se pueden comparar escuelas directamente con base en las puntuaciones obtenidas en distintos tests de rendimiento, si éstas acogen alumnos cuya procedencia socioeconómica y cultural es diferente. No se cumpliría el principio de justicia propuesto por House (1994). Una comparación justa implicaría neutralizar el efecto que la extracción social y cultural del alumnado tiene sobre el rendimiento. Como señalan diversos autores, una solución a este problema es la comparación de escuelas con características similares, es decir, valorar los resultados logrados a partir de un nivel teóricamente homogéneo (Webster, Mendro y Almaguer 1994; Fernández Díaz y González Galán, 1997; Tiana, 1997). En el modelo REDES así se hace. Cada una de las escuelas se clasifica en uno de los cuatro contextos que se han establecido (alto, medio-alto, medio-bajo y bajo), a partir de un cuestionario a los alumnos sobre la profesión de los padres, sus estudios, su situación laboral y algunos otros indicadores relacionados. En consecuencia, la información que se ofrece a cada uno de los centros es *contextualizada*, ya que se le presentan también datos de los otros centros que participan al mismo tiempo en esta evaluación. Por una parte tienen conocimiento de las puntuaciones medias de la totalidad de los centros, tanto a nivel de la comunidad autónoma como del Estado y, por la otra, las puntuaciones medias de los centros que tienen características similares en cuanto al contexto sociocultural en el que han sido clasificados. Esto permite a los centros tener un conjunto de referencias para llegar a conclusiones más significativas. En las figuras 1 y 2 que se presentan más adelante pueden verse sendos ejemplos de gráficas presentadas a los centros en los que aparecen las mencionadas comparaciones.

c) Información objetiva

La información de cada escuela se obtiene tras la aplicación de cuestionarios y pruebas estandarizadas, todas ellas previamente analizadas, con lo que se garantiza un alto grado de fiabilidad y validez de las mismas.

d) Información amplia y convergente

Además de ofrecer información sobre el rendimiento de los alumnos en las diferentes áreas curriculares, se evalúan también sus actitudes, satisfacciones, habilidades metacognitivas y estrategias de aprendizaje. De la misma manera, se recogen datos sobre los procesos del centro y de aula, y las valoraciones de los padres, profesores y alumnos. En la Tabla I se señalan los principales niveles y dimensiones del modelo propuesto.

e) Longitudinalidad

Fernández Díaz y González Galán (1997) señalan las características de los procesos de escuelas eficaces que bien pueden ser asumidas por el modelo de evaluación REDES. En este tipo de evaluación se considera que los procesos de mejora no terminan con la puesta en marcha de la renovación escolar, al contrario, no tienen fin, son procesos continuos que implican ajustes sucesivos para cundir un objetivo de mejora permanente, en constante revisión y, si se da el caso, con nuevos objetivos a alcanzar. Los resultados no suelen ser inmediatos y si lo fueran habría que evaluar la estabilidad del cambio; por ello, muchos autores abogan por estudios longitudinales a través de varios años (entre tres y cinco) para que los resultados obtenidos, tanto positivos como negativos, tengan cierta consistencia y estabilidad. El modelo de evaluación REDES es un proceso que abarca cuatro años de manera longitudinal. Ello permitirá ofrecer información a los centros acerca de su propia evolución, a través de una comparación de su avance entre un año y otro. Así por ejemplo, en el caso del alumnado que en el curso 1999-2000 se encontraba en 1º de educación secundaria, va a ser posible darle seguimiento total a toda su escolarización en dicho nivel educativo, ya que se conocerán sus puntuaciones al comienzo de 1º (nivel inicial), al final de 2º (nivel final/principio de ciclo) y al final de 4º (nivel final de la educación secundaria). De esta forma, nos estaríamos acercando al concepto de *valor añadido*, dado que permite comparar el rendimiento esperado de los alumnos con el rendimiento observado y, posteriormente, analizar las discrepancias. A este valor añadido Goldstein y Spiegelhalter (1996) prefieren llamarle “comparaciones ajustadas” (*adjusted comparisons*), debido a la influencia que ejerce el contexto en el rendimiento de los alumnos. En las figuras 3, 4, 5 y 6 puede verse de manera gráfica lo que acabamos de señalar, con el caso de un centro en concreto.

f) Confidencialidad de los datos

El papel de la administración educativa, en este caso el Departamento de Educación del gobierno del País Vasco, es además de promotor, el de facilitador y no controlador-supervisor. No se contempla la rendición de cuentas por parte de los centros evaluados y, de hecho, una de las características que lo garantiza es la *confidencialidad de los datos*. La información que reciben los distintos centros es confidencial, solamente puede ser conocida por el propio centro y los miembros

del equipo evaluador externo, de forma que ni siquiera la administración accede a ellos.

g) Evaluación formativa

El objetivo último de esta evaluación es el de trabajar junto a los centros para que se conozcan mejor a sí mismos y puedan diseñar estrategias de cambio. Desde este punto de vista, podemos hablar de una evaluación formativa. Las funciones de inspección y control están muy lejos de este plan de evaluación.

h) Interpretada por los propios centros

La información que se ofrece a los centros es de carácter descriptivo, sin interpretaciones. El análisis y la interpretación de la información deben ser hechas por los propios centros, ya que ellos comprenden mejor el alcance de los resultados obtenidos, porque conocen de manera más precisa la historia de la escuela, su cultura, los problemas acaecidos y las decisiones adoptadas.

i) Capaz de comprender los factores que influyen en la calidad de la enseñanza

En la medida en que la evaluación se aplica a un número representativo de centros educativos, proporciona información más científica de la influencia relativa de los diferentes niveles y dimensiones que constituyen el modelo de evaluación.

Diseño de la evaluación

Tal como señalan Marchesi y Martín (1999a), los modelos aplicados de evaluación pueden ser muy diversos en función de los niveles y dimensiones que se incluyan en ella y de sus características. Los modelos más simples son aquellos que sólo tienen en cuenta una dimensión (Escudero, 1997), normalmente la de los resultados académicos de los alumnos, mientras que los modelos más complejos tienen en cuenta el estudio del contexto sociocultural, los recursos disponibles, los procesos de centro y aula, los resultados académicos de los alumnos y las valoraciones que hacen del centro tanto los profesores, como los alumnos y sus padres. El diseño de la evaluación se basa en el modelo CIPP (*context, input, process, product*), en el que se tiene en cuenta contexto, proceso, entrada y producto. Este modelo es mejor en la práctica que aquellos otros más simples, pero exige más tiempo y un mayor coste económico. En la Figura 1 puede observarse el esquema del diseño de evaluación utilizado en REDES.

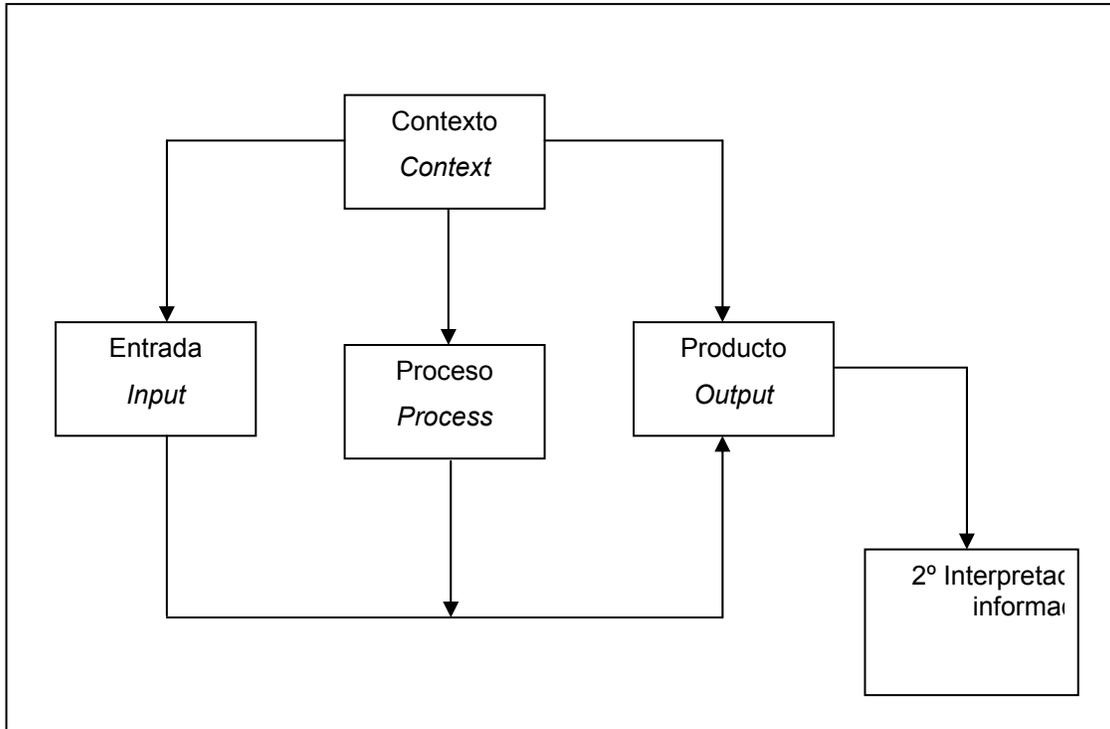


Figura 1. Esquema del diseño de evaluación del modelo REDES (adaptado de Reynolds, Teddlie, Creemers, Scheerens y Townsend, 2000).

Cada uno de los cuatro niveles señalados está constituido por varias dimensiones. En la Tabla I pueden verse las dimensiones seleccionadas en cada uno de los niveles.

Tabla I. Niveles y dimensiones seleccionados en el diseño de evaluación REDES

Niveles	Dimensiones
Entrada	Resultados iniciales de los alumnos
Contexto	Contexto socioeconómico y cultural
Procesos	Procesos del centro escolar Procesos del aula
Producto	Alumnos <ul style="list-style-type: none"> • Áreas curriculares • Actitudes y valores • Estrategias de aprendizaje • Habilidades metacognitivas • Valoración del centro escolar
	Padres <ul style="list-style-type: none"> • Valoración del centro escolar
	Profesores <ul style="list-style-type: none"> • Valoración del centro escolar

En el primer nivel, el de *entrada*, se recogen los resultados preliminares de los alumnos que inician la educación secundaria. Este nivel tiene mucha importancia para conocer los cambios que pueden darse a lo largo del tiempo y para determinar con mayor fiabilidad el valor añadido (*value-added*) del centro escolar. El valor añadido es la medida de la mejora que se da en el rendimiento del alumnado de un centro.

El nivel de *contexto* corresponde al estatus socioeconómico y cultural. Como ya se ha mencionado en diversas ocasiones, todos los estudios destacan la influencia que tiene el contexto socioeconómico y cultural en los resultados obtenidos por los alumnos.

El nivel de *procesos* recoge los procesos específicos del aula y los referidos a la organización y funcionamiento del centro. Algunos de los procesos que se estudian en esta dimensión del centro escolar son los siguientes: conocimiento del funcionamiento del centro, participación, clima de trabajo, valoración del funcionamiento del equipo directivo, funcionamiento de los seminarios o departamentos didácticos, expectativas hacia el centro y los alumnos, y las condiciones de trabajo. En lo que se refiere a procesos de aula se han tenido en cuenta, entre otros, la planificación de la enseñanza, la innovación en la evaluación, la capacidad del profesor de mantener un ambiente de trabajo ordenado, etcétera.

El nivel de *productos* pone énfasis principalmente en los alumnos, pero también recoge las opiniones de profesores y la familia. En lo referente a los alumnos, Marchesi (2002) señala que se defiende un concepto amplio y equilibrado de rendimiento escolar. En consecuencia, en este modelo los alumnos evalúan los aprendizajes de la mayoría de las áreas curriculares (Matemáticas, Ciencias Sociales, Ciencias de la Naturaleza, Tecnología, Lengua española, Lengua propia del país (Vasco) e Inglés; pero también se incluyen las habilidades metacognitivas, sus estrategias de aprendizaje y sus actitudes. También completan un cuestionario sobre el funcionamiento del centro en el que valoran su relación con el tutor, los profesores y el resto de compañeros; las actividades extraescolares, las calificaciones y la preparación recibida, etcétera. Los padres, por su parte, contestan un cuestionario sobre la valoración que hacen del funcionamiento del centro escolar: la atención que reciben, la relación con el tutor, el orden y la disciplina del centro, las calificaciones recibidas por sus hijos, etcétera. Por último, los profesores expresan su grado de satisfacción con el funcionamiento del centro y con las condiciones en las que realizan su trabajo.

Para llevar a cabo la evaluación de manera operativa, ésta se realiza en las siguientes cinco etapas:

- 1) Compromiso del centro. El grupo directivo recoge información acerca de las características del programa de evaluación. Después de que los órganos de

gobierno del centro escolar discuten y analizan el programa se toma la decisión de participar o no en él.

- 2) Recogida de datos.
- 3) Entrevista con el equipo directivo acerca del informe de su centro.
- 4) Interpretación de la información por parte del centro.
- 5) Proyecto de cambio y evaluación de las consecuencias posteriores.

Los responsables de la evaluación son conscientes de que una evaluación externa y cuantitativa, que busca ofrecer información relevante y contextualizada a los centros escolares aplicando las mismas pruebas en todos ellos, permite una mayor objetividad en el análisis de los resultados y les otorga una perspectiva más amplia sobre su situación; aunque este procedimiento tiene sus limitaciones (Marchesi, 2002). Por una parte, las pruebas son más reducidas y no pueden sustituir a la información que cada profesor tiene de sus alumnos y, por otra, el excesivo énfasis en los datos cuantitativos puede disminuir la percepción de la importancia de la evaluación cualitativa e interna del centro educativo. Por ello, se insiste en que esta evaluación externa es sólo una ayuda a la necesaria evaluación al interior de cada centro, recalcando que deben ser ellos mismos los que interpreten la información que se les proporcione. Por otro lado, esta información cuantitativa se complementa con las entrevistas mantenidas con los equipos directivos.

Recogida de la información

Como se ha indicado previamente, en cuanto a la información que se recoge se ha intentado evitar que los resultados solamente queden al nivel del rendimiento en las disciplinas académicas. Por ello, se han incorporado otros aspectos como actitudes, opiniones de alumnos, estrategias de aprendizaje y habilidades metacognitivas. Además, de cara a una evaluación más amplia que integre el mayor número de puntos de vista, se han incluido valoraciones del profesorado y de las familias acerca del funcionamiento del centro y del desarrollo de la enseñanza.

Los instrumentos que se aplican son los siguientes:

- Cuestionarios de: contexto socioeconómico y cultural, actitudes, estrategias de aprendizaje, procesos de centro y aula (profesorado), opinión de los alumnos y opinión de las familias.
- Pruebas de: habilidades metacognitivas, Lengua y Literatura Vasca, Lengua y Literatura Española, Inglés, Matemáticas, Ciencias Sociales, Ciencias Naturales, Tecnología, Geología-Biología y Física-Química.

Los cuestionarios aplicados son en su mayoría escalas de tipo Likert. Las pruebas curriculares son pruebas objetivas, que tratan de mantener un equilibrio entre los contenidos conceptuales y los procedimentales.

Un aspecto que merece destacarse es el de la preparación de las pruebas y los cuestionarios. Los instrumentos de recogida de datos son los mismos para todos los centros del Estado, exceptuando las pruebas de Ciencias Sociales y Lengua de la comunidad, por razones obvias. En una primera fase estos cuestionarios y pruebas son preparados en el IDEA por expertos en construcción de pruebas y profesorado de secundaria que pilotan dichas pruebas y realizan los análisis de ítems y de la prueba, correspondientes, siguiendo modelos de la Teoría clásica de los tests y de la Teoría de la respuesta al ítem (*Item Response Theory* o IRT).

De acuerdo con la clasificación de Jornet y Suárez (1996), podemos definir las pruebas utilizadas como *pruebas estandarizadas de indicadores de resultados*, ya que pretenden traducir los niveles de competencia que, en las distintas disciplinas y materias, una institución, centro o sistema educativo asume como objetivos a cumplir en el proceso educativo.

En la elaboración de las pruebas hay un componente muy importante que es el del *dominio educativo* al cual esperan responder. Se entiende por *dominio educativo* las dimensiones que se pretenden medir dentro de una materia o disciplina. La definición del dominio la realizan expertos en la materia que es objeto de evaluación y se cuenta con el apoyo de especialistas en medición y evaluación. El papel de estos profesionales es fundamental, ya que el trabajo realizado por ellos en la configuración y el desarrollo de las pruebas viene a compensar, de alguna forma, las dificultades surgidas por los débiles constructos teóricos en los que se suelen sustentar y por la falta de comprobación de su unidimensionalidad. Esto se debe a que las propuestas metodológicas específicas para el análisis empírico de los resultados con variables dicotómicas han dado resultados insatisfactorios al identificarse las dimensiones, en muchas ocasiones, con base en la dificultad de los ítems y no en los contenidos (Jornet y Suárez, 1996).

Además de la labor de los expertos en la determinación de la importancia y dificultad teórica de los ítems, otro elemento fundamental es la calidad técnica de las pruebas. Con el fin de garantizar dicha calidad, es importante el análisis psicométrico de la prueba, que asegure su fiabilidad y la de los ítems uno por uno. En este sentido y basándose en los datos ofrecidos por la prueba piloto, el análisis de distractores constituye el trabajo central en este aspecto, ya que va a permitir anticipar y controlar posibles dificultades que puedan surgir en la aplicación definitiva de las pruebas. Junto con esto, el análisis de los posibles sesgos es también un elemento clave para asegurar la validez de las pruebas.

Finalmente se podría concluir que las pruebas estandarizadas de rendimiento que se están aplicando responden a un planteamiento correcto, ya que al estar planteadas para poblaciones extensas y heterogéneas no son interpretadas individualmente y, por otra parte, la selección de indicadores, el análisis de ítems y las comprobaciones de fiabilidad y validez estimamos que son suficientes. A esto deberíamos añadir otro elemento que consideramos muy importante, el referido al propio uso de las pruebas, que pretenden tener una dimensión más formativa que sumativa para tomar decisiones acerca del grupo y no acerca de los individuos.

La información que hasta este momento se ha ofrecido a los centros participantes en la evaluación es de tipo normativo, en el sentido de que cada uno conoce su posición en relación con el resto de los centros participantes.

Con el objeto de completar el carácter normativo de la interpretación de las pruebas se ha planteado también el establecimiento de estándares (*standard*) y puntos de corte (*cut off score*) en las pruebas de rendimiento tal y como se hace en las pruebas de referencia criterial (*criterion-referenced test*), de tal manera que con los estándares establecidos podemos determinar el nivel a partir del cual podemos clasificar al alumnado en *aptos* y *no aptos* en la variable de rendimiento considerada.

En este sentido y teniendo en cuenta la importancia que tiene la variable *contexto sociocultural de las familias*, actualmente se está trabajando con el objetivo de poder ofrecer estándares diferenciados para sus distintos niveles. El proceso seguido en el establecimiento de los mencionados estándares y la influencia del contexto sociocultural en los mismos, se especifican en forma detallada en un artículo de Lizasoain *et al.* (en prensa) que próximamente se publicará.

En la recogida de datos se dan los siguientes pasos:

Las pruebas y cuestionarios referentes a 1º de Educación Secundaria se aplican en octubre, es decir, al inicio de la escolaridad en esta etapa. Junto con éstos, se recogen datos para determinar el contexto sociocultural del centro a través de un cuestionario que se aplica a todos los alumnos y alumnas de 1º, 2º y 4º de educación secundaria. La aplicación de estas pruebas al comienzo del curso tiene como objetivo realizar una evaluación inicial del alumnado, que será válida tanto para el profesorado de ese curso (en el sentido de que con estos primeros datos podrán adaptar las programaciones a la situación real) y para el propio centro, de tal forma que en un periodo de cuatro cursos escolares se podrán apreciar los avances de esos alumnos.

Previo a la recogida de datos se realiza una reunión en cada uno de los centros escolares, en la que participan la dirección del centro, el profesorado de educación secundaria y el equipo evaluador. En ella se informa al centro escolar todo lo relacionado con el plan de evaluación que se va a aplicar: objetivos, características, tipo de pruebas y cuestionarios, fechas y fases de la recogida de datos, etcétera. Es en este momento cuando pueden ser negociados diferentes aspectos de la evaluación; por lo que la reunión es muy importante, ya que se sientan las bases del clima de colaboración (entre los centros y en este caso los evaluadores), necesario para que se cumplan los objetivos planteados.

Durante los meses de enero y febrero el alumnado de 2º y 4º cumplimenta los cuestionarios de actitudes y opinión, el profesorado el cuestionario de procesos del centro y del aula, y las familias los cuestionarios sobre su valoración del

centro. Por último, en mayo se aplican las pruebas de rendimiento al alumnado de 2º y 4º de la educación secundaria.

Difusión y expansión de la información

Para la difusión de la información se utilizan los informes de evaluación, cuya elaboración representa la concreción de la tarea básica de la evaluación, esto es, transformar los datos en evidencias (Sáez, 1995). Estos informes elaborados en el modelo REDES tratan de seguir las siguientes directrices:

- Incluir toda la información que necesitan las distintas audiencias.
- Intentar ser pertinentes y relevantes.
- Devolverlos a la mayor brevedad posible (en el mismo trimestre en el que se aplican las pruebas).
- Intentar ser breves y concisos.
- Utilizar un lenguaje accesible.
- Utilizar recursos diversos de comunicación como tablas, gráficos, etcétera.
- No incluir recomendaciones, puesto que se considera que es una tarea que corresponde a los destinatarios del informe.
- Seguir un carácter expositivo.

Para ello, a lo largo del curso escolar se realizan dos reuniones del evaluador con la dirección del centro, una en noviembre y otra en junio. En ellas se hace entrega de los informes correspondientes a los resultados obtenidos.

Cada año se entregan dos informes a cada centro escolar: uno en noviembre con los resultados del alumnado de 1º y otro en junio con los resultados de las pruebas de 2º y 4º aplicadas por el profesorado; valoraciones realizadas por el alumnado y padres o madres. Son informes sencillos y esencialmente descriptivos de los datos presentados, sin ningún tipo de valoración. Las posibles interpretaciones se dejan para la reunión que se celebra con los responsables correspondientes. La estructura de los informes es la siguiente:

- Explicación de los datos acerca del contexto sociocultural en el que ha quedado enmarcado el centro.
- Puntuaciones individuales de cada uno de los alumnos en las distintas áreas evaluadas. Además de las puntuaciones totales, se recogen las subpuntuaciones en cada una de las subáreas.
- Puntuaciones acerca de las actitudes de los alumnos y de las valoraciones acerca del centro hechas por alumnos, profesores y familiares, tanto totales como sectoriales.
- Puntuaciones medias e histograma correspondiente a las puntuaciones medias en cada una de las áreas y a cada uno de los grupos-clase de cada nivel (1º, 2º y 4º).

- Tabla con las puntuaciones del centro en cada una de las áreas y subáreas por grupos-clase, junto con las puntuaciones medias de los centros clasificados en el mismo contexto sociocultural.
- Polígono de frecuencias en el que de forma gráfica y numérica se reflejan las puntuaciones medias del centro escolar específico, de los centros del mismo contexto sociocultural dentro de la comunidad autónoma, del total de los de la comunidad, de los centros del mismo contexto sociocultural dentro del Estado Español y del total que participan en la evaluación dentro del Estado.
- Descripción y esquema de cada una de las pruebas para ser utilizados en la interpretación de los resultados.
- Explicación del informe y de los resultados más importantes.

Como ya hemos señalado anteriormente, son los propios centros los que deben hacer las interpretaciones oportunas y valoraciones con base en sus propios criterios. Por ello uno de los objetivos más importantes de las reuniones que se celebran para la entrega y presentación de los informes es el de orientarles en esa interpretación y ayudarles a establecer proyectos de cambio y mejora.

A continuación se presentan ejemplos de dos gráficos que aparecen en los informes (ver Figuras 2 y 3). El primero muestra resultados obtenidos en el cuestionario “Procesos del centro” cumplimentado por el profesorado, comparado con los resultados de los centros de su mismo contexto y con los de la totalidad de los centros, tanto del País Vasco (PV) como del Estado Español.

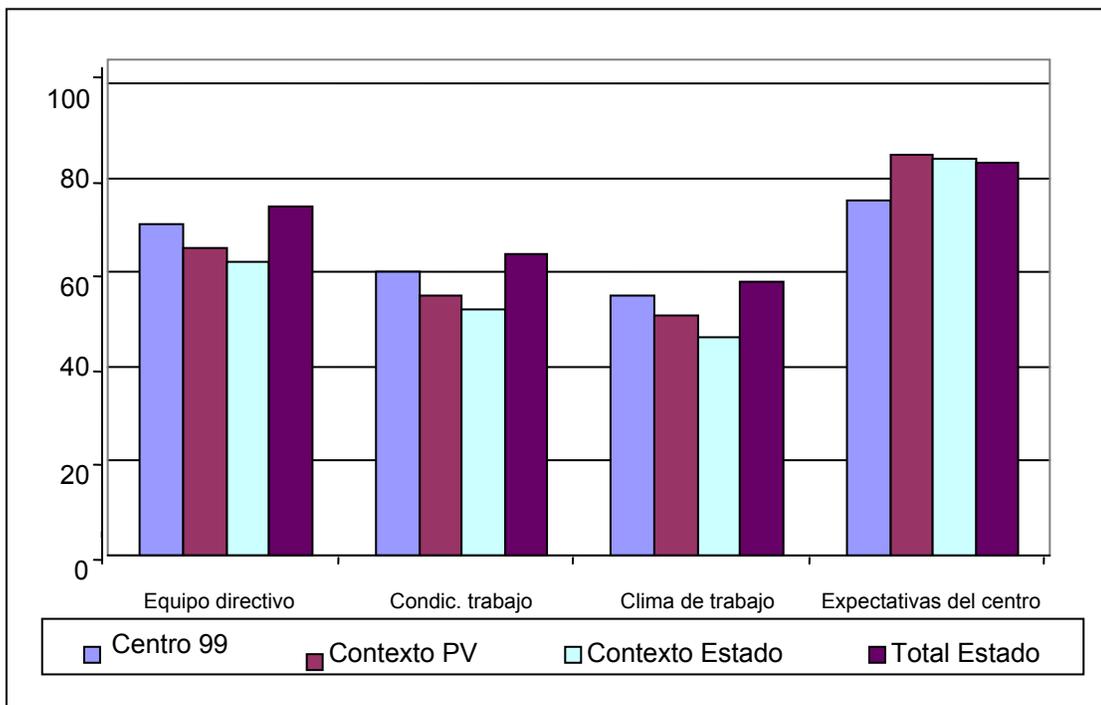


Figura 2. Resultados obtenidos por el centro en el cuestionario “Procesos del centro” (datos ficticios)

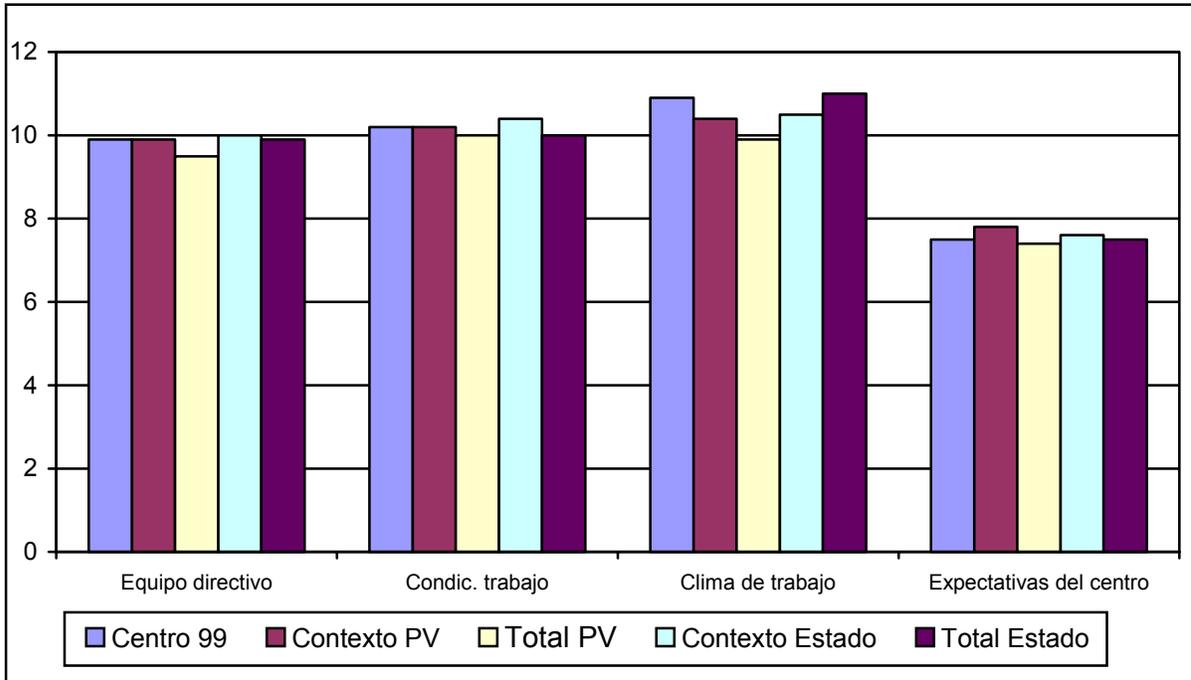


Figura 3. Resultados obtenidos por el centro en la prueba de Ciencias Sociales de 1º de enseñanza secundaria obligatoria (datos ficticios)

La Figura 3 muestra los resultados obtenidos por el centro en la prueba de Ciencias Sociales de para estudiantes de 1º de ESO, comparados con los resultados obtenidos en los centros de su mismo contexto y con los de la totalidad de los centros, tanto del País Vasco (PV) como del Estado Español.

No obstante, cuando un centro completa los cuatro años del programa de evaluación recibe un informe final, en el que se describe lo sucedido durante ese periodo. Básicamente, la estructura del informe es similar a la del resto de informes parciales, pero se ofrecen una serie de gráficas que resumen los resultados obtenidos durante esos cuatro cursos escolares en 1º, 2º y 4º de la ESO (Figuras 4, 5 y 6) en cada una de las dimensiones evaluadas. Además, se incluye un último gráfico en el que aparece el itinerario seguido por los alumnos en que en el primer año de aplicación del programa de evaluación se encontraban en 1º de la ESO y al final del programa de evaluación terminaron el 4º (ver Figura 7). Mediante dicho gráfico se puede tener una idea, aunque no del todo precisa, del valor añadido del centro escolar en cuestión.

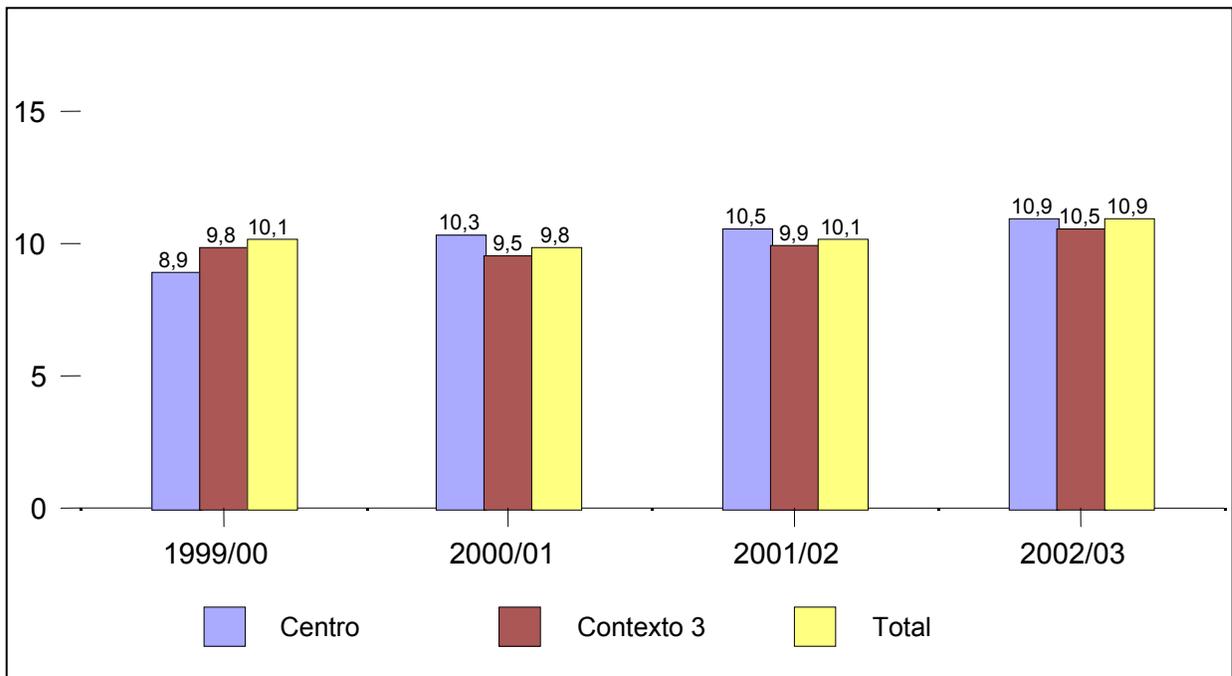


Figura 4: Matemáticas 1º de la ESO

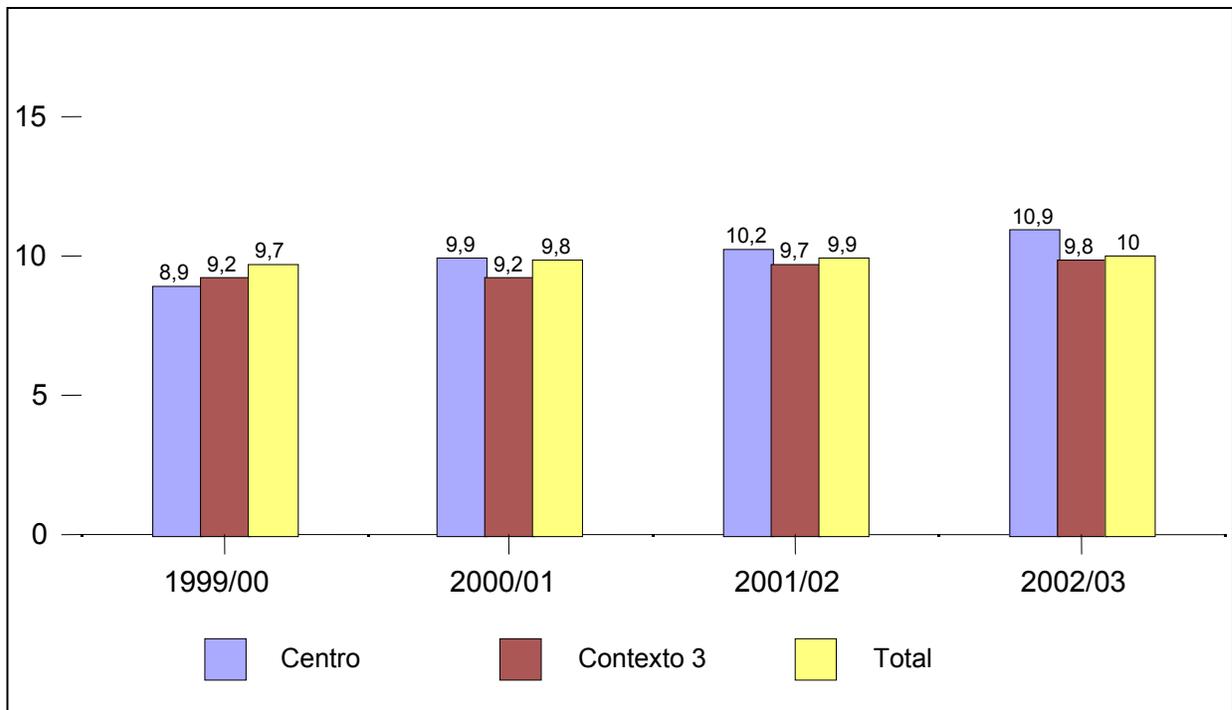


Figura 5. Matemáticas 2º de la ESO

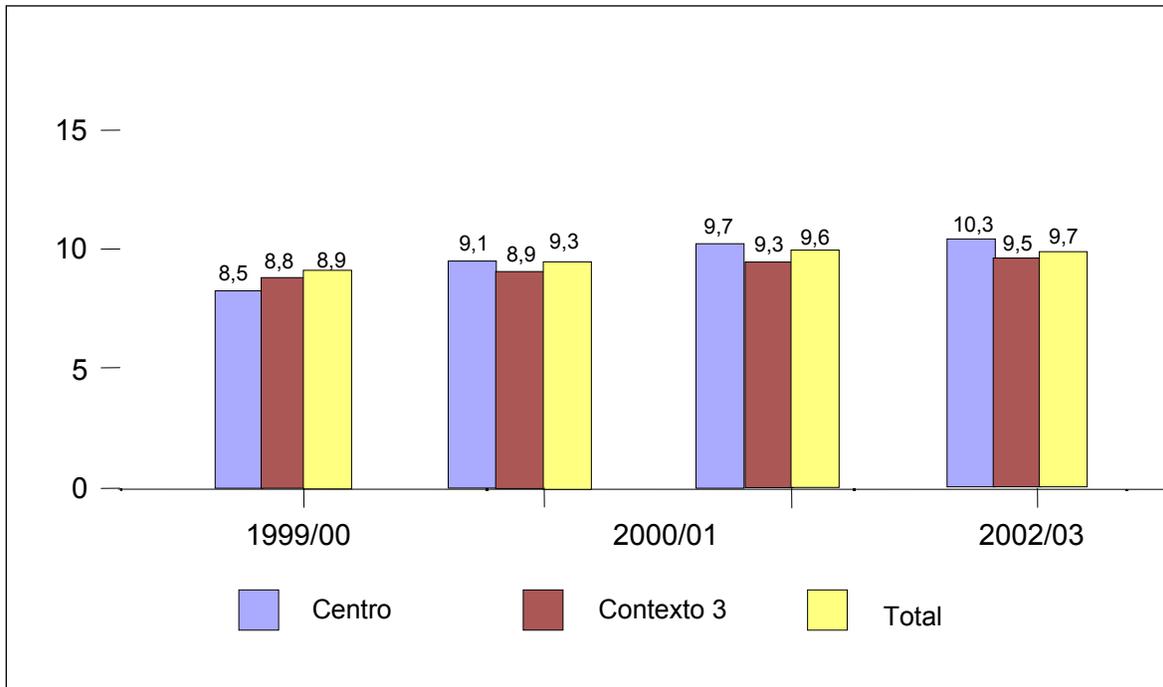


Figura 6. Matemáticas 4º de la ESO

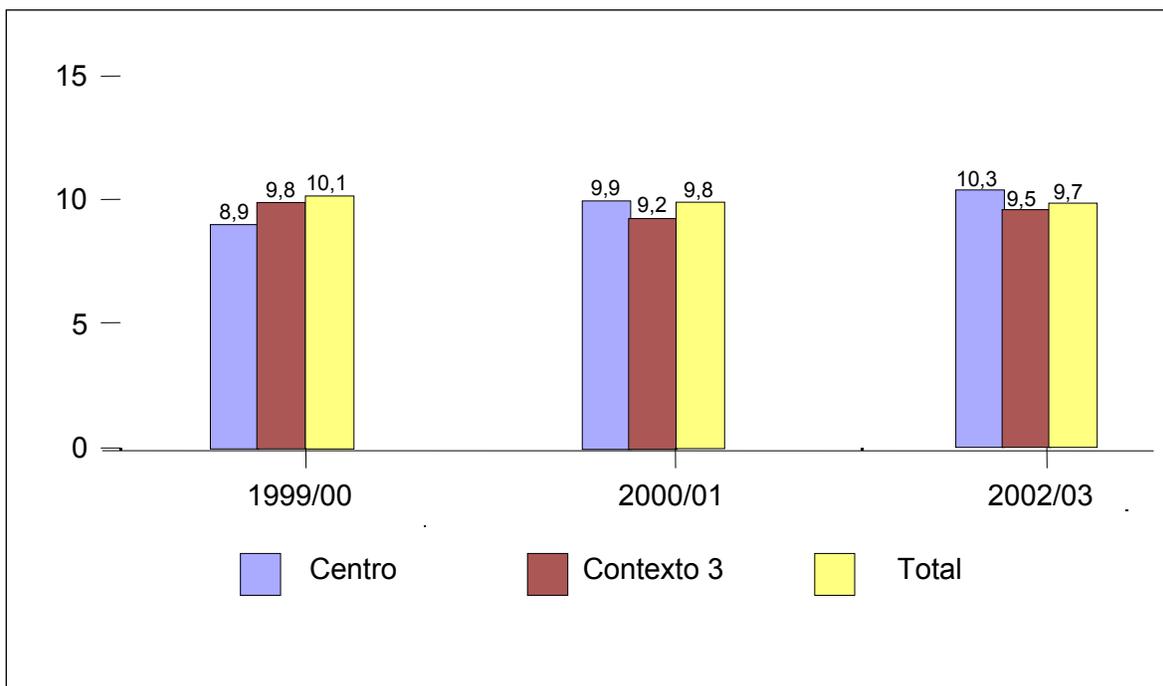


Figura 7. Desarrollo de las matemáticas (el mismo grupo de alumnos en 1º, 2º y 4º)

Toma de decisiones y planes de mejora

Hoy en día nadie duda de que una de las condiciones que ha de cumplir una evaluación es que ofrezca información y elementos de juicio a quienes deben tomar decisiones. Es decir, se da prioridad al elemento utilitarista de la evaluación. Como señala Tiana (1997) se supone que una evaluación debe proporcionar información rigurosa, válida y fiable a los encargados de un centro educativo o programa, a partir de la cual tomarán las decisiones oportunas. No obstante, esta concepción es tachada de ingenua debido a la escasa incidencia de la evaluación en los procesos de toma de decisión. De hecho, hoy en día se considera que la relación entre ambas prácticas es mucho más compleja. El proceso de toma de decisiones se apoya en elementos variados y de diverso orden, como los que Tiana (1997) señala:

- Información directa o indirecta procedente de diversas fuentes y con mayor o menor pertinencia y fiabilidad en cada caso.
- Condicionamientos de carácter normativo, institucional o personal que determinan el margen de maniobra existente.
- Información sobre el contexto de relaciones interpersonales, anticipando las reacciones que puedan producirse.
- Negociaciones multilaterales para que todos los actores implicados se sientan copartícipes de la decisión adoptada.

Es decir, nos movemos en una realidad en la que la información y el conocimiento son necesarios, pero no suficientes, dado que también la capacidad de negociar y de movilizar voluntades, el sentido de la oportunidad e incluso la intuición son muy importantes. En aras de facilitar esta labor, el modelo de evaluación REDES propone una serie de *orientaciones* para hacer de la evaluación un instrumento de mejora. A continuación se resumen.

La evaluación es sólo un primer paso, una fotografía inicial que desvela el conjunto y cada una de las partes con cierta precisión. Muestra los contornos firmes y borrosos. Pero esta fotografía es incapaz de hablar por sí misma, es necesario saber verla, comprenderla e interpretarla, lo que exige la participación de la dirección y del profesorado del centro. Por ello, como se ha mencionado en diversas ocasiones, la confluencia de la evaluación interna y externa es una garantía para el mejor conocimiento de la realidad educativa del centro.

Ahora bien, este primer paso no es el último. Para que la evaluación contribuya a generar procesos de cambio y mejora es necesario que demos al menos otros dos pasos. El segundo es interpretar y utilizar de forma adecuada la información obtenida. El tercero consiste en promover proyectos o iniciativas de cambio. Un último paso que cierra el círculo es volver a evaluarse para comprobar la dinámica que se está produciendo (ver Figura 8).

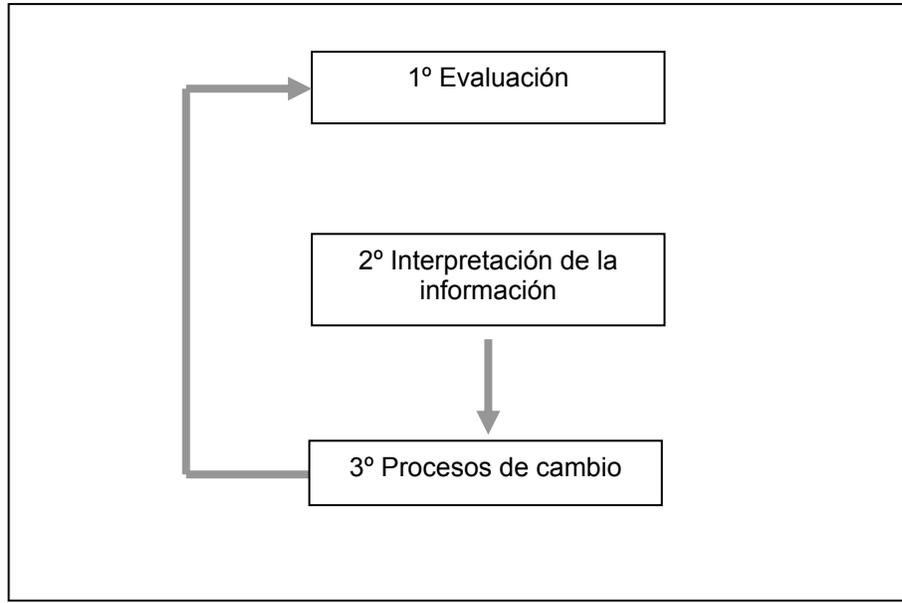


Figura 8. Proceso de evaluación y cambio

Garantizar que la información obtenida a partir de la evaluación se difunda entre la mayoría de las personas implicadas en el funcionamiento de la escuela va a posibilitar favorecer los procesos de cambio. Siempre y cuando, claro está, las personas que reciben la información sean capaces de interpretarla correctamente en beneficio del trabajo de los profesores y de la propia escuela.

El buen uso de la información es ya una parte importante del proceso de cambio. La reflexión compartida entre los diferentes miembros de la comunidad educativa, especialmente los equipos directivos y de profesores ayudará a tomar conciencia de la situación y a modificar actitudes, estilos de enseñanza y formas de trabajo. De la misma manera, los padres y los alumnos modificarán sus actitudes y expectativas sobre la escuela cuando se sienten partícipes de un proceso de evaluación.

Sin embargo, aunque se reconoce su indudable interés y ventaja, este análisis colectivo debe dar un paso más y conducir a reforzar los aspectos positivos del centro y a resolver los problemas diagnosticados. Tal y como se señala en el párrafo anterior, es labor del equipo directivo determinar los posibles proyectos de mejora de la escuela y establecer los objetivos, los plazos, las condiciones, las responsabilidades y las acciones para llevarlos a cabo. Ahora bien, nunca se debe olvidar que los cambios en educación son lentos y laboriosos, y hay muchas condiciones que deben tenerse en cuenta (la cultura de la escuela, las actitudes de los profesores, el tiempo disponible, las relaciones humanas, las expectativas profesionales, etc.), ya que de ellas dependerá que el plan de acción tenga éxito finalmente y se incremente la satisfacción y la moral de la comunidad educativa escolar.

Recomendaciones

Para concluir, se presenta una serie de sugerencias que consideramos pueden ayudar a mejorar el modelo de evaluación REDES.

1. Si el modelo de evaluación REDES prestara únicamente atención a los productos, estaría aplicando un modelo economicista que desvirtuaría incluso el propio concepto de educación. Tener en cuenta además de los productos los antecedentes, sería utilizar un enfoque eficaz en cuanto a medición de los rendimientos del sistema, pero limitado porque la detección de los problemas no siempre nos señala su ubicación ni la forma de resolverlos. Por ello, cualquier modelo de evaluación como el de REDES debe dedicarse también al *proceso*, ya que esto permitirá obtener un enfoque más educativo y completo para la mejora real del centro escolar. Por ello, aunque el modelo REDES contempla también el proceso, es conveniente que tenga un peso específico mayor.

2. La mayoría de los centros está de acuerdo con este programa de evaluación. En ese sentido, puede consultarse el artículo de Lukas *et al.* (2000), en el que se presenta más información acerca de la metaevaluación realizada. Puede observarse, entre otros aspectos, que la mayoría de los alumnos quiere seguir, pero se encuentran en dificultades para explotar la información a favor de instaurar procesos de mejora. Los evaluadores externos deben emplear parte de sus energías en ayudar a los centros en la difusión y diseminación de la información obtenida, en su interpretación y en la instauración de procesos de cambio y mejora.

3. Existe la necesidad de complementar la evaluación externa con la autoevaluación del centro que en la mayoría de las ocasiones no se realiza; en muchos casos debido a la falta de preparación del profesorado. Por ello, como señala Escudero (1997), las administraciones educativas deberían potenciar la formación de evaluadores, sobre todo, entre los profesores de los centros educativos, dado que no se puede seguir demandando y construyendo modelos que partan de la participación del personal de los centros, si los profesores no tienen la debida preparación técnica para ello.

4. La complementariedad entre la evaluación externa e interna puede realizarse de la siguiente manera (Mateo, 2000, p. 190): a) La evaluación externa debe partir de la documentación generada por la interna; b) en la evaluación interna debe participar algún agente externo; c) en la externa, alguno interno; d) los dos informes, internos y externos deben fusionarse en uno final aceptado por las partes; e) en las propuestas de mejora deben participar en igualdad de condiciones agentes externos e internos.

5. La evaluación del centro debe considerar la contextualización para poder adaptarlo a sus objetivos, entorno, posibilidades, etcétera, con el fin de buscar su mejora como institución individual. No obstante, programas estandarizados como

el de REDES dificultan en ocasiones la obtención de beneficios relevantes para cada centro, de acuerdo con su situación y sus necesidades particulares. De ahí que una vez más subrayemos la necesidad de complementar la evaluación externa que ofrece REDES con la autoevaluación del centro.

6. Aunque en el modelo REDES a lo largo de cuatro años se entregan ocho informes a cada centro y se hace un seguimiento de la difusión de los mismos, conviene proponer un cambio de perspectiva en donde el término *difusión* dé paso al de *diseminación*, es decir, intentar transmitir “el mensaje de que la evaluación debe depositar su germen de mejora en aquellos terrenos mejor dispuestos para hacerlo fructificar” (Tiana, 1997, p. 34). Para ello, es necesario adoptar una actitud activa y utilizar una diversidad de recursos comunicativos que garanticen la difusión o diseminación de la información en las distintas audiencias implicadas en el centro educativo.

7. Hasta ahora las comparaciones de los centros escolares han sido de tipo normativo. Es decir, cada escuela se compara en las distintas dimensiones evaluadas con el resto de los centros de su contexto o de su comunidad. Esta información, aunque es interesante, no es suficiente. Es cierto que al cabo de los cuatro años de evaluación el centro escolar recibe información adicional sobre el progreso que han tenido los alumnos que entraron en 1º de la ESO hasta llegar a 4º, como un intento aproximativo de ver cuál ha sido el valor añadido del centro. Es decir, se trata de una comparación *idiográfica* en la que se contrasta a los alumnos consigo mismos. No obstante, se debe profundizar en este aspecto utilizando la Teoría de respuesta a los ítems para conseguir modelos que señalen el valor añadido de cada escuela de manera más clara. También puede experimentarse con los Modelos de curva de crecimiento latente, aplicaciones estadísticas emergentes que se presentaron en el simposium *Modelos estadísticos para el estudio del cambio* en el XI Congreso Nacional de Modelos de Investigación Educativa organizado por la Asociación Interuniversitaria de Investigación Pedagógica (AIDIPE) en colaboración con la Universidad de Granada (España). Por último, para completar las comparaciones conviene establecer estándares y puntos de corte para cada una de las pruebas curriculares utilizando medios técnicos como los que habitualmente se emplean en la confección de los tests criteriosales. De esta forma, un centro escolar puede tener tres criterios de comparación a partir de los cuales la interpretación de la información sea más justa y más apegada a la realidad.

Referencias

Coleman, J. S., Campbell, E., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D. et al. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

De Miguel, M. (1997). La evaluación de los centros educativos. Una aproximación a un enfoque sistémico. *Revista de Investigación Educativa*, 15 (2), 145-178.

Escudero, T. (1997). Enfoques modélicos y estrategias en la evaluación de centros educativos. *Revista Electrónica de Investigación y Evaluación Educativa*, 3 (1). Consultado el 10 de febrero de 2004 en:
http://www.uv.es/RELIEVE/v3n1/RELIEVEv3n1_1.htm

Fernández Díaz, M. J. y González Galán, A. (1997). Desarrollo y situación actual de los estudios de eficacia escolar. *Revista Electrónica de Investigación y Evaluación Educativa*, 3 (1). Consultado el 15 de enero de 2004 en:
http://www.uv.es/RELIEVE/v3n1/RELIEVEv3n1_3.htm

Goldstein, H. y Spiegelhalter, D. J. (1996). League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistics Society*, 159, 385-443.

House, E. R. (1994). *Evaluación. Ética y poder*. Madrid: Morata.

Jencks, C. S., Smith, M., Acland, H., Bane, M. J., Cohen, D., Ginits *et al.* (1972). *Inequality: A reassessment of the effect of family and schooling in America*. Nueva York: Basic Books.

Jornet, J. M. y Suárez, J. M. (1996). Pruebas estandarizadas y evaluación del rendimiento: usos y características métricas. *Revista de Investigación Educativa*, 14 (2), 141-163.

Lizasoain, L., Joaristi, L., Santiago, C., Lukas, J. F., Moyano, N. y Sedano, M (en prensa). Establecimiento de estándares y puntos de corte en pruebas de rendimiento. Una aplicación en el área de matemáticas de 2º y 4º de ESO. *Bordón*, 56 (2).

Lukas, J. F., Santiago, C., Munarriz, B., Moyano, N. y Sedano, M. (2000) Euskal Autonomia Erkidegoko ikastetxeen kanpo-ebaluazioa. *Tantak*, 24, 39-63.

Marchesi, A. (2002). El proyecto REDES. En A. Marchesi y E. Martín (Comps.), *Evaluación de la educación secundaria. Fotografía de una etapa polémica* (pp. 21-33). Madrid: Fundación Santa María.

Marchesi, A. y Martín, E. (1999a). *Tendencias actuales en la evaluación de los centros docentes. Infancia y Aprendizaje*, 85, 5-18.

Marchesi, A. y Martín, E. (Coords.). (1999b). Una red de evaluación de centros de educación secundaria. *Infancia y Aprendizaje*, 85, 59-73.

Marchesi, A. y Martín, E. (Comps.). (2002). *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid: Fundación Santa María.

Mateo, J. (2000). *La evaluación educativa, su práctica y otras metáforas*. Barcelona: ICE-Horsori.

Pérez Juste, R. (1999). Nuevas perspectivas de gestión, control y evaluación. *Bordón*, 51 (4), 449-477.

Reynolds, D., Teddlie, C., Creemers, B. P. M., Scheerens, J. y Townsend, T. (2000). An introduction to school effectiveness research. En C. Teddlie y D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 3-25). Londres: Falmer Press.

Sáez, M. J. (1995). Caracterización y problemática de los informes de evaluación. En M. J. Sáez (Ed.), *Conceptualizando la evaluación en España* (pp. 83-93). Madrid: Servicio de Publicaciones de la Universidad de Alcalá.

Tiana, A. (1997). *Tratamiento y usos de la información en evaluación*. Consultado el 20 de enero de 2004 en el sitio Web de la Organización de Estados Iberoamericanos para la Educación de la Ciencia y la Cultura en: <http://www.campus-oei.org/calidad/tiana.htm>

Tiana, A. (1999). Ikastetxeen kanpo-ebaluazioa Aukerak eta mugak. *Tantak*, 21, 7-32.

Webster, W. J., Mendro, R. L. y Almaguer, T. O. (1994). Effectiveness indices: A "value added" approach to measuring school effect. *Studies in Educational Evaluation*, 20, 113-145.