

Los retos de la evaluación de sistemas educativos

Josu Sierra Orrantia

Director del Instituto Vasco de Evaluación e Investigación educativa (ISEI-IVEI).

Resumen:

Tras el traspaso de competencias educativas, algunas administraciones autonómicas han creado sus propios organismos de evaluación. El Instituto Vasco de Evaluación e Investigación Educativa (ISEI-IVEI) centra su atención en la obtención de datos fiables y comparables, más allá de las medias estatales, para la construcción de indicadores que sirvan para promover la mejora del sistema educativo vasco. Se apuesta por la evaluación internacional, y especialmente por PISA, como una de las fuentes de orientación de la política educativa. El País Vasco ha ampliado la muestra en el ciclo de PISA 2003 y también lo hace en el actual ciclo PISA 2006. Se analizan los resultados obtenidos por los alumnos vascos en PISA 2003. Además, se plantea el riesgo de evaluaciones redundantes y la necesidad de pensar en soluciones más completas para la mejora educativa, como los sistemas de rendición de cuentas.

Palabras clave: evaluación comparada, PISA, diagnóstico, sistema educativo, País Vasco, competencias clave, rendición de cuentas.

Abstract: *The challenge of the evaluation of education systems*

After the transfer of the educational administration, some Autonomous Governments have created their own evaluation bodies. The Basque Institute of Evaluation and Educational Research (ISEI-IVEI) focuses its attention in the procurement of reliable and comparable data, further away of the State measures, for the construction of indicators that can be used to promote the improvement of the Basque Education System. They believe in an international assessment, and particularly in PISA, as one of the guidance sources of education policies. The Basque country extended its sample in PISA 2003, and does it, as well, in the current PISA 2006. The results obtained by the Basque pupils in PISA 2003 are analysed. It is considered, as well, the risk of superfluous evaluations, and the need of thinking about more comprehensive answers for educational improvement, as accountability systems are.

Key words: comparative assessment, PISA, diagnosis, education system, key competencies, accountability.

INTRODUCCIÓN

Es necesario enmarcar los retos de la Evaluación de Sistemas Educativos dentro de los cinco retos de los Sistemas Educativos Europeos definidos por la Comisión Europea en el 2000 (European Commission, 2000).

- EL CONOCIMIENTO: La explosión de la información demanda una redefinición del concepto tradicional de conocimiento. Matemáticas, lectura y ciencias son competencias clave como instrumentos esenciales de conocimiento y como base para el aprendizaje permanente.
- LA DESCENTRALIZACIÓN: Durante los últimos años los Sistemas Educativos han devuelto más autonomía y responsabilidad a los centros. Incluso las evaluaciones internacionales (OCDE, *Kiel Institute of World Economics*) avalan la bondad de esta política.
- LOS RECURSOS: La educación es vista cada vez más como una inversión. Crece la demanda en la educación preescolar y aumentan los usuarios de la educación superior y *lifelong learning*.
- LA INCLUSIÓN SOCIAL: Todos los sistemas educativos europeos tienen como objetivo ofrecer a los jóvenes igualdad de oportunidades y prepararlos para la vida adulta y para sobrevivir en una economía globalizada.
- LOS DATOS Y LA COMPARABILIDAD: El *benchmark*. Indicadores que nos permitan compararnos, utilizados de modo diagnóstico y formativo, para informar la política educativa y conocer las tendencias. El reto de la comparabilidad es crear un clima abierto y positivo de diálogo, viendo los indicadores como un punto de partida.

Quiero destacar, sobre todo, el quinto reto, ya que ésta es, hoy, la clave para la mejora en el Sistema Educativo de Euskadi.

EL PUNTO DE PARTIDA: UN SISTEMA CENTRALIZADO QUE DESCONOCE BÁSICAMENTE SUS RESULTADOS

El Sistema Educativo vasco ha sido hasta ahora mismo un sistema preocupado por el *input* (el currículo) normativo y centralizado. Todos los centros deben seguir los decretos curriculares y la gestión de los recursos corresponde a la administración.

No obstante, se desconocían los resultados del sistema y si todos los centros siguen o no los currículos, en qué medida, etc. Las decisiones de política educativa se tomaban intuitivamente o basándose en opiniones.

En 2001, cuando se creó el Instituto Vasco de Evaluación e Investigación Educativa, solamente teníamos datos de resultados de nuestro sistema a raíz de las ampliaciones de muestra efectuadas en evaluaciones estatales coordinadas por el Instituto Nacional de Evaluación y Calidad del Sistema Educativo (INECSE):

- Evaluación de la Educación Primaria 1995
- Evaluación del Diagnóstico del Sistema Educativo 1997

A partir de 2001, el Departamento de Educación, Universidades e Investigación, con la consejera Anjeles Iztueta a la cabeza, apostó de un modo decidido por la evaluación y la investigación como nunca antes se había hecho.

Los objetivos estratégicos eran la obtención de datos fiables y comparables del Sistema Educativo vasco para la construcción de indicadores que pudieran ser valiosos como base para la toma de decisiones y la diseminación de los resultados entre los agentes del sistema con el fin de crear una conciencia evaluativa y promover la mejora.

LA APORTACIÓN DE LAS EVALUACIONES ESTATALES

Su aportación era absolutamente insuficiente para conocer nuestros resultados y construir indicadores. Solamente nos proporcionan la media estatal, la propia y la tendencia en años posteriores. Además, sólo miden el «currículo común», lo que obliga a hacer pruebas complementarias.

En realidad, tras la filtración de los resultados por comunidades de la Evaluación de Primaria 1995, se produjo un uso malévolo de la evaluación, utilizándola para vapulear en la prensa a determinadas comunidades. Tras este incidente, el Consejo Rector del INECSE decidió no hacer muestras representativas de las comunidades en sus estudios.

Los informes publicados por el INECSE sólo hablan de resultados en el Estado y no están orientados a la política educativa, que además difiere según la comunidad.

Hoy día, se da la paradoja de que muchas comunidades no miden sus resultados con estas pruebas y el MEC sí, a nivel estatal.

Las pruebas estatales solamente evalúan el currículo «común» basado en un acuerdo entre las comunidades que aportaron los ítems y el INECSE.

Todo lo anteriormente dicho nos acabó de convencer de que nuestras evaluaciones, hoy por hoy, deben basarse en pruebas internacionales.

NUESTRA APUESTA POR LA EVALUACIÓN INTERNACIONAL COMO FUENTE DE INDICADORES Y DE INFORMES SOBRE EL SISTEMA, ORIENTADOS A LA POLÍTICA EDUCATIVA

A partir de PISA 2000, la evaluación internacional se ha convertido en uno de los principales motores de la mejora de los Sistemas Educativos. El Simposio PISA de Berlín en 2002 fue la prueba más palpable de ello. He aquí algunos de sus postulados:

- «Impulsar a la vez buenos resultados y equidad no es un sueño. Es posible y así lo hacen algunos países.
- Un papel esencial del Sistema Educativo es compensar las diferencias sociales de origen de los alumnos y dar igualdad de oportunidades a todos.
- Existe un debate enriquecedor en torno a PISA donde el objetivo es mejorar.
- Sabemos que el nivel socioeconómico de la familia es muy importante a la hora de explicar el rendimiento de los alumnos. Pero no es determinante. Su importancia varía de unos países a otros.

- Algunos Sistemas Educativos compensan las diferencias socioeconómicas mientras que otros las multiplican».

También son interesantes algunas de sus conclusiones, publicadas recientemente por la OCDE en «Messages from PISA 2000» (OCDE, 2004, p. 19):

- Los sistemas educativos que tienen éxito han extendido la autonomía de los centros escolares. Además, lo han combinado con la devolución de la responsabilidad a los mismos.
- Los sistemas que tienen éxito están comprometidos en la monitorización del rendimiento de los alumnos y del sistema. Aunque no existe una metodología única, los mejores países monitorizan el rendimiento de sus estudiantes y centros.
- El método usado por un sistema educativo para ayudar a los alumnos con bajo rendimiento es crucial en la mejora de los resultados. Además, las diferencias entre centros tienden a aumentar en los sistemas no comprensivos con temprana segregación de alumnos en itinerarios distintos.

Aunque ya habíamos decidido acudir a PISA 2003, nuestra participación en este Simposio nos acabó de convencer de la bondad de la apuesta por la Evaluación Internacional. Ello nos llevó a participar con mayor convencimiento en PISA y TIMSS 2003.

Necesitábamos indicadores europeos y mundiales, compararnos con aquellos países con los que compartimos experiencias en común y competimos en los mercados.

Las evaluaciones de la OCDE y la IEA nos ofrecían:

- Tecnología de evaluación del más alto nivel
- Pruebas mundialmente aceptadas
- Comparabilidad internacional
- Datos de nuestro sistema para la política educativa
- Indicadores europeos en Competencias Clave

PISA 2003: UN GRAN AVANCE

PISA 2003 nos ha permitido «situarnos», conocer nuestros resultados en Competencias Clave (indicador europeo), ver nuestros puntos débiles y las características del Sistema Educativo vasco.

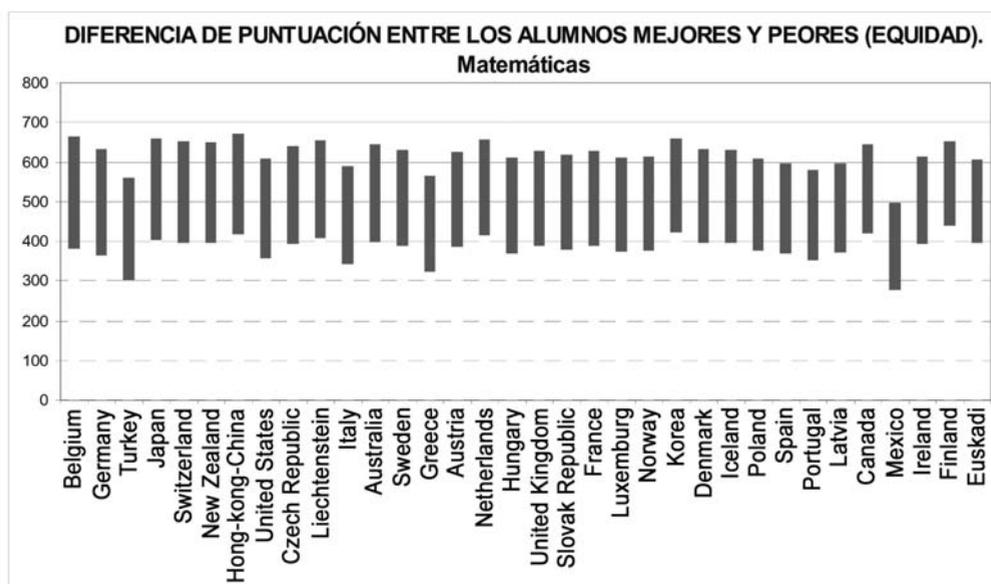
RESULTADOS EN SÍNTESIS:

- De las cuatro competencias clave, nos situamos en la media de la OCDE en tres:
 - Conocimiento Matemático
 - Lectura
 - Resolución de Problemas

- En Conocimiento Científico estamos por debajo de la media de los países de la OCDE. Ésta debe ser nuestra prioridad en los próximos años: mejorar en Conocimiento Científico.
- Nuestro sistema es de una gran equidad: existe muy poca dispersión en el rendimiento de los alumnos y nuestro porcentaje de alumnos de bajo rendimiento es solamente del 16% frente al 23% de España, por ejemplo.
- Tenemos menos alumnos de rendimiento bajo que la OCDE pero nos falta excelencia, es decir, tenemos muy pocos alumnos excelentes en matemáticas.

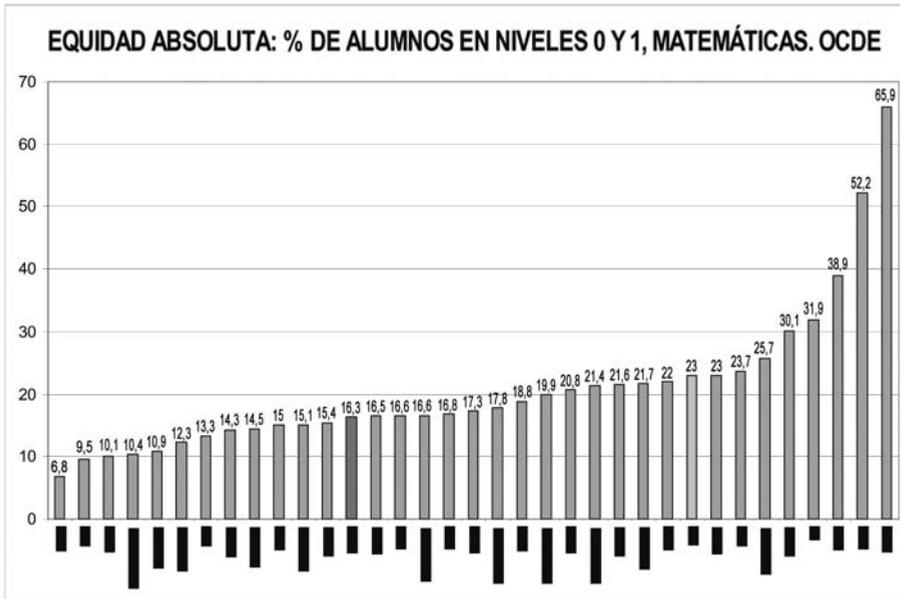
NUESTRA EQUIDAD RELATIVA ES LA MAYOR DE LA OCDE Y LA ABSOLUTA ES TAMBIEN IMPORTANTE

CUADRO I



Euskadi tiene la menor diferencia entre alumnos malos y buenos de toda la OCDE (diferencia entre el percentil 10 y el 90).

CUADRO II

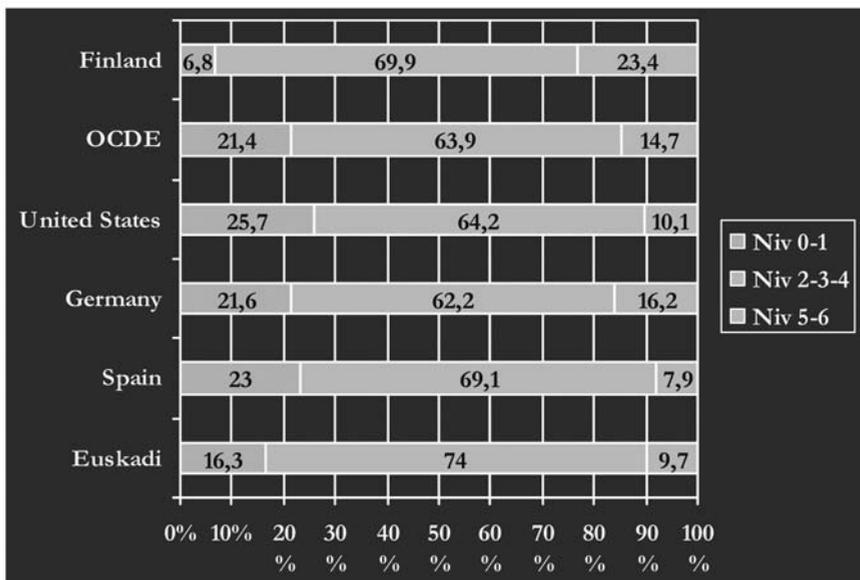


Además, el porcentaje de alumnos en los niveles 1 y menos que 1 de PISA (Equidad Absoluta) es del 16,3%; el 13°, de la OCDE.

El objetivo europeo para el 2010 es bajar este porcentaje al 15,5%. La media de la Unión Europea está en un 19,8% de alumnos en estos niveles 1 y menor de 1.

EUSKADI CUENTA CON UN SISTEMA DE RENDIMIENTO COMPACTO EN MATEMÁTICAS

CUADRO III



Un 74% del alumnado se sitúa en los niveles 2, 3 y 4 de PISA.

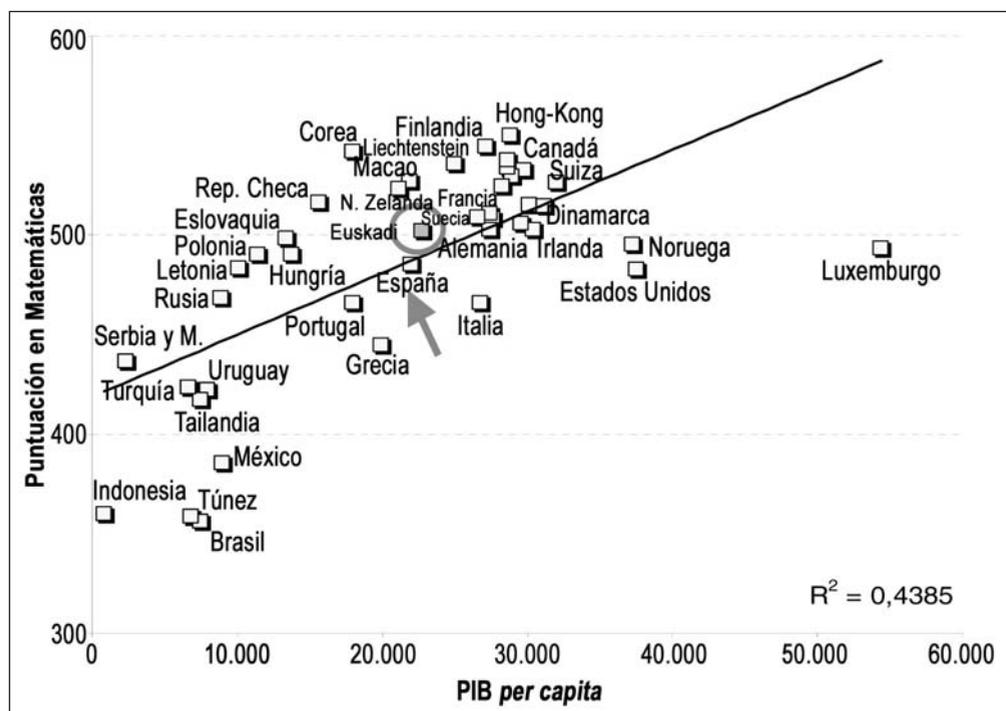
Tenemos menos excelencia que la media de la OCDE pero también menos alumnado de bajo nivel.

RESULTADOS SEGÚN EL NIVEL DE RIQUEZA EXPRESADO EN EL PIB PER CÁPITA

Una de las hipótesis sobre la que se ha venido opinando es que Euskadi obtiene el resultado ya descrito por su nivel de riqueza.

Comparando, en este gráfico realizado por el INECSE, nuestro PIB per cápita con nuestra puntuación en PISA 2003, nos situamos por encima de la recta de regresión, es decir, de nuestro rendimiento esperable. Hay algo más que la riqueza, por tanto.

CUADRO IV



RESULTADOS E ÍNDICE SOCIOECONÓMICO Y LABORAL DEL ALUMNADO

Tampoco parece que nuestro resultado se base en el nivel socioeconómico y laboral de las familias.

Como puede verse en la Tabla I, el índice socioeconómico y laboral de los padres de Euskadi es bastante inferior al de la OCDE, mientras que nuestros resultados se sitúan, en matemáticas, en la media de la OCDE.

TABLA I

PAÍS	ÍNDICE SEL
Canadá	52,6
Finlandia	50,2
Alemania	49,3
OCDE	48,8
Irlanda	48,3
Grecia	46,9
Italia	46,8
Euskadi	45,2
España	44,3

OTROS RESULTADOS IMPORTANTES SOBRE EL SISTEMA EDUCATIVO: SIMILAR RENDIMIENTO EN TODOS LOS CENTROS Y ESCASO EFECTO DEL NIVEL SOCIOECONÓMICO

En Euskadi, la variable Nivel Socioeconómico y Cultural del alumno y (medio) del Centro tiene una baja influencia en los resultados de los alumnos. Su nivel de importancia es equivalente al que tiene en Finlandia o Canadá, países modelo para nosotros.

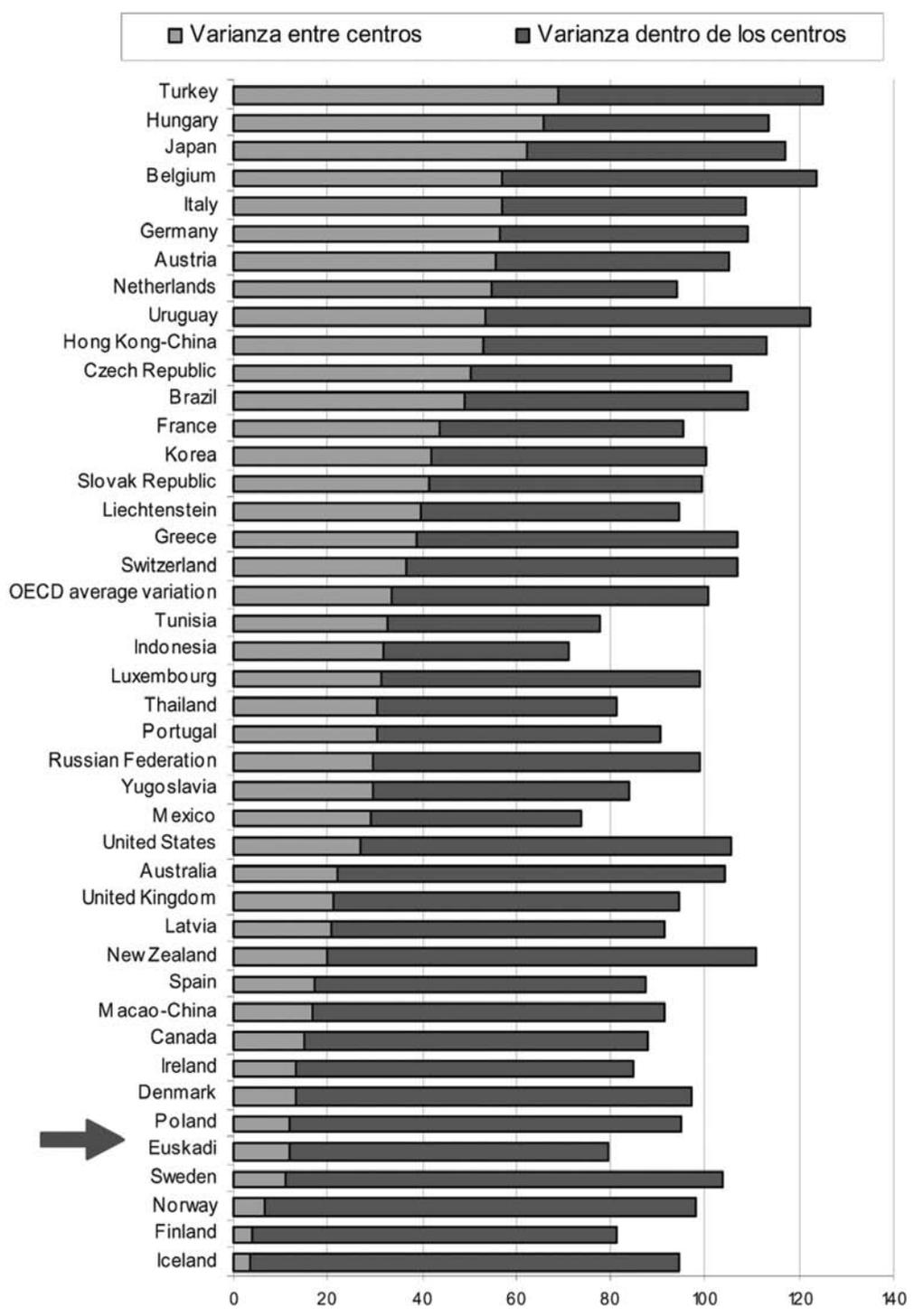
Pero, además, tenemos, en matemáticas, una de las menores diferencias entre centros de todos los países participantes: tras Islandia, Finlandia, Noruega y Suecia, nuestra diferencia de rendimiento entre centros es la más baja de PISA.

LAS DIFERENCIAS DE NIVEL SOCIOECONÓMICO Y CULTURAL ENTRE LA RED PÚBLICA Y LA CONCERTADA

Como es sabido, en Euskadi la red concertada agrupa a más de un 50% del alumnado en la muestra PISA, un porcentaje superior al de España y equivalente al de Irlanda, por ejemplo. Nuestro porcentaje de alumnos en la red privada (según definición PISA) es solamente del 3%.

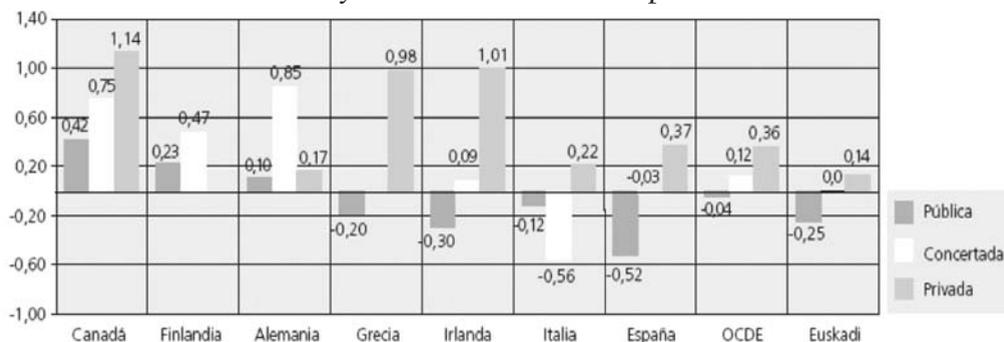
Sin embargo, las diferencias de nivel socioeconómico y cultural entre las redes pública y concertada en el índice calculado por PISA son en Euskadi de las más bajas de la OCDE.

FIGURA V
Diferencias entre centro y dentro del centro en matemáticas



CUADRO V

Nivel socioeconómico y cultural del alumnado por titularidad del centro



Todos estos datos conforman, por tanto, un sistema educativo en Euskadi de rendimiento compacto, de gran equidad y de escaso influjo del nivel socioeconómico y cultural. Un sistema que, al menos, no multiplica las diferencias sociales y que, probablemente es muy compensador.

LOS PUNTOS MANIFIESTAMENTE MEJORABLES

EL CONOCIMIENTO CIENTÍFICO

No obstante, no todo son luces en los resultados de la evaluación PISA.

El rendimiento obtenido por Euskadi en Conocimiento Científico está claramente por debajo de la media de la OCDE y contrasta con el resto de los rendimientos evaluados.

El Departamento de Educación está analizando las medidas que deben adoptarse para mejorar el rendimiento en esta área.

LA AUTONOMÍA DE LOS CENTROS

Otro de los aspectos mejorables que PISA nos ha mostrado es el de la autonomía de los centros.

Encuestados los directores de los centros participantes acerca de su capacidad de decisión en distintos ámbitos, he aquí el resultado:

CUADRO VI

CAPACIDAD DE DECISIÓN ÁMBITO	Publica %	Concertada %	TOTAL %
Contratación del profesorado	0.0	98.8	58.4
Despido del profesorado	3.6	100.0	60.3
Salarios iniciales del profesorado	0.0	38.3	22.6
Aumentos salariales del profesorado	0.0	34.6	20.4

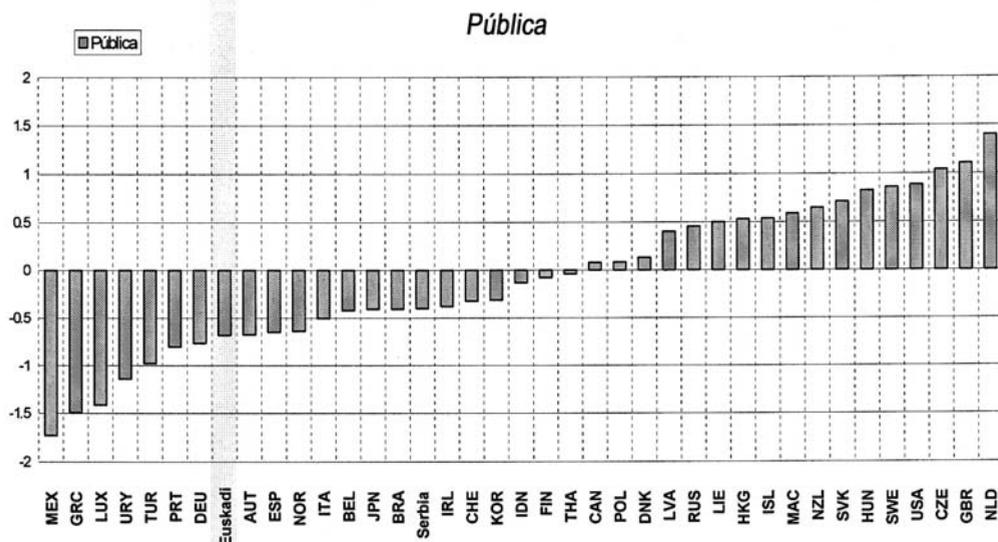
(continúa)

CUADRO VI (continuación)

Presupuesto	Elaboración del presupuesto del centro	59.3	100.0	83.6
	Distribución presupuestaria dentro del centro	100.0	100.0	100.0
Normas	Normas de disciplina para el alumnado	98.2	100.0	99.3
	Normas para la evaluación del alumnado	98.2	97.5	97.8
	Admisión del alumnado	50.0	93.8	76.3
Currículo	Elección de los libros de texto	100.0	100.0	100.0
	Determinación de los contenidos de los cursos	76.8	90.1	84.7
	Decisión sobre qué enseñanzas se ofrecen	46.4	95.0	75.0

En comparación con los países participantes, en la red pública, solamente tienen menos autonomía que nosotros México, Grecia, Luxemburgo, Uruguay, Turquía, Portugal y Alemania. Además, está claro que no se trata de los países con mejor rendimiento.

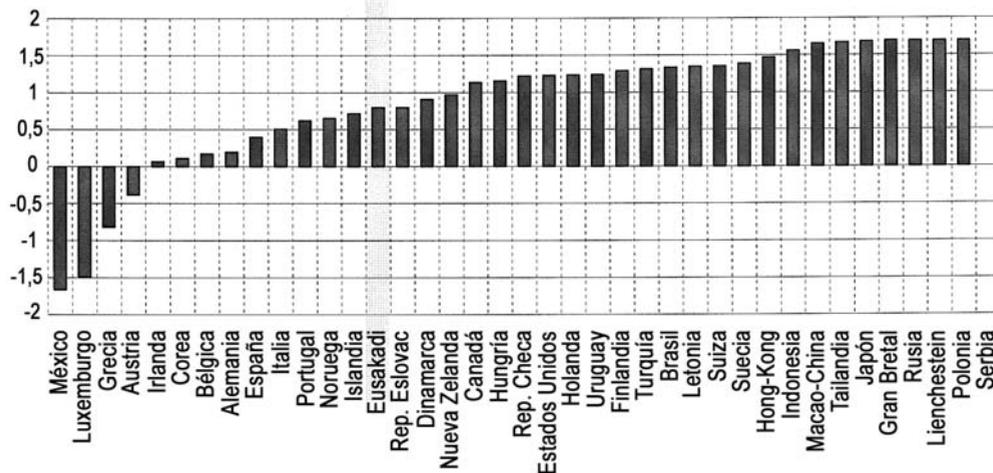
CUADRO VII



En la red concertada tampoco los resultados son muy buenos.

CUADRO VIII

Nivel de autonomía de centros concertados



DIFERENCIAS DE RENDIMIENTO POR SEXO DEL ALUMNO

En tres de las cuatro competencias evaluadas por PISA, los chicos vascos obtienen una puntuación más baja que las chicas, llegando a los 45 puntos en lectura. Esto nos sitúa junto a los países que mayor diferencia de rendimiento tienen entre sexos. Merece una reflexión por parte de nuestra comunidad educativa.

TAMBIÉN PARTICIPAMOS EN TIMSS 2003, EVALUACIÓN DE LAS MATEMÁTICAS Y LAS CIENCIAS EN 2º DE ESO

Euskadi ha participado también en esta evaluación impulsada por la IEA en solitario, ya que España abandonó el estudio poco antes de la aplicación de pruebas.

Resumiendo nuestros resultados en MATEMÁTICAS, estamos en la misma puntuación que Nueva Zelanda, Eslovenia e Italia. Por encima de la media internacional.

En CIENCIAS: en la misma puntuación que Noruega, Italia, Israel y Bulgaria. Por encima de la media internacional.

Se reafirman los resultados de PISA en cuanto a equidad del Sistema Educativo vasco.

No son coincidentes los resultados en CIENCIAS. Los resultados TIMSS no difieren mucho de los de MATEMÁTICAS. No se da la diferencia visible en PISA.

Hay que tener en cuenta que TIMSS es un estudio curricular y PISA no lo es. Además, TIMSS se aplica en 2º de la ESO y PISA va dirigido a una edad: 15 años.

En definitiva, hay un antes y un después en nuestro sistema a partir del 2004. Gracias a nuestra apuesta por la evaluación internacional, estamos en condiciones de tomar decisiones sobre propuestas para la mejora del rendimiento de nuestros alumnos y sabemos mucho más acerca nosotros mismos que antes.

LÍMITES DE LA EVALUACIÓN: MUESTRAS, CENSOS, PRUEBAS DIAGNÓSTICAS

Uno de los problemas de la Evaluación de Sistemas Educativos es el de la limitación de los estudios muestrales.

Los estudios basados en una muestra de centros y alumnos están limitados porque sólo sirven para dar una visión general del sistema. Permiten detectar problemas generales de rendimiento pero muchas veces no explican sus causas. Las variables explicativas se quedan en un plano general de nivel socioeconómico-cultural, formación de los padres, estructura del Sistema Educativo, etc.

Una de sus más severas limitaciones es que no se puede hablar de centros, y menos de alumnos de determinados centros.

Ante la publicación de los resultados, los centros escolares no se consideran involucrados y no se puede hacer un «mapa» de rendimiento del sistema o un «panel de mando» que permita actuar allá donde realmente se necesite en términos de grupos, centros o alumnado.

Esta limitación no es inherente a las evaluaciones internacionales sino a cualquier evaluación. Las evaluaciones internacionales permiten ser aplicadas a todos los alumnos de una edad o curso pero esto, naturalmente, encarece sensiblemente la aplicación y el análisis de datos.

LA TENDENCIA ACTUAL EN OTROS PAÍSES

La corriente actual en evaluación es claramente a favor de los sistemas de Rendición de Cuentas o *Accountability*.

Combina pruebas únicas, basadas en estándares casi siempre, aplicadas a todos los alumnos con una devolución de la autonomía y la responsabilidad a los centros educativos.

Estos sistemas utilizan pruebas censales.

Las pruebas (diagnósticas o no) censales:

- Permiten conocer dónde están los problemas y cuales son. Muchas veces, las evaluaciones muestrales detectan un problema pero no pueden explicar las causas.
- Sin embargo, el efecto de las pruebas diagnósticas es que se convierten *de facto* en «el currículum», pasando la evaluación a convertirse en norma para el profesorado. Por esto es importante que:
 - Las pruebas (los estándares en que se basan) no sean reduccionistas ni simplificadoras.
 - No abusen de los ítems de respuesta múltiple y contengan preguntas abiertas, de construcción compleja, de interpretación...
 - Sean fruto de un gran consenso.

EL VALOR AÑADIDO

Uno de los resultados menos deseables de las evaluaciones censales es el establecimiento de clasificaciones de rendimiento de centros educativos. Las administraciones se sienten muchas veces obligadas, si quieren evitar la injusticia de comparar centros con alumnado absolutamente dispar y situaciones económicas de privilegio con marginación social, a publicar informes introduciendo la variable nivel socioeconómico y cultural para explicar los resultados.

El paradigma que supera estas clasificaciones básicamente injustas es el del valor añadido. No se mide sólo el rendimiento sino la mejora.

Su aplicación masiva comenzó en el Estado norteamericano de Tennessee hace ya diez años, como medio para establecer complementos para el profesorado basados en la eficacia de su trabajo.

Tennessee aplica un test criterial llamado TCAP (1990. CTB/Mcgraw Hill) a todos los alumnos durante los niveles 3º al 5º basado en sus propios estándares (los estándares allí son estatales, no federales).

Estas aplicaciones permiten ver:

- El rendimiento de todo el alumnado con respecto a los estándares.
- *La ganancia año a año* atribuida a determinado profesorado, ya que se controlan las variables extraescolares que pudieran afectar al rendimiento.

Tomemos el ejemplo de un centro con un 96,7% de alumnado en nivel socioeconómico bajo y un 60% de afroamericanos. La escala va de la A (máxima puntuación) a la F (mínima). La publicación de los resultados en la web tiene este aspecto:

CUADRO IX

Student Performance				
Grades K-8: Academics				
	Achievement		Value-Added	
	2002 Grade	2003 Grade	2002 Grade	2003 Grade
Math	F	F	C	C
Reading	F	F	A	B
Language	F	F	D	F
Social Studies	F	F	C	A
Science	F	F	A	A
Writing 4th/5th	C	C		

(Note: Based on norm referenced test 3-year averages.)

Escala A-F: malos resultados en rendimiento

Bastante mejores resultados en valor añadido, con la excepción de lengua.

Lo que ocurre en Tennessee es que son los centros más privilegiados los que se quejan porque su rendimiento es excelente pero su valor añadido es muy bajo. La polémica allí es ésta: ¿para qué necesitamos valor añadido si tenemos unos resultados inmejorables? He aquí el ejemplo de un centro de alto nivel socioeconómico y alumnado blanco.

CUADRO X

Student Performance				
Grades K-8: Academics				
	Achievement		Value-Added	
	2002 Grade	2003 Grade	2002 Grade	2003 Grade
Math	A	A	B	B
Reading	A	A	F	D
Language	A	A	F	F
Social Studies	A	A	F	F
Science	A	A	F	F
Writing 4th/5th	A	A		

(Note: Based on norm referenced test 3-year averages.)

Los centros con excelentes resultados, por tanto, apenas tienen mejora año a año. ¿Es realmente más difícil mejorar cuando los resultados son ya excelentes que cuando estos son malos? Ésta es una de las cuestiones controvertidas del valor añadido.

Pero este sistema del valor añadido tiene también otros problemas.

Hay muchas formas de calcular el valor añadido pero el único sistema «comercial», con capacidad para hacerlo en breve espacio de tiempo y para miles de sujetos, es el de la empresa SAS, para quien trabaja uno de los científicos pioneros de este sistema: William Sanders.

La dependencia de esta empresa, al cabo de los años, llega a ser total y no revelan su software ni publican en revistas especializadas cómo lo hacen.

LAS PRUEBAS DIAGNÓSTICAS CENSALES

Parece que las PRUEBAS DIAGNÓSTICAS propuestas para 4º de Primaria y 2º de Secundaria propuestas en la LOE serían similares, en alguna medida, a las citadas más arriba, al hablar de las tendencias actuales en evaluación.

La LOCE ya propugnaba las pruebas diagnósticas. Pero no había prevista una devolución de autonomía y responsabilidad a los centros, cosa que ocurre en todos los países con un sistema de *accountability*. El control sin la autonomía, en mi opinión, no funcionará.

Si lo centros no asumen responsabilidad alguna sobre sus resultados, el sistema de rendición de cuentas no funciona bien.

Hay otra cuestión clave también: «La evaluación sin consecuencias no es evaluación», dicen en Alberta (Canadá), luego es necesario prever mecanismos de mejora una vez que se conozcan los resultados (esto vale para todas las evaluaciones). No se puede dejar todo como está después de la evaluación, sobre todo en aquellos centros que obtengan pobres resultados.

El problema es que la estructura actual del Sistema Educativo vasco no permite, más que de un modo limitado, la intervención asimétrica en algunos centros y en otros no y esto es especialmente grave en cuestiones de gestión de personal. Habría que preguntarse si el sistema funcional actual, tal y como está en la red pública, resulta compatible con este tipo de rendición de cuentas. Incluso en Francia disponen de un sistema más descentralizado teniendo en cuenta el tipo de dirección de centros con que cuentan y las llamadas Zonas Educativas Prioritarias con una asignación de recursos muy especial.

También hay que tener en cuenta el diferente uso que se hace en distintos países de las pruebas censales.

Mientras en Alberta o Suecia, por poner un ejemplo de países con un sistema absolutamente descentralizado del ministerio, los resultados se hacen públicos centro por centro y la ciudadanía los puede consultar en Internet con las consecuencias que esto tiene en la elección de centro por los padres, en otros países como Francia, más centralizados, el Ministerio de Educación se queda con una muestra de la que publica un informe anual y no se dan a conocer los resultados de todos los centros.

Alemania ha sido el último país que ha construido un sistema de rendición de cuentas, a raíz de PISA 2000, habiendo comenzado la aplicación de pruebas en septiembre de 2004.

En Canadá o Suecia, el ministerio dispone cada año de un «panel de mando» con todos los resultados de todos los centros y actúa en consecuencia donde hace falta.

En Francia se pretende más bien que el propio profesorado y el centro saquen sus conclusiones con las referencias que se les facilitan sobre la evaluación¹.

Por otro lado, hay efectos colaterales previsibles. Los centros que obtengan buenos resultados en las pruebas diagnósticas los harán públicos con el fin de demostrar a los padres su alto nivel de rendimiento, lo cual suele exigir hacer públicos algunos informes explicando las diferencias entre centros a raíz del nivel socioeconómico de los alumnos o el nivel cultural de las familias.

Otra consideración es que estas pruebas deben tener estándares adaptados a la realidad de cada comunidad, aunque se puede establecer un mecanismo de comparación o equivalencia. También existe el problema de la recogida de los datos (normalmente vía Internet), el del análisis de una cantidad ingente de ellos en un tiempo razonable y, además, el de la financiación de todo ello. Es necesario creer en la evaluación y apostar decididamente por ella como paso previo a la mejora para dedicar los recursos necesarios a tal fin. Dejo para el final el pro-

¹ Ver: «Évaluations diagnostiques»: <http://educ-eval.education.fr/evaldiag.htm> y <http://www.evace26.education.gouv.fr>.

blema de las dificultades técnicas de la construcción, aplicación y corrección de estas pruebas con garantías.

¿QUÉ HACER ANTE TANTA PROLIFERACIÓN DE EVALUACIONES?

Muchos países tienen complejos sistemas de evaluación que, sin embargo, no se solapan. En Tennessee, por ejemplo, el NAEP es una evaluación muestral a nivel federal que permite comparar los resultados de los Estados. El programa *No Child Left Behind* es de financiación federal pero se basa en los estándares de cada Estado y los sistemas como el del valor añadido son exclusivamente estatales y cada vez se aplica en más Estados.

Uno de los problemas de las evaluaciones diagnósticas propuestas en la LOE para 4º de primaria y 2º de ESO es que, a la vez, parece que se propone seguir con las evaluaciones estatales y con las internacionales.

Esto supone EVALUACIONES REDUNDANTES, con el peligro de dilapidar los recursos y encontrarnos con resultados dispares que serían difíciles de explicar.

Que las futuras pruebas diagnósticas coincidan con evaluaciones internacionales en el mismo nivel puede ser una duplicidad innecesaria.

Mucho más lo es que las evaluaciones estatales coincidan con evaluaciones internacionales en los finales de algunas etapas. Sería más práctico someterse al dictamen de PISA en cuanto a competencias clave, por ejemplo, ampliándolo con otras pruebas en el caso de que se considere necesario.

Parece, en definitiva, que sería necesario reflexionar acerca de lo que queremos realmente en evaluación y qué podemos llevar a cabo.

Las evaluaciones diagnósticas censales tienen una potencia evaluativa enorme pero, en mi opinión, deben combinarse con la autonomía y responsabilidad de los centros sobre sus resultados.

Las evaluaciones internacionales son pues nuestra apuesta actual incluso como evaluación de final de etapa. PISA por ejemplo tiene una opción denominada de 10º nivel que permite hacer una muestra de 4º de ESO y convertirla en evaluación de secundaria obligatoria.

CONSIDERACIONES FINALES

Creo, en definitiva, que hemos comenzado a transitar en Euskadi por un camino que nos mostrará sus primeros resultados a medio y largo plazo. Comenzamos a ser conscientes de nuestros resultados y esto, además de crear una cultura evaluativa e interpelarnos (así lo espero) acerca de nuestro trabajo en los centros, en los servicios de apoyo y en la administración, es el primer e imprescindible paso para la mejora. A partir de los resultados es posible y deben tomarse decisiones. No obstante, debemos pensar ya en soluciones más perfeccionadas como los sistemas de rendición de cuentas. Éstos nos permitirían un diagnóstico más exacto y una intervención más precisa a todos los niveles, llevando los recursos allá donde sean más necesarios pero estableciendo compromisos. Los cambios nunca son fáciles pero nada puede mejorar sin ellos.

REFERENCIAS BIBLIOGRÁFICAS

EUROPEAN COMMISSION. DIRECTORATE-GENERAL FOR EDUCATION AND CULTURE (2000):
European Report on the quality of School Education.

OCDE (2004): *Messages from PISA 2000.* Paris.

PRIMER INFORME DE LA EVALUACIÓN PISA 2003. Resultados en Euskadi.

SEGUNDO INFORME DE LA EVALUACIÓN PISA 2003. Resultados en Euskadi.

INECSE (2004): EVALUACIÓN PISA 2003. Resumen de los primeros resultados. Madrid.

TENNESSEE EVALUATION AND ASSESSMENT.

THE TENNESSEE VALUE-ADDED ASSESSMENT SYSTEM.

PÁGINAS WEB

<http://europa.eu.int/comm/education/policies/educ/indic/rapinen.pdf>

<http://www.pisa.oecd.org/dataoecd/31/19/34107978.pdf>

<http://www.isei-ivei.net/cast/pub/indexpub.htm>

<http://www.ince.mec.es/pub/pubintn.htm#ref13>

<http://www.state.tn.us/education/assessment>

http://www.shearonforschools.com/TVAAS_index.html