

TIMSS Y PISA. DOS PROYECTOS INTERNACIONALES DE EVALUACIÓN DEL APRENDIZAJE ESCOLAR EN CIENCIAS

José Antonio Acevedo Díaz
Consejería de Educación de la Junta de Andalucía.
Inspección de Educación. Delegación Provincial de Huelva.
E-mail: ja_acevedo@vodafone.es

RESUMEN

Durante la última década se han incrementado las aplicaciones de proyectos nacionales e internacionales de evaluación del aprendizaje escolar en ciencias y otras materias, posiblemente como consecuencia de las presiones políticas a las que se ven sometidos los sistemas educativos para evaluar su rendimiento. Los estudios e informes derivados de la puesta en práctica de las evaluaciones transnacionales han permitido la comparación entre países y, más allá de la comprobación de los niveles de conocimiento y comprensión del alumnado, se suelen usar como una medida global de la calidad de los sistemas educativos. En este artículo se muestran las principales características de los dos proyectos de este tipo más conocidos: TIMSS y PISA. Se pretende con ello animar al profesorado a que haga una lectura crítica más profesional y un análisis más reflexivo de los informes derivados de estas evaluaciones, para así llegar a sacar conclusiones por sí mismo.

Palabras clave: *aprendizaje de las ciencias, evaluación, proyectos internacionales, TIMSS, PISA.*

"La radiografía del estado de la educación en los países de la OCDE, el Informe PISA 2003, ha hecho saltar la alarma sobre los malos resultados obtenidos por el sistema español. Aunque los conocimientos y habilidades de los españoles en matemáticas, comprensión lectora y cultura científica ya aparecían a la cola de los países desarrollados en 2000, el macroestudio hecho público el lunes revela un preocupante estancamiento, e incluso un empeoramiento en lectura..."

(Editorial de El País 8-12-2004)

INTRODUCCIÓN

Noticias como la de la entradilla y los correspondientes artículos de opinión brotan en la prensa impresa y digital durante unos cuantos días por su actualidad y cierta dosis de catastrofismo, oyéndose también toda clase de opiniones –a veces con muy poco pudor– en las tertulias radiofónicas, que se difuminan en pocos días para dar paso a otras informaciones. En consecuencia, también la mayor parte de la sociedad tiende a olvidar pronto estas noticias, hasta que la publicación de otro informe similar vuelve a dar la señal de alarma. Frente a esta actitud, los profesionales de la educación

deberíamos esforzarnos en hacer un análisis mucho más sereno, profundo y riguroso de estos informes, opinando con mayor conocimiento del tema sobre lo que nos dicen.

Desde hace aproximadamente tres lustros, los sistemas educativos están siendo objeto de muchas presiones políticas para evaluar su rendimiento. Como resultado de ello, han aumentado los sistemas nacionales e internacionales de evaluación, los estudios transnacionales y las comparaciones entre países, que se utilizan como una medida relativa de la calidad global de la educación. De este modo, la evaluación externa ha ido mucho más allá de la comprobación de los niveles de conocimiento y comprensión del alumnado, habiendo llegado a ser un indicador de la capacidad de los profesores, del rendimiento de las escuelas y del propio sistema educativo (Pilot, 2000). Así, durante la última década se han puesto en marcha varios proyectos internacionales para evaluar los resultados del aprendizaje escolar en ciencias, entre otras materias, siendo los dos más importantes el TIMSS (López-Varona y Moreno-Martínez, 1996, 1997; Vázquez, 2000) y el PISA (Harlen, 2001, 2002).

Ciertamente, los análisis transnacionales del rendimiento que logran los estudiantes permiten ampliar y enriquecer la visión nacional, proporcionando un contexto más amplio en el que interpretar los resultados propios. Aunque no son estudios académicos, porque están dirigidos sobre todo a los administradores y gestores de la educación –*policy oriented studies*–, pueden servir para orientar a los centros educativos y detectar los puntos fuertes y débiles del currículo, pues no sólo son esenciales las deliberaciones teóricas e ideológicas, sino que también hacen falta pruebas empíricas (Sjøberg, 2004). Con los incentivos apropiados, pueden motivar la mejora del aprendizaje de los alumnos, de la enseñanza de los profesores y la eficacia de los centros educativos. Así mismo, aportan diversos instrumentos para el seguimiento de los niveles de rendimiento escolar por parte de las autoridades educativas, incluso en el caso de que las competencias de la administración educativa estén descentralizadas. El valor de las evaluaciones y las investigaciones comparativas transnacionales es indudable, porque permiten aprender de los demás y entender mejor que la educación se basa en preferencias políticas y tradiciones culturales, comprendiendo también que la situación propia puede ser diferente.

En la evaluación se vierten diferentes creencias sobre lo que es y no es importante de una materia –sólo se considera verdaderamente importante lo que se evalúa–, acerca de qué es y cómo se produce el aprendizaje y sobre cómo se puede medir. Bien planteadas, las evaluaciones externas pueden servir de acicate para orientar la enseñanza de las ciencias hacia las innovaciones de los currículos reformados en consonancia con las aportaciones de la investigación en didáctica de las ciencias (Nieda, Cañas y Martín-Díaz, 2004; Martín-Díaz, 2004), pero con diseños mucho más tradicionales pueden contribuir a todo lo contrario, como sucede con las pruebas de acceso a la universidad, que refuerzan la enseñanza propedéutica y condicionan las innovaciones curriculares hasta llegar a reprimirlas (Oliva y Acevedo, 2005).

Las evaluaciones externas son importantes, por lo que es necesario un conocimiento más profesional de los informes transnacionales de evaluación del aprendizaje escolar. No obstante, al contrario de lo que suele suceder en la bibliografía especializada anglosajona, las publicaciones españolas de didáctica de las ciencias prestan muy poca

atención a los proyectos internacionales de evaluación externa y a los informes que se derivan de sus aplicaciones, como puede comprobarse revisando la correspondiente bibliografía. Para contribuir a paliar en parte esta situación, en este artículo se mostrarán algunas de las principales características de esos proyectos con la intención de animar al profesorado a que haga una lectura más crítica de ellos, los analice de una manera más reflexiva y saque sus propias conclusiones.

EL PROYECTO TIMSS¹

TIMSS es el acrónimo de *Third International Mathematics and Science Study*, un proyecto de evaluación internacional del aprendizaje escolar en matemáticas y ciencias realizado por la *International Association for the Evaluation of Educational Achievement* (IEA)², que se aplicó en 1995 (Beaton *et al.*, 1996; Martin *et al.*, 1997; Mullis *et al.*, 1998). Puesto que en 1990 la Asamblea General de la IEA³ decidió hacer sus evaluaciones de manera regular cada cuatro años, volvió a aplicarse en 1999 con el nombre de *TIMSS Repeat* (Martin *et al.*, 2000). En el año 2003 cambió su nombre a *Trends in International Mathematics and Science Study*, pero manteniendo el acrónimo que lo identifica (figura 1), por lo que éste se conoce como *TIMSS Trends* (Mullis *et al.*, 2002). España participó en 1995 (7º y 8º EGB), pero no lo ha hecho en las dos ocasiones siguientes, excepto la Comunidad Autónoma del País Vasco que sí ha participado en el año 2003.



Figura 1.- Libros del TIMSS Trends 2003.

Las personas que contribuyen al desarrollo de las diversas partes del TIMSS proceden de todas las áreas de evaluación educativa, incluyendo expertos en el análisis de políticas educativas, educación matemática, educación científica, diseño del currículo, supervisión de la investigación, construcción de pruebas, psicometría, encuestas, muestreo y análisis de datos.

El proyecto TIMSS evalúa el rendimiento de los estudiantes en matemáticas y ciencias para aprender más de la naturaleza y el alcance del aprendizaje de los estudiantes en estas dos materias, así como del contexto en que ello ocurre. Pretende encontrar factores directamente relacionados con el aprendizaje de los estudiantes en ambas materias que puedan modificarse por la política educativa, tales como el currículo, la asignación de recursos o las prácticas de enseñanza.

Durante el año 2003, TIMSS *Trends* ha evaluado el rendimiento del alumnado en matemáticas y ciencias en dos cursos: 4º grado (4º de Primaria en España) y 8º grado (2º de ESO en España). Como estos cursos son los mismos que se evaluaron en 1995 y 1999, TIMSS *Trends* proporciona información de la tendencia a lo largo de un período de ocho años (Martin *et al.*, 2004). Aunque de naturaleza muy distinta, estos datos también pueden servir de complemento a los que proporciona el proyecto internacional PISA de la OCDE, que se aplica a poblaciones de estudiantes de 15 años de edad.

Los resultados de las pruebas del TIMSS *Trends* se pueden emplear para diversos fines. Los responsables de las políticas educativas y los investigadores académicos obtienen datos generales sobre el rendimiento del alumnado en matemáticas y ciencias que:

- Amplían las evaluaciones realizadas en matemáticas y ciencias iniciadas con el TIMSS 1995 y continuadas con el TIMSS 1999.
- Permiten hacer comparaciones contrastadas del rendimiento en matemáticas y ciencias entre los países participantes y, junto a otros datos que proporciona el TIMSS 2003, sugieren posibles razones para explicar las diferencias encontradas.
- Pueden mejorar la evaluación de la eficacia de la enseñanza y el aprendizaje de las matemáticas y las ciencias en cada país participante.
- Resaltan aspectos del crecimiento de los conocimientos y destrezas matemáticos y científicos desde cuarto a octavo curso.
- Proporcionan información para otros análisis relacionados con la mejora de los niveles de rendimiento en los sistemas educativos, los centros de enseñanza y las prácticas didácticas mediante políticas educativas más y mejor informadas.

El modelo de evaluación curricular del TIMSS

TIMSS parte de la idea de que puede conseguirse una comprensión adecuada de los elementos que influyen en el aprendizaje de las matemáticas y las ciencias mediante un análisis cuidadoso del rendimiento de los estudiantes, las características de éstos, el currículo que siguen, las metodologías didácticas de sus profesores y los recursos disponibles en sus escuelas y aulas. Las características del aula y de la escuela son un reflejo de la comunidad y del sistema educativo, que, a la vez, también son aspectos particulares de la sociedad del país en el que viven los estudiantes.

Para conseguir lo anterior, los diseñadores del TIMSS han establecido un modelo conceptual curricular como amplio factor explicativo que subyace al rendimiento de los

estudiantes (Robitaille y Garden, 1996). Desde esta perspectiva, se consideran tres niveles del currículo (véase la figura 2):

- (i) El *currículo pretendido*, es decir, el que se planifica para la enseñanza.
- (ii) El *currículo aplicado*, esto es, el que se enseña en el aula.
- (iii) El *currículo logrado*, que es el que los estudiantes consiguen aprender.

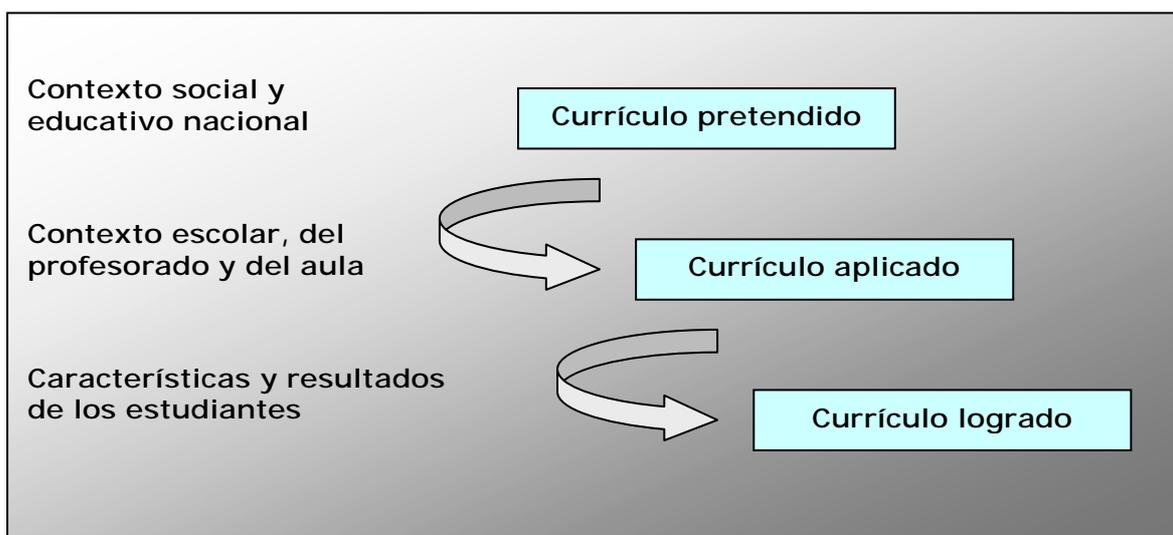


Figura 2.- Modelo curricular del TIMSS.

El *currículo pretendido* es la respuesta planificada a las finalidades educativas de la sociedad, en general, y de una comunidad educativa, en particular, para la enseñanza y el aprendizaje escolar; unas metas que reflejan los ideales y las tradiciones sociales y comunitarias, pero que se ven limitadas por los recursos del sistema educativo. De otro modo, representa lo que la sociedad pretende que aprendan los estudiantes y cómo debería organizarse y planificarse el sistema educativo para facilitar este aprendizaje.

El *currículo aplicado* se corresponde con lo que realmente se enseña en el aula, con quién lo imparte y cómo lo hace. Aunque se supone que se inspira en la planificación del currículo pretendido, hay que tener en cuenta que lo que sucede en el aula suele responder en gran parte a la actuación del profesor, cuyo comportamiento puede estar muy influido por su formación y experiencia, la naturaleza y estructura orgánica de la escuela, su interacción con los compañeros y las características del alumnado, entre otros factores.

El *currículo logrado* hace referencia a lo que los estudiantes consiguen aprender y a lo que piensan sobre las materias aprendidas. El rendimiento de los estudiantes depende en parte del currículo aplicado y de su contexto social y educativo, así como de las características de cada estudiante, incluyendo entre ellas su capacidad, actitud, intereses, motivación y esfuerzo.

Con este modelo, TIMSS utiliza pruebas de rendimiento en matemáticas y ciencias para describir el aprendizaje de los estudiantes de los países participantes, además de

cuestionarios de contexto que proporcionan una gran cantidad de información. Estos últimos preguntan sobre la estructura y el contenido del currículo pretendido en matemáticas y ciencias, la preparación, experiencia y actitudes de los profesores, los contenidos de matemáticas y ciencias que realmente se imparten, los enfoques didácticos empleados, la organización y los recursos de los centros y las aulas y, por último, las experiencias y actitudes de los estudiantes en los centros docentes.

El conocimiento del currículo pretendido en cada país participante se obtiene en una fase inicial de aplicación del TIMSS a partir de un análisis minucioso de las orientaciones de los documentos curriculares, los libros de texto y otros materiales de apoyo, así como de los cuestionarios respondidos por expertos en educación y especialistas del currículo. Su finalidad es mostrar con detalle las intenciones curriculares de los países participantes.

Los datos para el estudio del currículo aplicado se toman en la fase de aplicación de la evaluación del rendimiento de los estudiantes. Los cuestionarios contestados por el profesorado de matemáticas y ciencias, los estudiantes y los directores de sus escuelas contienen información sobre características del profesorado –formación, experiencia, actitudes y opiniones–, los temas de ciencias y matemáticas que se enseñan y aquellos en los que se pone más énfasis, los enfoques metodológicos empleados en el aula, las estructuras organizativas y los recursos que sirven de apoyo a la enseñanza y los aspectos que facilitan o dificultan ésta y el aprendizaje.

La evaluación del rendimiento de los estudiantes suministra numerosos datos para el estudio del currículo logrado. Por un lado, la información de las características de los estudiantes, sus actitudes, creencias y experiencias se obtiene por medio de los cuestionarios que responde cada alumno participante en el estudio; esta información sirve para identificar los principales rasgos de los estudiantes relativos al aprendizaje y permite establecer un contexto para el posterior análisis de lo que han aprendido. Por otro lado, los resultados de las pruebas aplicadas de matemáticas y ciencias proporcionan información sobre el conocimiento del currículo logrado en cada país, lo que posibilita la elaboración de una amplia base de datos para hacer comparaciones transnacionales del rendimiento de los estudiantes.

Las pruebas del TIMSS

El principio fundamental para elaborar las pruebas de aplicación del TIMSS es la producción de instrumentos de evaluación que generen datos de rendimiento fiables y válidos para los fines que se pretenden. A partir de los marcos teóricos expuestos, las pruebas se elaboran mediante un proceso de consenso internacional, con aportaciones de expertos en educación, matemáticas, ciencias y evaluación. Estas pruebas tienen preguntas cerradas de opción múltiple, algunas de respuesta abierta y otras de resolución de problemas. La mayoría de los ítems de las pruebas del TIMSS *Trends* se centran en un contenido concreto, pero también se necesitan conocimientos o destrezas procedentes de otras áreas de contenidos. Además, determinados temas se han enunciado de forma más general que en ocasiones anteriores y algunos de los ítems de nueva creación exigen a los estudiantes que sinteticen conocimientos y tengan que poner en práctica más de una destreza. En la aplicación del año 2003, se

ha hecho más hincapié que otras veces en preguntas y tareas que impliquen destrezas y capacidades analíticas, de resolución de problemas y de investigación.

La evaluación de ciencias del TIMSS *Trends* aplicada en el año 2003 se basa en dominios de contenido y dominios cognitivos. Los primeros incluyen diversas áreas específicas de las materias de ciencias evaluadas, mientras que los segundos corresponden a las destrezas necesarias para abordar los contenidos de ciencias. En la tabla 1 se indican los porcentajes del tiempo dedicado en la evaluación a ambos dominios. Además de los dominios de contenido señalados, la aplicación del año 2003 también ha incluido la evaluación de la investigación científica como área separada, dedicándose a la misma aproximadamente el 15% del tiempo total de las pruebas.

POBLACIÓN EVALUADA	4º CURSO	8º CURSO
Porcentajes de tiempo de dominios de contenido⁴ (%)		
Ciencias de la vida	45	30
Física y Química	35	–
Física	–	25
Química	–	15
Ciencias de la Tierra	20	15
Ciencias medioambientales	–	15
Porcentajes de tiempo de dominios cognitivos (%)		
Conocimiento factual	40	30
Comprensión conceptual	35	35
Razonamiento y análisis	25	35

Tabla 1.- Porcentajes del tiempo dedicado en cada curso a los dominios de contenido y cognitivos en la evaluación del rendimiento en ciencias del TIMSS *Trends*.

Los ítems de conocimiento factual son: (i) preguntas de recuerdo y reconocimiento, (ii) definiciones, (iii) descripciones y (iv) cuestiones de conocimiento sobre el uso de instrumentos y procedimientos de medidas científicas.

Las preguntas de comprensión conceptual requieren: (i) ilustrar con ejemplos, (ii) comparar, contrastar y clasificar, (iii) representar y modelizar, (iv) relacionar, (v) obtener y aplicar información, (vi) encontrar soluciones y (vii) explicar.

Las preguntas de razonamiento y análisis implican: (i) analizar, interpretar y resolver problemas, (ii) asociar, integrar y sintetizar, (iii) formular hipótesis y predecir, (iv)

diseñar y planificar, (v) recopilar, analizar e interpretar datos, (vi) sacar conclusiones, (vii) generalizar, (viii) evaluar y (ix) justificar.

Por último, los dominios de contenido evaluados por el TIMSS *Trends*, incluyendo la investigación científica, y sus principales áreas temáticas se muestran en la tabla 2.

DOMINIOS DE CONTENIDO DE CIENCIAS Y ÁREAS TEMÁTICAS PRINCIPALES	
<p style="text-align: center;">Ciencias de la vida</p> <ol style="list-style-type: none"> 1. Tipos, características y clasificación de los seres vivos. 2. Estructura, función y procesos vitales de los organismos. 3. Las células y sus funciones. 4. Desarrollo y ciclos de vida de los organismos. 5. Reproducción y herencia. 6. Diversidad, adaptación y selección natural. 7. Ecosistemas. 8. La salud humana. 	<p style="text-align: center;">Física</p> <ol style="list-style-type: none"> 1. Estados físicos y cambios en la materia. 2. Tipos, fuentes y conversiones de energía. 3. Calor y temperatura. 4. Luz. 5. Sonido y vibración. 6. Electricidad y magnetismo. 7. Fuerzas y movimiento.
<p style="text-align: center;">Química</p> <ol style="list-style-type: none"> 1. Clasificación y composición de la materia. 2. La estructura de partículas de la materia. 3. Propiedades y usos del agua. 4. Ácidos y bases. 5. El cambio químico. 	<p style="text-align: center;">Ciencias de la Tierra</p> <ol style="list-style-type: none"> 1. Estructura y rasgos físicos de la Tierra (litosfera, hidrosfera y atmósfera). 2. Procesos, ciclos e historia de la Tierra. 3. La Tierra en el sistema solar y en el universo.
<p style="text-align: center;">Ciencias medioambientales</p> <ol style="list-style-type: none"> 1. Cambios en la población. 2. Utilización y conservación de los recursos naturales. 3. Cambios en el medio ambiente. 	<p style="text-align: center;">Investigación científica</p> <ol style="list-style-type: none"> 1. Formular preguntas e hipótesis. 2. Diseñar las investigaciones. 3. Recopilar y representar los datos. 4. Analizar e interpretar los datos. 5. Sacar conclusiones y elaborar explicaciones.

Tabla 2.- Dominios de contenido correspondientes a la evaluación del rendimiento en ciencias del TIMSS *Trends*.

EL PROYECTO PISA⁵

PISA es el acrónimo de *Programme for International Student Assessment*, un proyecto promovido por la *Organization for Economic Cooperation and Development* (OECD en inglés; OCDE en castellano)⁶ para evaluar el resultado de los sistemas educativos relativo a la formación de los alumnos necesaria para la vida adulta (figura 3). La evaluación se centra en la alfabetización⁷ lectora (año 2000), matemática (año 2003) y científica⁸ (año 2006)⁹. La iniciativa de esta evaluación comparativa transnacional

corresponde a los gobiernos de los países miembros de la OCDE y está diseñada para orientar sus políticas educativas.

En las diversas ocasiones en que se ha aplicado PISA hay cuestiones de los tres tipos de alfabetización señalados, pero cada nueve años se pone el acento en un área de conocimiento¹⁰, a la que se dedica dos tercios del tiempo total de las pruebas de evaluación (OECD, 1999, 2000). Esta actualización de resultados permite a los países participantes hacer un seguimiento cíclico de los progresos conseguidos en cada área.

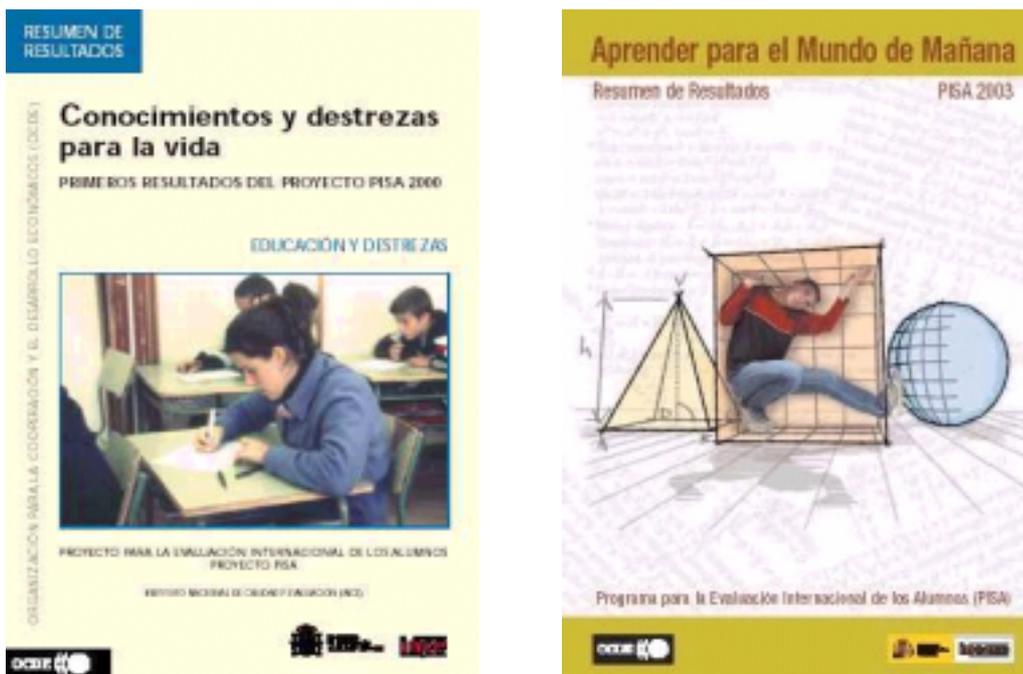


Figura 3.- Libros en castellano con los resultados PISA 2000 y 2003.

La evaluación del alumnado que está terminando su escolarización obligatoria es un indicador útil del rendimiento de los sistemas educativos. Aunque la mayor parte de los jóvenes de los países de la OCDE continúa su educación después de los 15 años, esta edad está próxima al final del período de escolarización obligatoria en el cual los estudiantes siguen, por lo general, un currículo básico común. En ese momento de su formación es conveniente determinar los conocimientos y destrezas que han adquirido para desenvolverse en la vida y para la elección de estudios posteriores. En cada país se evalúan entre 4500 y 10000 alumnos¹¹ de 15 años de edad¹². PISA ha evaluado durante el año 2003 a 276165 estudiantes, provenientes de 10104 centros, de un total de 42 países: los 30 miembros de la OCDE y otros 12 (OECD, 2003). España ha participado en las dos evaluaciones realizadas hasta ahora (años 2000 y 2003), obteniendo en ambas resultados mediocres (MEC, 2004; OECD, 2001, 2004a,b). Los resultados obtenidos en el año 2003 por España y las tres Comunidades Autónomas¹³ que participaron con una muestra ampliada se exhiben en la figura 4, donde también se indican los países con mejor y peor puntuación en cada área evaluada y la media correspondiente a los países de la OCDE.

Los resultados de las evaluaciones PISA, que se publican cada tres años junto a otros indicadores de los sistemas educativos, permiten a los administradores y gestores de la política educativa comparar el funcionamiento de sus sistemas educativos con los de los otros países y tomar decisiones en su ámbito nacional. También pueden ayudar a impulsar y orientar las reformas de la enseñanza y la mejora de los centros educativos, sobre todo en aquellos casos en que se obtienen peores resultados disponiendo de recursos similares. Además, estos resultados también son importantes para mejorar la evaluación y el seguimiento de la eficacia de los sistemas educativos.



Figura 4.- Comparación de resultados de la aplicación PISA 2003.

PISA no es el único estudio transnacional para comparar el rendimiento del alumnado. Como se ha mencionado en la sección anterior, durante los últimos 40 años se han realizado diversos estudios, la mayoría de ellos dirigidos por la IEA –como el TIMSS– y por la *International Assessment of Educational Progress* (IAEP) del *Educational Testing Service* (ETS). No obstante, estas otras evaluaciones se han centrado en resultados directamente relacionados con el currículo y, por tanto, sólo en aquellas partes de éste que son comunes a todos los países participantes. Por el contrario, el objetivo general del proyecto PISA es conocer la preparación de los estudiantes de 15 años de edad para afrontar los retos cotidianos de la vida adulta. Como PISA no es un proyecto rigurosamente vinculado al currículo, se facilita la comparación de resultados entre países, con independencia de la organización de las enseñanzas que éstos tengan.

El modelo de evaluación del PISA

Los marcos conceptuales¹⁴ del proyecto PISA definen cada una de las áreas que se evalúan –lectura, matemáticas y ciencias–, explican *qué* y *cómo* se evalúa y describen el contexto de las evaluaciones, incluyendo las posibles limitaciones que éste impone.

En PISA subyace la idea de un modelo dinámico de aprendizaje a lo largo de toda la vida para adaptarse con más éxito a las circunstancias personales en cada momento. Como es obvio, los estudiantes no pueden aprender en las escuelas todo lo que necesitarán para desenvolverse en la vida como adultos, pero sí deberían adquirir las capacidades necesarias –cognitivas y emotivas– para seguir aprendiendo en el futuro. Por este motivo, se considera deseable que los alumnos sean capaces de organizar y regular su propio aprendizaje, tanto individualmente como en grupo. Para ello, es preciso que lleguen a ser conscientes de sus propios procesos cognitivos, estrategias y métodos de aprendizaje. Así pues, la evaluación de la capacidad de autorregulación en el aprendizaje tiene mucha importancia en el proyecto PISA¹⁵ (OECD, 2003).

La evaluación PISA se ocupa de los procedimientos o destrezas, la comprensión de conceptos y la capacidad para responder a diversas situaciones reales en cada área evaluada. Sin embargo, los conocimientos y las destrezas no se definen, como ocurre en el TIMSS, a partir de los contenidos comunes de los diversos currículos nacionales, sino en función de lo que se considera necesario para la vida cotidiana. A pesar de que las tres áreas que evalúa se corresponden con asignaturas concretas, PISA va mucho más allá de los contenidos curriculares específicos, prestando atención sobre todo a los conocimientos y destrezas relevantes para la vida adulta (Acevedo, 2004)¹⁶.

Por otro lado, PISA también incluye la evaluación de ciertas capacidades transversales que no son específicas de una materia determinada, pero sí son esenciales para los alumnos¹⁷; por ejemplo, comunicación, adaptación, flexibilidad, resolución de problemas, uso de las tecnologías de la información, etc.

Los indicadores¹⁸ que proporcionan las evaluaciones del PISA están diseñados para lograr una mejor comprensión de hasta qué punto los sistemas educativos de los países participantes prepararan a sus alumnos para continuar aprendiendo a lo largo de sus vidas y puedan tener un papel activo en la sociedad como ciudadanos. Con otras palabras, PISA pretende más la evaluación del rendimiento de los sistemas educativos respecto a sus objetivos subyacentes –tal como los define la sociedad– que con referencia en la enseñanza y el aprendizaje de un cuerpo disciplinar de conocimientos. Esta forma de entender la evaluación es necesaria para estimular a los sistemas educativos y a los centros a enfrentarse a los retos educativos actuales, por lo que es, sin duda, la característica más importante y novedosa del proyecto PISA.

Las preguntas de ciencias del PISA

El tipo de preguntas del proyecto PISA es, en general, distinto al de otros proyectos de evaluación internacional como el TIMSS, los cuales usan sobre todo cuestiones cortas de elección múltiple basadas en conocimientos curriculares. Por el contrario, muchas preguntas del PISA requieren la combinación de diferentes tipos de conocimientos y destrezas y, a veces, la valoración de diversas decisiones para las que no hay una única respuesta correcta.

Algunas cuestiones del proyecto PISA son cerradas, en las que los alumnos sólo tienen que elegir entre varias opciones o dar respuestas relativamente simples que pueden compararse directamente con una respuesta correcta. Otras son preguntas abiertas, en las cuales los alumnos tienen que dar respuestas mucho más elaboradas; de esta

forma, se pueden evaluar aspectos más amplios que los que suelen recogerse en otros estudios más convencionales. Una importante innovación de PISA es la evaluación de destrezas de orden superior, lo que suele hacerse mediante preguntas abiertas.

De acuerdo con la definición de la alfabetización científica propuesta en PISA, se establecen tres grandes dimensiones para la evaluación (véase la tabla 3):

1. *Destrezas*. Los procedimientos científicos necesarios para la resolución de una pregunta o un problema¹⁹.
2. *Conceptos*. El conocimiento científico y la comprensión conceptual que se requieren para usar los procedimientos anteriores²⁰.
3. *Áreas de aplicación del conocimiento*. El marco teórico de referencia de los casos o situaciones del mundo real²¹ en los que se aplican las destrezas y los conceptos.

DESTREZAS	CONCEPTOS	ÁREAS DE APLICACIÓN
1. Descripción, explicación y predicción de fenómenos científicos. 2. Comprensión de la investigación científica. 3. Interpretación de pruebas y conclusiones científicas.	1. Estructura y propiedades de la materia. 2. Cambio atmosférico. 3. Cambios físicos y químicos. 4. Transformaciones de la energía. 5. Fuerzas y movimiento. 6. Función y forma. 7. Biología humana. 8. Cambio fisiológico. 9. Biodiversidad. 10. Control genético. 11. Ecosistemas. 12. La Tierra y su lugar en el universo. 13. Cambio geológico.	1. Ciencias de la vida y de la salud: (i) Salud, enfermedad y nutrición. (ii) Preservación y uso sostenible de las especies. (iii) Interdependencia de los sistemas físicos y biológicos. 2. Ciencias de la Tierra y del medio ambiente: (i) Contaminación. (ii) Producción y pérdida de suelo. (iii) Tiempo y clima. 3. Ciencias en las tecnologías: (i) Biotecnología. (ii) Uso de materiales y recogida de residuos. (iii) Utilización de la energía. (iv) Transporte.

Tabla 3.- Dimensiones del PISA para la evaluación de la alfabetización científica.

En la mayor parte de los casos, las cuestiones de la evaluación se organizan en grupos de preguntas basados en un texto relativo a un mismo asunto, problema o tema, el cual refleja una situación de la vida real. Este formato permite emplear cuestiones que profundizan más en el tema que si cada pregunta introdujera un nuevo contexto. Así, el alumno tiene tiempo suficiente para asimilar el material que se emplea para evaluar distintos aspectos del rendimiento.

En la tabla 4 se muestra una cuestión abierta de ciencias utilizada en la prueba piloto del proyecto PISA realizada antes de su aplicación del año 2000²², así como algunas

respuestas que dan los estudiantes y la puntuación otorgada a cada una de ellas en la evaluación realizada²³.

La cuestión presenta una posición que defienden algunas personas: "los trolebuses no contribuyen a la contaminación del aire". Los alumnos tienen que evaluar la validez de esta afirmación utilizando la información que se suministra en el propio texto y sus conocimientos sobre los productos desprendidos en la combustión del carbón en las centrales eléctricas (OECD, 2000, p. 81; p. 122 de la traducción castellana, 2001). La destreza necesaria para ello es la capacidad para obtener conclusiones válidas, el concepto científico se refiere a transformaciones de la energía y el área de aplicación del conocimiento es la contaminación, correspondiente a "Ciencias de la Tierra y del medio ambiente" (véanse en la tabla 3). La situación del mundo real descrita en la pregunta es la contaminación atmosférica provocada por los autobuses, un asunto controvertido sobre el que los estudiantes deben tomar decisiones razonadas.

Ejemplo de pregunta abierta
<p>El autobús de Julián, como la mayoría de los autobuses, funciona con un motor diesel. Estos autobuses contribuyen a la contaminación del medio ambiente.</p> <p>Un compañero de Julián trabaja en una ciudad donde se usan trolebuses que funcionan con un motor eléctrico. El voltaje necesario para este tipo de motores eléctricos es suministrado por cables eléctricos (como en los trenes eléctricos). La electricidad procede de una central que utiliza carbón. Los partidarios del uso de trolebuses en la ciudad argumentan que este tipo de transporte no contribuye a la contaminación del aire.</p> <p>¿Tienen razón los partidarios del trolebús? Explica tu respuesta.</p>
Ejemplos de respuestas dadas por los estudiantes (entre paréntesis se indica la puntuación otorgada)
<p><i>"Sí y no. Los trolebuses no contaminan la ciudad, lo que es bueno, pero las centrales eléctricas contaminan y eso no es muy bueno."</i> (1)</p> <p><i>"Los trolebuses contribuyen a la contaminación del aire por usar combustibles fósiles, pero no tanto como los autobuses normales con todos sus gases."</i> (1)</p> <p><i>"Bien, no tienen problema hasta que el humo dañino no vaya al aire y perjudique la capa de ozono, y tener electricidad a partir de los combustibles fósiles es mucho más cuidadoso con el medio ambiente."</i> (0)</p> <p><i>"Sí, tienen razón ya que la electricidad no es perjudicial para el medio ambiente, sólo consumimos el gas terrestre."</i> (0)</p> <p><i>"Pienso que los partidarios del trolebús tienen razón ya que los autobuses de diesel contaminan más que los trolebuses que son eléctricos."</i> (0)</p> <p><i>"Sí, porque si se quema carbón los gases no contaminantes se liberarán."</i> (0)</p>

Tabla 4.- Pregunta abierta de ciencias del PISA.

Los cuestionarios de contexto del proyecto PISA

Con el fin de encontrar algunas de las principales características asociadas al éxito escolar, el proyecto PISA tiene en cuenta el rendimiento de los alumnos junto con las características de sus centros de enseñanza y sus entornos familiares. Para ello, los alumnos y los directores de los centros educativos responden a cuestionarios sobre su entorno²⁴. Estos cuestionarios no son un simple anexo del proyecto PISA, sino que permiten recopilar información del contexto, la cual es muy útil para el análisis de los resultados en función de las características de los alumnos y del centro educativo²⁵. También se incluye un cuestionario para obtener información de la capacidad de autorregulación del aprendizaje²⁶ por parte de los alumnos²⁷.

La información contextual recogida en los cuestionarios del alumno y del director del centro educativo abarca sólo una parte de toda la información de este tipo disponible en el proyecto PISA. Los indicadores relativos a la estructura general de los sistemas educativos –sus contextos demográficos y económicos, tales como costes, matrículas, características de los centros educativos y del profesorado...– y al funcionamiento del mercado laboral los desarrolla la OCDE regularmente por otra vía.

En conjunto, los cuestionarios de contexto de PISA proporcionan una completa y detallada base de datos para el análisis, orientado a la política educativa, de los resultados de la evaluación realizada. Junto con la información obtenida mediante otros programas e instrumentos de la OCDE, permiten:

- Comparar diferencias entre los resultados de los alumnos de distintos sistemas educativos y con diversos entornos académicos.
- Comparar diferencias entre los resultados de los alumnos que estudian distintos contenidos curriculares y siguen diversos procedimientos pedagógicos.
- Considerar posibles relaciones entre el rendimiento de los alumnos y los factores organizativos, tales como el tamaño de los centros educativos y los recursos disponibles, así como las diferencias entre países en estos aspectos.
- Examinar diferencias entre países respecto al grado en que los centros educativos reducen o incrementan los efectos de los factores contextuales que influyen en el nivel de rendimiento del alumnado.
- Considerar diferencias entre los sistemas educativos y el contexto nacional respecto a las diferencias en el rendimiento de los estudiantes dentro de cada país.

REFLEXIONES FINALES

Si se desea hacer con cierta objetividad, la medición absoluta del rendimiento de un sistema educativo es prácticamente una utopía, pues aún son muchos los factores desconocidos que pueden incidir en ello. Por tanto, no se debería minusvalorar la importancia de los estudios comparativos de evaluación entre países, porque se trata de una medición relativa más factible. No obstante, las evaluaciones transnacionales comparativas han recibido muchas críticas por sus marcos teóricos y diversos aspectos de la metodología empleada, los cuales condicionan sus resultados y pueden producir algunos efectos indeseables. Por ejemplo, la diversidad de currículos de los países

participantes restringió mucho el diseño de las pruebas y cuestiones del TIMSS aplicado en 1995, limitándolas sobre todo a los contenidos más convencionales, con un elevado porcentaje de preguntas de recuerdo de información factual y conceptual presentadas con un formato de opción múltiple, lo cual reforzó en muchos países la creencia de que era necesario insistir más en los aprendizajes tradicionales (Sjøberg, 1999; Fensham, 1999, 2000, 2004). Además, la publicación de la tabla de resultados de esa aplicación del TIMSS sirvió más para fortalecer en muchos países participantes la enseñanza propedéutica de las ciencias que a afrontar los defectos encontrados (Vázquez, 2000). TIMSS *Trends* ha hecho un importante esfuerzo por mejorar algunos de esos defectos, pero conserva parte de ellos porque sigue siendo un estudio de evaluación del currículo común a países muy distintos.

Para intentar superar estas limitaciones, PISA presenta un enfoque diferente basado en el concepto de alfabetización para la vida adulta, pretendiendo la evaluación de otros aprendizajes que van mucho más allá de los conocimientos factuales y conceptuales. Como se ha mostrado aquí, las cuestiones del PISA tienen un formato distinto al habitual de los exámenes escolares, con más preguntas abiertas basadas en un texto relativo a una situación real. El diseño de la evaluación PISA correspondiente al año 2006 también pretende prestar atención a las actitudes relacionadas con la ciencia y el medio ambiente, incluyendo así los aspectos emotivos y afectivos, que no se consideran sólo un complemento informativo de la evaluación, como en el TIMSS, sino una importante dimensión de la alfabetización científica (Fensham, 2004). Sin duda, se trata de una importante novedad, que realza la educación de actitudes como objetivo valioso del aprendizaje de las ciencias, lo cual viene siendo reclamado desde hace tiempo por el movimiento CTS –Ciencia, Tecnología y Sociedad– para la enseñanza de las ciencias (Acevedo, Vázquez y Manassero, 2002, 2003). Sin embargo, esto puede suponer una seria dificultad metodológica, porque la evaluación se hace mediante una encuesta de opinión –probablemente usando una escala de Likert o similar– y, por tanto, no es del mismo tipo que la correspondiente a las demás cuestiones. Estas preguntas actitudinales han sido incluidas en las pruebas piloto que se han realizado recientemente antes del verano. El análisis de los resultados de estas pruebas permitirá decidir la conveniencia o no de mantenerlas en la prueba definitiva a celebrar en el año 2006.

Por último, como recomendación general, sería muy deseable diferenciar entre las conclusiones que, de modo razonable, pueden extraerse de los datos suministrados por los informes derivados de las aplicaciones de los proyectos comparativos transnacionales de evaluación y las interpretaciones que se hacen de ellos, en particular de las relaciones causa-efecto. Estas interpretaciones y atribuciones de causalidad pueden tener una gran carga ideológica y una importante intencionalidad política, lo que aumenta el riesgo de oscurecer el debate y manipular las posteriores decisiones educativas que se tomen (Vázquez, 2000). Por tanto, parece necesario seguir analizando con más calma, detenimiento y rigor las conclusiones y sugerencias que proporcionan los proyectos de este tipo que se están llevando a cabo para comprobar sus valores potenciales para las reformas de la enseñanza de las ciencias (Vázquez, Acevedo y Manassero, 2005). El primer paso para ello debe ser conocer algunas características de esos proyectos, que es lo que se pretendía con este artículo.

REFERENCIAS BIBLIOGRÁFICAS

- ACEVEDO, J.A. (2004). Reflexiones sobre las finalidades de la enseñanza de las ciencias: Educación científica para la ciudadanía. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 1(1), 3-16, <http://www.apac-eureka.org/revista/Larevista.htm>.
- ACEVEDO, J.A., VÁZQUEZ, A. y MANASSERO, M.A. (2002). El movimiento Ciencia, Tecnología y Sociedad y la enseñanza de las ciencias. En *Sala de Lecturas CTS+I de la OEI*, <http://www.campus-oei.org/salactsi/acevedo13.htm>. Versión en castellano del capítulo 1 del libro de Manassero, M.A., Vázquez, A. y Acevedo, J.A.: *Avaluació dels temes de ciència, tecnologia i societat*. Palma de Mallorca: Conselleria d'Educació i Cultura del Govern de les Illes Balears, 2001.
- ACEVEDO, J.A., VÁZQUEZ, A. y MANASSERO, M.A. (2003). Papel de la educación CTS en una alfabetización científica y tecnológica para todas las personas. *Revista Electrónica de Enseñanza de las Ciencias*, 2(2), <http://www.saum.uvigo.es/reec/>.
- BEATON, A.E., MARTIN, M.O., MULLIS, I.V.S., GONZÁLEZ, E.J., SMITH, T.A. y KELLY, D.L. (1996). *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.
- BYBEE, R.W. (1997). *Achieving scientific literacy: From purposes to practices*. Portsmouth, NH: Heinemann.
- FENSHAM, P.J. (1999). What do the findings of TIMSS mean – science-wise and educationally? En M. Komorek, H. Behrendt, H. Dahncke, R. Duit, W. Gräber y A. Kros. (Eds.): *Research in science education. Past, present, and future. Proceedings Second International Conference of the European Science Education Research Education (ESERA)*, pp. 609-611. Kiel: IPN, University of Kiel.
- FENSHAM, P.J. (2000). International success, but is it Science? – Identifying strengths and weaknesses in Australian primary school science from TIMSS and other data. *Australian Science Teachers Association Journal*, 45(2), 39-44.
- FENSHAM, P.J. (2004). Beyond Knowledge: Other Scientific Qualities as Outcomes for School Science Education. En R.M. Janiuk y E. Samonek-Miciuk (Ed.), *Science and Technology Education for a Diverse World – dilemmas, needs and partnerships*. International Organization for Science and Technology Education (IOSTE) XIth Symposium Proceedings, pp. 23-25. Lublin, Poland: Maria Curie-Sklodowska University Press.
- HARLEN, W. (2001). The assessment of scientific literacy in the OECD/PISA project. *Studies in Science Education*, 36, 79-104.
- HARLEN, W. (2002). Evaluar la alfabetización científica en el programa de la OCDE para la evaluación internacional de estudiantes (PISA). *Enseñanza de las Ciencias*, 20(2), 209-216.
- LÓPEZ-VARONA, J.A. y MORENO-MARTÍNEZ, M.L. (1996). Tercer estudio internacional de Matemáticas y Ciencias (TIMSS). *Revista de Educación*, 311, 315-336.

- LÓPEZ-VARONA, J.A. y MORENO-MARTÍNEZ, M.L. (1997). *Resultados de Ciencias. Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS)*. Madrid: INCE/MECD.
- MEC (2004). *Evaluación PISA 2003. Resumen de los primeros resultados en España*. INECSE/MEC.
- MARTIN, M.O., MULLIS, I.V.S., BEATON, A.E., GONZÁLEZ, E.J., SMITH, T.A. y KELLY, D.L. (1997). *Science Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.
- MARTIN, M.O., MULLIS, I.V.S., GONZÁLEZ, E.J. y CHROSTOWSKI, S.J. (2004). *TIMSS 2003 International Science Reports. Findings From IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- MARTIN, M.O., MULLIS, I.V.S., GONZÁLEZ, E.J., GREGORY, K.D., SMITH, T.A., CHROSTOWSKI, S.J., GARDEN, R.A. y O'CONNOR, K.M. (2000). *TIMSS 1999 International Science Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. Chestnut Hill, MA: Boston College.
- MARTÍN-DÍAZ, M.J. (2004). El papel de las ciencias de la naturaleza en la educación a debate. *Revista Iberoamericana de Educación*, edición electrónica, <http://www.campus-oei.org/revista/deloslectores/692MartinDiaz.PDF>.
- MILLAR, R. y OSBORNE, J., Eds. (1998). *Beyond 2000. Science education for the future*. London: King's College London School of Education.
- MULLIS, I.V.S., MARTIN, M.O., BEATON, A.E., GONZÁLEZ, E.J., KELLY, D.L. y SMITH, T.A. (1998). *Mathematics and Science Achievement in the Final Year of Secondary School: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.
- MULLIS, I.V.S., MARTIN, M.O., SMITH, T.A., GARDEN, R.A., GREGORY, K.D., GONZÁLEZ, E.J., CHROSTOWSKI, S.J. y O'CONNOR, K.M. (2002). *TIMSS Assessment Frameworks and Specifications 2003*. Chestnut Hill, MA: Boston College. Traducción de M. Angstadt (2002), *Marcos teóricos y especificaciones de evaluación de TIMSS 2003*. Madrid: INCE/MECD.
- NIEDA, J., CAÑAS, A. y MARTÍN-DÍAZ, M.J. (2004). *Actividades para evaluar las Ciencias en secundaria*. Madrid: Visor – Cátedra UNESCO.
- OECD (1999). *Measuring student knowledge and skills. A new framework for assessment*. París: OECD. Traducción de G. Gil Escudero, J. Fernández García, F. Rubio Miguelsanz y C. López Ramos (2000), *La medida de los conocimientos y las destrezas de los alumnos. Un nuevo marco de evaluación*. Madrid: INCE/MECD.
- OECD (2000). *Measuring student knowledge and skills: The PISA assessment of reading, mathematical and scientific literacy*. París: OECD. Traducción de G. Gil Escudero, J. Fernández García, F. Rubio Miguelsanz, C. López Ramos y S. Sánchez Robles (2001), *La medida de los conocimientos y las destrezas de los alumnos: La evaluación de la lectura, las matemáticas y las ciencias en el proyecto PISA 2000*. Madrid: INCE/MECD.

- OECD (2001). *Knowledge and skills for life: First results from PISA 2000. Executive Summary*. París: OECD. Traducción de G. Gil Escudero (2001), *Conocimientos y destrezas para la vida: primeros resultados del proyecto PISA 2000. Resumen de resultados*. Madrid: INCE/MECD.
- OECD (2003). *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*. París: OECD. Traducción de E. Belmonte (2004): *Marcos teóricos de PISA 2003: la medida de los conocimientos y destrezas en matemáticas, lectura, ciencias y resolución de problemas*. Madrid: INECSE/MEC.
- OECD (2004a). *First results from PISA 2003: Executive Summary*. París: OECD. Traducción de E. Belmonte (2004), *Aprender para el mundo de mañana. Resumen de resultados PISA 2003*. Madrid: INECSE/MEC.
- OECD (2004b). *Learning for Tomorrow's World: First Results from PISA 2003*. París: OECD.
- OLIVA, J.M. y ACEVEDO, J.A. (2005). La enseñanza de las ciencias en primaria y secundaria hoy. Algunas propuestas de futuro. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 2(2), 142-151, <http://www.apac-eureka.org/revista/Larevista.htm>.
- PILOT, A. (2000). The concept of "basic scientific knowledge" through some of the reforms recently undertaken in science and technology teaching in European States. En M. Poisson (Ed.), *Science education for contemporary society: problems, issues and dilemmas*. Final report of the International Workshop on The reform in the teaching of science and technology at primary and secondary level in Asia: Comparative references to Europe. Part IV: *New approaches in science and technology education*, pp. 104-110. Beijing, China (27-31 March 2000). International Bureau of Education, The Chinese National Commission for UNESCO. En <http://www.ibe.unesco.org/National/China/chifinal.htm>.
- ROBITAILLE, D.F. y GARDEN, R.A., Eds. (1996). *TIMSS Monograph No. 2: Research Questions y Study Design*. Vancouver, Canadá: Pacific Educational Press.
- SJØBERG, S. (1999). TIMSS and consequences drawn in Australia, Austria and Germany. Paper presented at the 2nd Conference of the European Science Education Research Association (ESERA): *Research in Science Education. Past, present, and future*. Kiel: IPN, University of Kiel.
- SJØBERG, S. (2004). Science Education: The voice of the learners. Contribution to the Conference on Increasing Human Resources for Science and Technology in Europe. Bruselas, Unión Europea (2 de abril de 2004). En <http://europa.eu.int/comm/research/conferences/2004/sciprof/pdf/sjoberg.pdf>.
- VÁZQUEZ, A. (2000). Análisis de los datos del tercer estudio internacional de Matemáticas y Ciencias (TIMSS) desde la perspectiva del sistema educativo español. Memoria de investigación. Madrid: MEC/CIDE.
- VÁZQUEZ, A., ACEVEDO, J.A. y MANASSERO, M.A. (2005). Más allá de una enseñanza de las ciencias para científicos: hacia una educación científica humanística (enviado a publicación).

NOTAS

¹ La web del proyecto TIMSS es <http://timss.bc.edu/>. Algunas de las referencias que se citan en castellano sobre este proyecto se pueden descargar en la web del Instituto Nacional de Evaluación y Calidad del Sistema Educativo (INECSE): <http://www.ince.mec.es/index.htm>.

² Desde hace más de 40 años, la IEA se ha ocupado de hacer estudios sobre el rendimiento de los estudiantes en matemáticas y ciencias y recopilar material contextual para la mejora del aprendizaje en estas materias, cumpliendo así su objetivo de hacer avanzar la educación científica y matemática. El primer estudio internacional sobre matemáticas –*First International Mathematics Study (FIMS)*– se realizó en 1964. Las ciencias se evaluaron por primera vez como parte del un estudio del rendimiento en seis materias en 1970-71. Las matemáticas y las ciencias volvieron a ser el centro de evaluaciones de este tipo en 1980-82 y en 1983-84, respectivamente.

³ La IEA fue fundada en 1959 para realizar estudios comparativos de investigación sobre políticas, prácticas y resultados educativos. Hoy día están asociadas a ella alrededor de 60 países. El Centro de Estudios Internacionales de la IEA tiene su sede en el Boston College (EE.UU.), pero también hay una secretaria permanente en Amsterdam (Holanda) y un centro de proceso de datos en Hamburgo (Alemania).

⁴ En 4º curso, Física y Química se evalúa como una única área de contenido que incluye temas de ambas materias. En el mismo curso, algunos temas relativos a las Ciencias medioambientales se evalúan como parte de los correspondientes a Ciencias de la vida y Ciencias de la Tierra.

⁵ La web del proyecto PISA es <http://www.pisa.oecd.org/>. Algunas de las referencias que se citan en castellano sobre este proyecto se pueden descargar en la web del Instituto Nacional de Evaluación y Calidad del Sistema Educativo (INECSE): <http://www.ince.mec.es/index.htm>.

⁶ El diseño y la ejecución de estos estudios se llevan a cabo bajo la dirección del Secretariado de la OCDE por medio de un consorcio internacional encabezado por el *Australian Council for Educational Research (ACER)*. Este organismo produce los indicadores, realiza los análisis y prepara los informes y publicaciones internacionales en contacto directo con los países miembros. Así mismo, es el responsable general de la gestión del proyecto PISA y del control de su puesta en práctica.

⁷ Traducción literal del término anglosajón *literacy*. Sin embargo, en el proyecto PISA su sentido va mucho más allá de lo que se entiende estrictamente por alfabetización e, incluso, por alfabetización funcional. De manera general, en PISA la alfabetización se refiere a la formación acumulada que es útil para enfrentarse a la vida real.

⁸ El proyecto PISA define la alfabetización científica como “*la capacidad de emplear el conocimiento científico para identificar cuestiones y obtener conclusiones a partir de pruebas, con el fin de comprender y ayudar a tomar decisiones acerca del mundo natural y de los cambios artificiales que la actividad humana produce en él.*” (OECD, 2000, p. 76; p. 115 de la traducción castellana, 2001). Esta forma de entender la alfabetización científica se corresponde aproximadamente con el tercero de los cuatro niveles del modelo continuo desarrollado por Bybee (1997) para explicarla (OECD, 2003), nivel que se denomina alfabetización científica conceptual y procedimental (Acevedo, Vázquez y Manassero, 2003).

⁹ En esa ocasión, PISA prestará especial atención a la evaluación de la alfabetización científica. Los resultados se harán públicos en el año 2007.

¹⁰ Aunque este período de tiempo puede parecer bastante dilatado, hay que tener en cuenta que los cambios en educación son lentos.

¹¹ Como norma general, las especificaciones internacionales de muestreo exigen un mínimo de 4500 alumnos y 150 centros por país.

¹² Como la evaluación PISA no está estrictamente centrada en el currículo, sino en la formación para la vida adulta, la muestra de este estudio no selecciona alumnos de un curso determinado, sino aquellos que tienen cumplidos 15 años de edad en el momento de la aplicación de las pruebas. Con ello se salva la dificultad de lograr equivalencias válidas entre cursos académicos, debida a las diferencias organizativas de los distintos sistemas educativos de los países participantes.

¹³ Castilla y León, Cataluña y País Vasco han participado en la aplicación del proyecto PISA 2003 con una muestra ampliada respecto a la de otras Comunidades Autónomas de España, lo que ha permitido la posterior presentación separada de sus resultados. En la próxima aplicación del año 2006 pretenden participar 58 países y hasta 9 Comunidades Autónomas españolas ampliarán su muestra para poder obtener resultados propios. Además de las tres que participaron de esta manera en PISA 2003, se incorporarán a ello Andalucía, Aragón, Asturias, Galicia, La Rioja y Navarra.

¹⁴ Los marcos conceptuales del proyecto PISA han sido desarrollados por grupos de expertos bajo la dirección de Raymond Adams, del ACER. El grupo de expertos de ciencias, ha estado a cargo de la profesora Wynne Harlen, del *Scottish Council for Research in Education (SCR)*.

¹⁵ Mientras que esta capacidad se evaluó de manera opcional en la aplicación PISA del año 2000, ha sido un objetivo central en la correspondiente al año 2003.

¹⁶ Por ejemplo, para favorecer la participación ciudadana en los debates tecnocientíficos, la comprensión de temas y conceptos generales, como el consumo de energía, la biodiversidad o la salud y la alimentación, se considera mucho más útil que otros contenidos convencionales.

¹⁷ En el año 2003 se ha evaluado la solución de problemas como área transversal. Sin embargo, la evaluación de destrezas y conocimientos transversales de distinto tipo no es periódica, sino ocasional.

¹⁸ Hay tres tipos de indicadores: (i) *básicos*, que proporcionan un perfil de los conocimientos y destrezas de los estudiantes de 15 años; (ii) *contextuales*, que muestran la relación de esos conocimientos y destrezas con las características de los centros y las variables demográficas, sociales, económicas y educativas más importantes de cada país; y (iii) *de tendencias de los resultados educativos*, que reflejan los cambios producidos en esos resultados durante los ciclos de tiempo establecidos.

¹⁹ Los procedimientos científicos cubren un amplio conjunto de destrezas y saberes necesarios para recopilar e interpretar los hechos del mundo natural y extraer conclusiones de ellos. Puesto que la gran mayoría de los estudiantes no serán científicos en el futuro, PISA da prioridad a los procedimientos sobre cómo funciona la ciencia en vez de a los procedimientos que se emplean en la ciencia. Así, los ciudadanos necesitan saber cuándo resulta pertinente usar el conocimiento científico y discernir entre las preguntas que la ciencia puede responder y aquellas a las que no puede dar respuesta. También tienen que ser capaces de juzgar cuándo una prueba científica es válida, tanto en lo relativo a su pertinencia como al modo en que se ha obtenido. Y, quizás lo más importante de todo, necesitan ser capaces de relacionar los hechos y los datos con las conclusiones extraídas de ellos, así como sopesar las pruebas a favor y en contra de determinadas acciones que pueden afectar a su vida personal o familiar, a la comunidad local o al planeta. En definitiva, este enfoque del currículo de ciencias no se centra en cómo "hacer ciencia" ni en cómo "crear conocimientos científicos", algo que debería reservarse a los pocos alumnos que serán científicos en el futuro, sino en la capacidad de evaluar las pruebas aportadas, distinguir entre teorías y observaciones, y valorar el grado de confianza que hay que conceder a las explicaciones proporcionadas (Millar y Osborne, 1998).

²⁰ Los conceptos científicos seleccionados en PISA se expresan como grandes ideas integradoras que ayudan a entender algunos aspectos del mundo natural.

²¹ Las situaciones del mundo real empleadas en las preguntas de ciencias del proyecto PISA corresponden a asuntos relevantes desde un punto de vista (i) personal o familiar (*v.g.*, cuestiones de salud, alimentación, consumo...), (ii) local (*v.g.*, instalación de una central eléctrica o un vertedero industrial, abastecimiento y depuración de aguas, transporte público...), (iii) mundial (*v.g.*, calentamiento global del planeta, disminución de la biodiversidad...) y (iv) histórico (*v.g.*, influencia social en la ciencia y de la ciencia en la sociedad, evolución del conocimiento científico...).

²² Pueden consultarse más ejemplos de preguntas de ciencias y las puntuaciones asignadas a las respuestas en OECD (2000, 2003).

²³ Para conseguir puntuar en la respuesta debe mencionarse la contaminación causada por la central eléctrica de carbón que produce la electricidad.

²⁴ Estos cuestionarios permiten obtener información sobre: (i) aspectos socioeconómicos y culturales de las familias de los alumnos y su implicación en la educación, (ii) las costumbres y actitudes de los alumnos en el ambiente familiar y en el centro educativo, y (iii) la titularidad pública o privada, organización, estructura y recursos de los centros educativos.

²⁵ En comparación con TIMSS, PISA aporta muy pocos datos sobre el profesorado, lo que constituye una desventaja del proyecto. Los responsables del diseño de PISA justifican la ausencia de cuestionarios específicos para el profesorado por el hecho de que no se evalúan los conocimientos y las destrezas curriculares del alumnado de un curso determinado, sino las acumuladas hasta los 15 años de edad.

²⁶ Esta información se refiere a: (i) motivación para el aprendizaje, (ii) autoconcepto, (iii) influencia de los factores emocionales, (iv) capacidad y esfuerzo para superar las dificultades de aprendizaje, (v) estrategias de control de las actividades de aprendizaje, (vi) preferencias por diferentes tipos de situaciones de aprendizaje, (vii) estilos de aprendizaje preferidos, (viii) estrategias usadas para sistematizar y analizar la información, y (ix) destrezas sociales para el aprendizaje cooperativo en grupo.

²⁷ En la aplicación del año 2003, se incluyeron dos cuestionarios opcionales para el alumnado. El primero sobre la familiaridad con las TIC (disponibilidad, utilización, actitud, confianza y formación en las TIC). El segundo acerca de la formación educativa (antecedentes escolares, formación actual, formación prevista y puesto de trabajo que se espera conseguir en el futuro).