

La evaluación PISA 2006 en España



Ramón Pajares Box

*Jefe del área del INECSE y coordinador nacional
del estudio PISA en España*

La evaluación PISA 2006 en España

Con una nueva ley de educación recién aprobada y su posterior desarrollo normativo y curricular en plena efervescencia, la aplicación en España de una evaluación internacional de Ciencias es un acontecimiento ciertamente oportuno. Un ejercicio de comparación internacional ayudará a perfilar mejor las alternativas y sugerirá posibles vías de avance ante la serie de retos que la enseñanza de las Ciencias se verá obligada a afrontar. Los siguientes párrafos están dedicados a examinar cómo aborda la enseñanza de las Ciencias una evaluación internacional como PISA, cuyos rasgos fundamentales recordamos muy rápidamente.

1. ¿Qué es PISA?

PISA son las siglas del estudio de la OCDE: *Programme for International Student Assessment* (Programa para la Evaluación Internacional de Alumnos). Descrito muy brevemente, se trata de un estudio internacional de evaluación del rendimiento educativo de los alumnos de 15 años, aplicado del modo más similar posible en los más de cincuenta países participantes con el fin de realizar una comparación rigurosa del rendimiento de los distintos sistemas educativos. Es un estudio impulsado por la OCDE, organización de ayuda a los gobiernos en la formulación y optimización de sus decisiones políticas. En consecuencia, PISA se realiza con la finalidad de obtener

datos y enseñanzas para la política educativa y no es, por tanto, un estudio orientado al micronivel educativo de centros docentes y de procesos de enseñanza-aprendizaje, sino al macronivel de la definición y formulación de políticas educativas de largo alcance.

Los estudios PISA se aplican cada tres años. En cada aplicación se estudian los rendimientos del alumnado en tres materias: Lectura, Matemáticas y Ciencias, pero una de ellas, de forma rotatoria, recibe una atención más profunda, mientras que con las otras dos solo se realiza un sucinto sondeo¹. El primer estudio PISA, cuya toma principal de datos se realizó en el año 2000, tuvo como materia principal la Lectura. PISA 2003 tuvo como materia principal las Matemáticas y el presente estudio PISA 2006, cuya recogida de datos en España finaliza en estos días, está centrado en las Ciencias y a ellas se dedica unos dos tercios del tiempo de la prueba.

La recogida de datos consiste en que los alumnos contesten a una prueba de papel y lápiz² de dos horas de duración, seguida de un cuestionario de datos personales y

- (1) En PISA 2003 se añadió con carácter experimental una materia transversal: la Solución de problemas. Esta materia no volverá a ser evaluada en las aplicaciones PISA futuras.
- (2) Las pruebas de papel y lápiz se consideran una limitación y se están explorando otras vías de recogida de datos. En esta evaluación PISA 2006 algunos países –pero no España– ensayan la recogida de datos por ordenador. Este medio favorece unos estímulos más ricos en las preguntas,

familiares junto a opiniones y actitudes de los alumnos ante los estudios y el centro escolar. El cuestionario dura una media hora. El director del centro también contesta un cuestionario de unos 20 minutos, pero no hay previsto ningún cuestionario para los profesores. Se han preparado 13 cuadernos de examen distintos, con el contenido parcialmente solapado según la técnica del muestreo matricial. El conjunto completo de preguntas requeriría un tiempo de examen de 7 horas: el muestreo matricial permite que puedan aplicarse todas las preguntas a subgrupos distintos de alumnos sin que el tiempo de examen individual exceda de dos horas.

Existen varios tipos de preguntas³. Por un lado, hay preguntas cerradas en dos

que se pueden beneficiar de las capacidades multimedia presentes en los ordenadores actuales, y resultan más motivadores para los jóvenes alumnos, normalmente muy enterados de todas las facetas informáticas. Pero el principal beneficio futuro de una prueba por ordenador estaría en su capacidad de adaptación sobre la marcha a la capacidad demostrada por el alumno, de modo que el grado de acierto a las preguntas anteriores podría influir en la selección de preguntas sucesivas, explorando así mejor la capacidad individual del alumno que responde.

- (3) Al finalizar cada aplicación PISA, un cierto número de preguntas se hacen públicas, mientras que el resto se reservan para futuras aplicaciones, pues son la base de la medición de tendencias entre sucesivas ediciones de este estudio. En las páginas web del INECSE (<http://www.ince.mec.es/pub/>) pueden encontrarse distintas recopilaciones en castellano de las preguntas liberadas hasta el momento en Lectura, Matemáticas y Ciencias.

versiones: simples –o de elección múltiple entre cuatro opciones de respuesta distintas– y complejas, que combinan diferentes opciones de respuesta binaria (sí/no, verdadero/falso). También hay preguntas abiertas, de respuesta corta o larga, que necesitan la intervención de correctores, cuidadosamente entrenados y equipados con pormenorizados criterios de valoración. Por último, las preguntas sobre actitudes utilizan escalas tipo Likert.

Es específico de PISA que las preguntas no se presenten nunca aisladas, sino formando unidades, esto es, pequeñas agrupaciones de preguntas bajo una introducción textual y/o gráfica común, llamada «estímulo». Esto permite minimizar los cambios de contexto de las preguntas y también permite que el estímulo común de la unidad pueda ser explorado o aprovechado mejor. Este estímulo común trata de presentar las preguntas en un contexto «auténtico», es decir, realista y cotidiano, algo no tan habitual en la práctica docente de nuestras aulas.

Es también característico de PISA el que no se pretenda evaluar los contenidos previamente definidos en el currículo. La experiencia de la IEA, la otra gran agencia internacional de evaluación comparada, ha mostrado que es muy difícil encontrar el factor común a la disparidad curricular existente en los países participantes. Por ello PISA ha renunciado desde su comienzo a seleccionar a los alumnos por el curso académico en el que están escolarizados y se ha decantado por la edad biológica como

criterio de selección de su población objetivo. La edad elegida, 15 años, tiene la virtud de coincidir, en la mayoría de los países, con el último curso de la escolaridad obligatoria.

2. ¿Evaluación genérica o evaluación específica?

El contenido de las pruebas no corresponde, pues, a ningún currículo, sino al juicio de un grupo internacional de expertos que las formulan de acuerdo con el acervo de conocimientos y destrezas esperables en un alumno de la edad referida. Un acierto, a mi juicio, de PISA es que sus expertos buscan en los alumnos las capacidades que les permitirán integrarse a corto plazo en la vida adulta, y no las de integrarse a corto plazo en estudios más avanzados. Este sesgo poco academicista de PISA ha sido objeto de críticas desde concepciones tradicionales de la educación, pero para otras concepciones más actualizadas ha supuesto un estímulo adicional para la renovación de nuestra enseñanza secundaria.

De hecho, también supone un cierto estímulo renovador en nuestro contexto el estilo PISA de hacer preguntas. La búsqueda del elemento «auténtico» las hace alejarse de la práctica machacona del ejercicio o procedimiento rutinario, o de la búsqueda de la repetición de lo recién explicado en clase. Una de las consecuencias de este estilo de hacer preguntas es que, con mucha frecuencia, se suele encontrar la respuesta buscada en el enunciado de la pregunta o en el texto del estímulo común de la

unidad, con lo que el alumno tiene que aportar pocos conocimientos previos para proporcionar una respuesta correcta. Esto tiende a convertir a la mayor parte de las preguntas de PISA en ejercicios de lectura, que miden unas competencias generales o transversales más que unas competencias específicas de la materia objeto de examen.

En efecto, un repaso a las correlaciones entre las puntuaciones alcanzadas por los alumnos españoles en las diversas materias examinadas en PISA 2003 muestra que las pruebas de PISA parecen medir más una capacidad subyacente común y no capacidades específicas propias de cada asignatura, puesto que los índices de correlación son realmente muy altos para una investigación de ciencias sociales.

Correlaciones de las puntuaciones obtenidas por los alumnos españoles entre las distintas materias evaluadas en PISA 2003

Matemáticas con Lectura	0,70
Matemáticas con Ciencias	0,75
Matemáticas con Problemas	0,85
Lectura con Ciencias	0,80
Lectura con Problemas	0,78
Ciencias con Problemas	0,71

De ahí que quizá no haya mucho que esperar de los resultados de PISA para iluminar los pormenores del contenido curricular o de los procesos de enseñanza-aprendizaje propios de cada materia. A este respecto, los análisis realizados

sobre las preguntas de las ediciones anteriores de PISA han resultado ser menos fecundos de lo esperado. Y tampoco es de esperar que en esta edición de PISA los análisis secundarios vayan a proporcionar grandes conclusiones sobre los diversos campos temáticos de las Ciencias.

Ahora bien, esta cualidad generalista de PISA ¿debe ser considerada como un defecto o como una virtud? ¿Es un fallo de diseño por parte de PISA el que sus resultados no ofrezcan un poder de discriminación más acusado entre los distintos «saberes»? ¿O es más bien el reflejo de una característica de la formación de los alumnos de la enseñanza secundaria inferior, que necesitan un aprendizaje menos focalizado y de más amplio espectro? Si, como parece cada vez más evidente, no debe esperarse de los alumnos a esta edad una asimilación y un trabajo tan específico como el que pretende en estos momentos la organización de la enseñanza secundaria y la práctica didáctica de su profesorado, quizá fuera oportuno empezar a cuestionarse si no habría que acometer reformas estructurales hacia un modelo de secundaria menos atomizado, con menos materias y con cuerpos de profesorado menos especializados.

La secundaria tiene su origen histórico en la preparación de los alumnos para su entrada a la universidad. Es aún muy reciente, en términos históricos, la extensión, con obligatoriedad, de la enseñanza secundaria a todos los alumnos, tanto a los que han de cursar estudios universitarios como a los que no. Y esta universalización ha supuesto

un cambio de naturaleza. Sin embargo, parece que se mantiene, más en la práctica que en la normativa, una organización y una orientación afín a la propedéutica universitaria ahora que la nueva secundaria, universal y obligatoria, ya no tiene su única salida natural en la universidad. Aun a riesgo de hacer caricatura de una realidad mucho más compleja, parece como si el honor y el ideal profesional de los profesores de secundaria fueran más afines a los de sus colegas universitarios, con su parcelación de materias, su especialización, sus ideales de enseñanza centrados en hacer avanzar el contenido científico de las disciplinas. Y en este esfuerzo parece que se olvidan o se minusvaloran los rendimientos genéricos y transversales, como el de saber leer comprensivamente, que son precisamente los que necesitan los jóvenes a punto de entrar en la vida adulta y de convertirse en ciudadanos de pleno derecho, algo que las pruebas PISA ayudan a recordarnos.

Pero volvamos nuestra atención a otras características de este estudio PISA 2006, centrado en la evaluación de las Ciencias, y examinemos las novedades que contiene y los principales puntos de interés para un país como España. A mi juicio, la edición 2006 de PISA nos presenta los siguientes rasgos: a) la incorporación de preguntas actitudinales, b) la importancia dada a los procedimientos metodológicos de la ciencia, c) una batería de preguntas que refleja una concepción de la ciencia tendente a lo experimental y utilitario, y d) el permanente riesgo de sesgo cultural anglosajón.

3. La incorporación de preguntas actitudinales

Siempre han estado presentes en los estudios PISA determinadas preguntas sobre las actitudes de los alumnos y sobre la valoración que les merecen las materias que estudian. Pero el lugar en el que estas preguntas se ubicaban era el cuestionario de contexto. Ahora, en PISA 2006, el nuevo equipo internacional de expertos ha insistido en aumentar el número de este tipo de preguntas y trasladarlas a los cuadernos de prueba, en un esfuerzo por contextualizarlas mejor y por hacer aparecer las actitudes como un resultado de la acción educativa al mismo nivel que los conocimientos y las destrezas.

Las dimensiones en las que PISA pretende proporcionar resultados sobre las actitudes de los alumnos son: a) el interés por la ciencia, b) el apoyo a la investigación científica y c) la responsabilidad hacia el desarrollo sostenible. Tras los resultados de la prueba piloto realizada en 2005, las preguntas relativas a las dos primeras dimensiones han mostrado suficiente solidez psicométrica como para poder aparecer en los cuadernos de prueba junto a las preguntas de conocimientos. No ha ocurrido lo mismo con las de la dimensión de responsabilidad, cuyos problemas de escalamiento han aconsejado mantenerlas en el cuestionario de contexto.

El énfasis en las actitudes de los alumnos ante la ciencia parece responder a una creciente inquietud en buena parte de los países de la OCDE por la disminución del

atractivo de las carreras científicas como una opción de futuro profesional. En otras palabras, parece como si se percibiera una disminución progresiva de las vocaciones científicas en las nuevas generaciones. Ciertamente, dentro del campo educativo, ya aparecen dificultades para reclutar profesores de Matemáticas y Ciencias en determinados países europeos, y también se detecta una cierta huida de candidatos ante las carreras de contenido científico, normalmente más largas y exigentes que otras carreras de origen más reciente, pero no peor remuneradas.

Sin embargo, la nueva ubicación de las preguntas actitudinales en las pruebas de PISA ha encontrado resistencias para su aceptación en un número no mayoritario, pero sí significativo de países. Una consecuencia de esas resistencias es el retraso en la publicación oficial por parte de la OCDE de los marcos teóricos de la evaluación PISA 2006, que aún no se ha producido. Los críticos han señalado varios aspectos que no suscitan consenso:

- a) No es lo mismo tratar las actitudes como variables de contexto, es decir, como condicionantes del proceso de aprendizaje en Ciencias, que como variables de rendimiento, es decir, como componentes de los resultados alcanzados. No están claras las consecuencias de este cambio de naturaleza en el tratamiento dado a las preguntas sobre actitudes y las exploraciones previas no son lo suficientemente concluyentes como

para hacer de los elementos actitudinales un objetivo deseable de la comparación internacional.

- b) Se introduce un aspecto subjetivo y valorativo en una evaluación que hasta ahora quería ser objetiva y libre de valores. Para contrarrestar esta crítica, en los cuadernos de prueba las preguntas actitudinales se presentan con un fondo gris junto con instrucciones explícitas de que dichas preguntas no contribuyen a la puntuación final del alumno, puesto que no contienen respuestas objetivamente correctas. Pese a lo cual, en la prueba piloto se ha detectado que las respuestas de los alumnos tendían a orientarse en el sentido de la aceptabilidad social.
- c) La presencia de estas preguntas, que suponen algo menos de la cuarta parte de la batería final de preguntas de Ciencias, provoca la disminución del tiempo disponible para explorar el dominio de otras áreas cognitivas que no encuentran lugar en los cuadernos de prueba de esta edición de PISA, y ha supuesto importantes restricciones a la hora de la selección de las unidades de Ciencias para la prueba definitiva.

El Consorcio internacional que diseña y analiza los datos recogidos en las pruebas PISA tiene intención de construir una escala TRI de las respuestas actitudinales, pese a que determinados presupuestos, como la ausencia de una respuesta correcta y, por tanto, la ausencia de un parámetro de dificultad

para las preguntas, convierten a este intento en una iniciativa estadística hasta cierto punto sorprendente. Todo ello hace que el énfasis otorgado a las actitudes en el estudio de las Ciencias sea una de las facetas de mayor novedad e interés en esta edición de PISA.

4. La importancia concedida a los conocimientos sobre metodología científica

Los marcos teóricos de PISA 2006, que aún se encuentran en fase de borrador, dividen los contenidos objeto de examen en conocimientos *de* Ciencias (los sistemas vivos, los sistemas físicos, los sistemas de la Tierra y el espacio) y en conocimientos *sobre* Ciencias (la indagación científica, las explicaciones científicas, la ciencia y la tecnología en la sociedad).

Es de destacar que los temas en torno a cómo se hace y se utiliza la ciencia ocupan un lugar tan importante como los temas más clásicos en torno a qué se sabe sobre temas científicos. Y esto puede plantear algún problema a los alumnos españoles, pues aún existen tradiciones docentes que tienden a dejar de lado, en la enseñanza de las Ciencias, los aspectos metodológicos en beneficio del conocimiento del *corpus* de adquisiciones teóricas. Ciertamente, en este aspecto la práctica docente es desigual, pero constituye uno de los puntos de interés de esta edición de PISA el saber si los alumnos españoles se mueven con la misma soltura en ambos tipos de temas.

Esperar que los alumnos estén enterados sobre el modo en que los científicos hacen progresar las Ciencias y consiguen ponerse de acuerdo en dar por probadas o refutadas sus controversias teóricas tiene una mayor tradición fuera de nuestras fronteras que en España. Se erige ciertamente en un aspecto importante de la formación en Ciencias si uno de los objetivos que se persiguen es el de promover la aparición de vocaciones científicas. Pero en nuestro país no está claro que ese objetivo se persiga con alta prioridad, siendo de temer que el unamuniano «¡que inventen ellos!» no haya pasado del todo a la historia, ya que el grueso de nuestra investigación científica parece consistir todavía en digerir la investigación ajena más que en promover la propia. En una palabra, no parece que nos sintamos concernidos por la responsabilidad de que nuestras omisiones investigadoras nos lleven a una parálisis de creatividad científica: somos colectivamente conscientes de no ser competitivos en tareas de investigación y por ello las cuestiones metodológicas no parecen recibir suficiente atención docente hasta etapas más avanzadas de los estudios.

Esto no es así en otros países y esta convicción ha influido en el diseño de las pruebas, que son realmente insistentes en evaluar el conocimiento de los requisitos de los diseños experimentales controlados y, por extensión, en el conocimiento del criterio de demarcación entre las afirmaciones que pueden ser consideradas científicas y las que no. Detrás de todo ello está un loable intento ilustrador por evaluar las capacidades de los

alumnos para razonar basándose en hechos, y no solo en opiniones, y para argumentar con apoyo de pruebas verificables olvidándose de las tentaciones de apoyar las propias convicciones en fundamentos supersticiosos, aún fuertes en esas edades.

5. La concepción de las Ciencias en PISA 2006

Recordemos que PISA es una evaluación guiada por expertos. Estos expertos hacen un esfuerzo para definir una prueba que represente lo que se considera deseable que los alumnos de 15 años conozcan, sepan aplicar y valoren. El grupo de expertos, junto con los miembros del Consorcio internacional de PISA y en comunicación constante con los representantes de los países, termina elaborando un conjunto de preguntas que, en el caso de la prueba de Ciencias, es de una calidad muy notable, pese a la dificultad de conseguir una representación perfecta de las distintas áreas temáticas consideradas relevantes.

Pero también hay que señalar que el peso del grupo de expertos es lo suficientemente importante como para que los cambios en su composición lleguen a tener consecuencias en la orientación de las evaluaciones. Para esta edición de PISA el grupo de expertos ha incorporado nuevos miembros y en particular un nuevo presidente, quien ha impulsado una dinámica que ha terminado empujando al resto de los expertos, al Consorcio de PISA, al Secretariado de la OCDE y a los países participantes a cambiar –de un modo

innecesario, en mi opinión– la definición de lo que debe entenderse por Ciencias. El cambio ha consistido en incluir la dimensión actitudinal⁴ como componente del rendimiento de los alumnos, inclusión que, como hemos mencionado, no se ha logrado sin tensiones ni resistencias.

Otro motivo de incomodidad para algunos países participantes está en la ambigüedad constante en los documentos originales, en inglés, entre el uso global y genérico del término «ciencia» y su uso específico y curricular en entornos escolares. El primer uso se traduce al castellano como «ciencia» y el segundo es mejor entendido si se lo traduce como «Ciencias», en plural y con mayúsculas, como ha venido siendo tradicional en nuestros planes de estudios desde hace muchas décadas. Y si en inglés la unicidad de términos hace que el paso de uno a otro uso sea habitual y constante, ese paso no es tan fácil en otros idiomas, pues no es lo mismo hablar de la enseñanza de la ciencia, que para nosotros parece referirse más a la metodología científica, que hablar de la enseñanza de las Ciencias, que evocan en nuestro país los contenidos de las Ciencias de la Naturaleza, edificados en torno a la polaridad Biología y Física-Química o ciencias de lo vivo y de lo inerte. En inglés, sin embargo, no existe esa distinción y al tener que traducir a otros idiomas determinadas asociaciones o connotaciones se pierden.

(4) En su redacción inicial se denominaba «dimensión afectiva» y su carácter valorativo era mucho más marcado.

Por otro lado, si examinamos el conjunto de preguntas sobre Ciencias preparada para esta edición de PISA, emerge una cierta concepción de lo que debe ser considerado ciencia que quizá no se ajusta del todo a nuestras expectativas o tradiciones y en todo caso merece la pena ser comentado. A mi juicio, la concepción de la ciencia presente en la batería de preguntas de esta edición de PISA parece pecar ligeramente de un sesgo instrumental y experimentalista:

- a) Un sesgo instrumental, y quizá debería decir utilitario, en cuanto que las preguntas tienden siempre a conectar el tema visitado con los intereses, beneficios o perjuicios del alumno, dejando de algún modo traslucir el mensaje de que el conocimiento científico tiene interés por sus beneficios prácticos, más que por el crecimiento del conocimiento mismo o por la superior imagen que nos proporciona del mundo en que vivimos.
- b) Un sesgo experimentalista, más presente durante la elaboración de las pruebas que en el conjunto final, que trasluce un criterio de demarcación entre lo científico y lo extracientífico muy dependiente de la posibilidad de experimentar controladamente y con exclusión de las observaciones no repetibles. Así, en un principio, las disciplinas descriptivas, como la Geografía, no son consideradas científicas y se ha pretendido en algunas preguntas de la prueba piloto que las conclusiones de los economistas

y de los arqueólogos quedaran excluidas del campo de lo científico. La reacción internacional ha hecho que esas preguntas hayan desaparecido de la batería final, lamentablemente eliminando con ellas unas unidades de contenido valioso e interesante.

En general puede decirse que, en las unidades de Ciencias, las referencias a los temas adoptan un estilo poco distanciado: son siempre las consecuencias para *mi* salud, *mi* seguridad, *mi* entorno y *mi* comunidad lo que constituye el marco contextual elegido. Quizá este estilo esté muy adaptado al público escolar –aunque más al de la escuela primaria que al de la secundaria–, pero para un joven que está a punto de entrar en la vida adulta, también sería deseable que el conocimiento científico no estuviera siempre bajo el prisma de la utilidad inmediata, sino bajo el interés ilustrado por adquirir una visión del mundo más adulta.

En este sentido, destacan ciertas ausencias: la poca presencia de preguntas sobre Química, la muy escasa mención al tema de la evolución y la total inexistencia de preguntas sobre cosmología, esto es, sobre cuestiones astronómicas o espaciales que vayan más allá del Sistema Solar. Y aunque sean bastante carentes de carácter utilitario, parece ser deseable en los alumnos de 15 años una cierta familiaridad con nociones como «agujero negro» o «Big Bang» o con las ideas de «conservación de la materia» o de «selección natural». Estas nociones o ideas deberían formar parte del marco conceptual

de referencia de un adulto en nuestros tiempos, pues le ayudan a situar sus ideas sobre el puesto del hombre en el mundo de un modo ilustrado y alejado de tentaciones oscurantistas.

6. El permanente riesgo de sesgo cultural

Es habitual, al comentar los resultados poco brillantes que suelen obtener los alumnos españoles en las evaluaciones internacionales, achacarlos a estos motivos: a) al estilo novedoso e inhabitual de las pruebas, que no conectan demasiado con las prácticas escolares habituales, y b) al sesgo anglosajón que manifiestan. A mi entender, la primera acusación es cierta, pero no por ello lo que denuncia deja de ser beneficioso. A nuestras prácticas establecidas les viene muy bien el contacto con nuevos modos de presentar las preguntas y, a través de ellas, con nuevos estilos didácticos. Es una bocanada de aire fresco proveniente del exterior que renueva la atmósfera algo cargada de nuestras tradiciones didácticas.

La segunda acusación, la del sesgo cultural que invalida las pruebas, es globalmente injusta, aunque ciertamente apunte a determinados elementos que pueden provocar una cierta alarma. Me refiero a la presencia de una cierta «jerga» profesional bien establecida en inglés, pero que resulta menos habitual y algo difícil de traducir en castellano.

He aquí algunos ejemplos:

- a) el área temática «Ciencias de la Tierra», que no corresponde a ninguna materia habitual en nuestro currículo y que incorpora elementos de diversas disciplinas (Geografía, Geología, Meteorología y Astronomía de corto alcance),
- b) la referencia constante al razonamiento basado en «evidencias», modo de expresión que resulta en nuestro idioma poco «evidente» y que, al traducirlo, provoca pesados rodeos para incluir las connotaciones de lo fáctico y de lo fundamentado en pruebas,
- c) la expresión «pregunta científica», que ha adquirido en inglés una consolidación fija y muy concreta como denominación del objetivo o meta final de una indagación científica, pero que no tiene en nuestros ambientes escolares el mismo estatus de frase hecha y de denominación familiar.

Además de la presencia de una cierta «jerga» de origen anglosajón, otro de los motivos de alarma radica en que, en PISA, la mayor parte de los expertos internacionales y de los técnicos del Consorcio internacional responsable de la organización y coordinación de las distintas operaciones (muestreo, verificación de la traducción, corrección de preguntas abiertas, limpieza y análisis de datos) provienen de países de habla inglesa. Esta internacionalización insuficiente del Consorcio de PISA se deja notar en detalles como que se hayan rechazado 53 unidades

presentadas por México para formar parte de las pruebas, o de que no haya prácticamente ninguna bibliografía citada en los documentos internos y públicos de PISA que no haya sido redactada originalmente en inglés. Y si bien no parece demasiado halagador que se ignore la investigación pedagógica en lengua española, es desde luego a todas luces injusto que se haga lo mismo con la producida originalmente en francés o alemán, que suponen un excelente *corpus* de conocimientos y estudios.

Pero todos estos indicios no deben arrastrarnos hacia la sospecha generalizada. Todo lo contrario. Hay que decir bien alto y claro que el proceso de elaboración de las pruebas no ha podido ser más abierto y transparente. En Ciencias, se han elaborado muchas más preguntas de las necesarias: después de haber desechado un buen número en las primeras fases de consulta internacional, todavía se pusieron a prueba, en la evaluación piloto realizada el año pasado, un número superior en más de tres veces al necesario para la prueba final⁵. Durante la fase de elaboración se han

(5) En la prueba piloto de PISA 2006, solo se pusieron a prueba preguntas de Ciencias. Se incluyeron 376 preguntas, de las que finalmente han pasado a la prueba final 118, algo menos de la tercera parte. La prueba final consta de 140 preguntas de Ciencias (las mencionadas 118 más otras 22 preguntas presentadas en ediciones anteriores de PISA), de las que 32 preguntas versan sobre actitudes. Además se utilizan 48 preguntas de Matemáticas y 32 de Lectura, todas ellas provenientes de anteriores evaluaciones PISA.

producido consultas repetidas a los países participantes, y las opiniones provenientes de estas consultas se han tenido en cuenta mediante un sistema estructurado de puntuaciones y comentarios. Incluso se organizaron, en dos ocasiones, reuniones internacionales específicas, con el nombre de Foro de Ciencias, para que los países pudieran debatir directamente con los equipos del Consorcio sobre las unidades y preguntas propuestas. Se han corregido multitud de pequeños detalles que suscitaban recelo en algunos países e incluso han desaparecido unidades enteras de contenido valioso por detectarse sesgos culturales de difícil conciliación⁶.

En suma, las características de una evaluación internacional que incluye a multitud de países con rasgos culturales diversos hacen que sea prácticamente imposible conseguir un conjunto de ítems totalmente satisfactorio para todos los países participantes. Pero hay que reconocer al Consorcio internacional de PISA el esfuerzo que ha realizado y los recursos que ha

(6) Un ejemplo: una unidad que versaba sobre los grupos sanguíneos hubo de ser retirada por no poder conciliar las diferentes denominaciones usuales en los distintos países. En particular, el grupo sanguíneo que no contiene antígenos suele ser denominado como grupo «cero» en los países europeos y como grupo «o» en los países anglosajones. Aunque la primera denominación parezca ser más correcta desde el punto de vista científico, incrementaría apreciablemente la dificultad de la pregunta para los alumnos de los países que se expresan en inglés.

invertido para lograr que ese conjunto esté lo más libre posible de condicionantes culturales que pudieran interferir en el valor comparativo de los datos obtenidos. Pretender invalidar o minusvalorar los resultados de la evaluación que realiza PISA basándose en la contaminación cultural de sus pruebas es, en mi opinión, una reacción de tipo primario, desinformada e injusta.

7. A modo de conclusión: el interés de estas pruebas PISA para España

Además de lo expuesto, conviene tener en cuenta otras características específicas de los alumnos españoles que añaden interés a esta evaluación PISA, tanto en lo que supone un mayor conocimiento de la enseñanza de las Ciencias como el conocimiento de los puntos fuertes y débiles de nuestro sistema educativo, aspecto este en el que las anteriores ediciones PISA han arrojado una bienvenida iluminación.

A diferencia de otros países donde los alumnos de 15 años pueden estar cursando programas de estudios distintos, en España todos ellos están cursando la Educación Secundaria Obligatoria. Las diferencias de rendimiento observadas no quedan, por tanto, oscurecidas por factores diferenciales atribuibles a la estructura del sistema educativo.

Por otro lado, en nuestro país el ámbito popular de las Ciencias incluye también a las Matemáticas en un conjunto globalmente opuesto al ámbito de las Letras.

Pero en su especificación curricular, las Ciencias son unas veces una sola asignatura, y otras veces se desglosan en varias, y no todas de estudio obligatorio. En esto seguimos una pauta muy común con otros países. Será un reto añadido para los analistas el lograr detectar las diferentes pautas de rendimiento con independencia del número y modalidad de materias de Ciencias cursadas por los alumnos.

También despierta un interés específico en esta nueva evaluación de PISA el perfil actitudinal de nuestros alumnos, que nos mostrará si ellos también se enfrentan a los estudios de contenido científico con una entrega e interés disminuidos y si, en suma, nuestro país va a sufrir también un agravamiento en la escasez de vocaciones científicas.

Asimismo, estamos pendientes del papel que realizarán nuestros alumnos en las preguntas de metodología científica, que nos mostrarán si a este particular contenido de la enseñanza de las Ciencias se le debe otorgar una mayor atención curricular y un tiempo más extenso en la práctica de aula. Y en cuanto a esta práctica, será interesante estudiar, a partir de la información de contexto proporcionada por alumnos y directores, las particularidades de nuestra enseñanza de las Ciencias y su comparación con los enfoques didácticos de otros países, como, por ejemplo, la importancia relativa dada a la realización de prácticas de laboratorio como parte de la enseñanza de esta materia.

Y, por supuesto, estaremos muy interesados en la medida del rendimiento general de nuestros alumnos en Ciencias en comparación con los de otros países, e internamente dentro de nuestro propio país. Pues una de las novedades de esta edición de PISA en España es el aumento muy considerable del número de Comunidades Autónomas que han decidido, con cargo a su presupuesto, ampliar la muestra de centros y alumnos evaluados por PISA en su territorio con el fin de que los datos recogidos tengan precisión suficiente para ser comparados con fiabilidad estadística con los de otros territorios y países.

En PISA 2003, Castilla y León, Cataluña y el País Vasco aumentaron su muestra. Ahora, en PISA 2006, se les unen las Comunidades Autónomas de Andalucía, Aragón, Asturias, Cantabria, Galicia, La Rioja y Navarra. Con la excepción de Andalucía, todas ellas se encuentran situadas en la mitad norte de España. Este incremento de participación tendrá el beneficio añadido de que se empezará a romper el prejuicio existente contra la comparación interterritorial de los resultados educativos, hasta ahora uno de los escollos importantes a los que se ha enfrentado nuestro Instituto de Evaluación. En esta ocasión, la comparación será realizada por el Consorcio de PISA y publicada por la OCDE, lo que parece dar un marchamo de aceptabilidad. Sería deseable de todos modos que el hábito de la comparación interterritorial se extendiera como recurso habitual a la publicación de resultados de alcance menos internacional.

Y por último, un escollo o un reto final que ronda a esta evaluación de PISA es el de acertar con una fórmula de difusión de los resultados que sepa destacar los mensajes apropiados para las distintas audiencias. Es deseable, en especial, que esta fórmula sepa sortear las simplificaciones a las que nos tienen acostumbrados los medios, a quienes hay que recordar que PISA no es una competición internacional entre sistemas educativos, sino un instrumento de observación que, gracias a su componente comparativo, se beneficia de un claro aumento de resolución y, por ello, de mayor penetración analítica que el que se consigue con otros estudios de la realidad educativa no comparativos. Y también que, en contra de una cierta tendencia primaria al catastrofismo, los resultados alcanzados hasta ahora por España –y los que previsiblemente alcanzará en esta edición 2006– muestran que el rendimiento educativo de nuestro país es plenamente homologable con el de los países de la OCDE. No obstante, y partiendo de esta constatación, sigue siendo deseable mantener una cierta tensión en los diferentes resortes del sistema educativo para conseguir elevar unos resultados que aún son susceptibles de mejora.