

El Programa Internacional para la Evaluación de los Alumnos (PISA). Una perspectiva general

Ross Turner

Investigador principal del Australian Council for Educational Research (ACER)

Resumen: En este artículo se ofrece una perspectiva general del proyecto PISA con el fin de establecer un contexto para el debate de algunos resultados de la evaluación de 2003. Dicha perspectiva general describe los objetivos del proyecto, sus participantes, sus procedimientos y organización, sus principales componentes y metodologías, y las características concretas de las matemáticas en PISA, que fue la principal área de evaluación en 2003. Asimismo, se ofrece un brevísimos resumen de algunos resultados claves de la evaluación PISA 2003 en España, extrayendo conclusiones relevantes en materia de política educativa.

Palabras clave: evaluación internacional, PISA, matemáticas, muestreo, idoneidad cultural, diseño de tests, análisis de datos, características contextuales.

Abstract: *Programme for International Student Assessment. An overview*

This article provides a general overview of the PISA project with the aim of setting up an appropriate context for debating on some specific outcomes of 2003 assessment survey. This general perspective focuses on the goals of the project, its participants, procedures and organisation, its major components and methodologies as well as on the precise characteristics of Mathematics in PISA, which was the main assessment subject in 2003. Similarly, a very brief summary on certain key outcomes of PISA 2003 assessment in Spain is also provided.

Key words: international evaluation, PISA, mathematics, sampling, cultural appropriateness, test design, data analysis, background characteristics.

¿QUÉ ES PISA?

El Programa para la Evaluación Internacional de los Alumnos (PISA) fue diseñado y puesto en marcha por la OCDE a finales de los años noventa como un estudio comparado, internacional, periódico y continuo sobre determinadas características y competencias de los alumnos. Su finalidad es generar indicadores sobre aspectos del rendimiento educativo. Por tanto, no es esencialmente un proyecto de investigación, aunque los datos generados pueden resultar de gran interés para los investigadores. Son los países miembros de la OCDE quienes gestionan y dirigen cooperativamente PISA, en colaboración con un número elevado y cada vez mayor de países asociados. La OCDE, a través de un concurso público, designa a un proveedor de servicios externo para llevar a cabo el proyecto en cada ciclo del estudio.

Los estudios PISA se realizan cada tres años. Las primeras evaluaciones tuvieron lugar en el año 2000 y las segundas en 2003, y están previstos ya los estudios para 2006 y 2009. Los procedimientos científicos de muestreo se emplean para determinar qué centros educativos y qué alumnos se incluirán en cada estudio. Cada país participante utiliza un conjunto típico de instrumentos de evaluación y estudio según unas condiciones comunes y controladas que permiten la comparación basándose en los datos resultantes. Se hace uso de técnicas analíticas que permiten la realización de comparaciones en cada país participante, entre todos ellos y entre los distintos ciclos del estudio.

PISA evalúa a escolares de 15 años. Estos alumnos se acercan al final de la escolarización obligatoria en la mayoría de los países participantes, lo que los convierte en un grupo de edad adecuado para realizar una evaluación sobre su grado de preparación con el fin de afrontar los desafíos diarios de nuestras sociedades modernas.

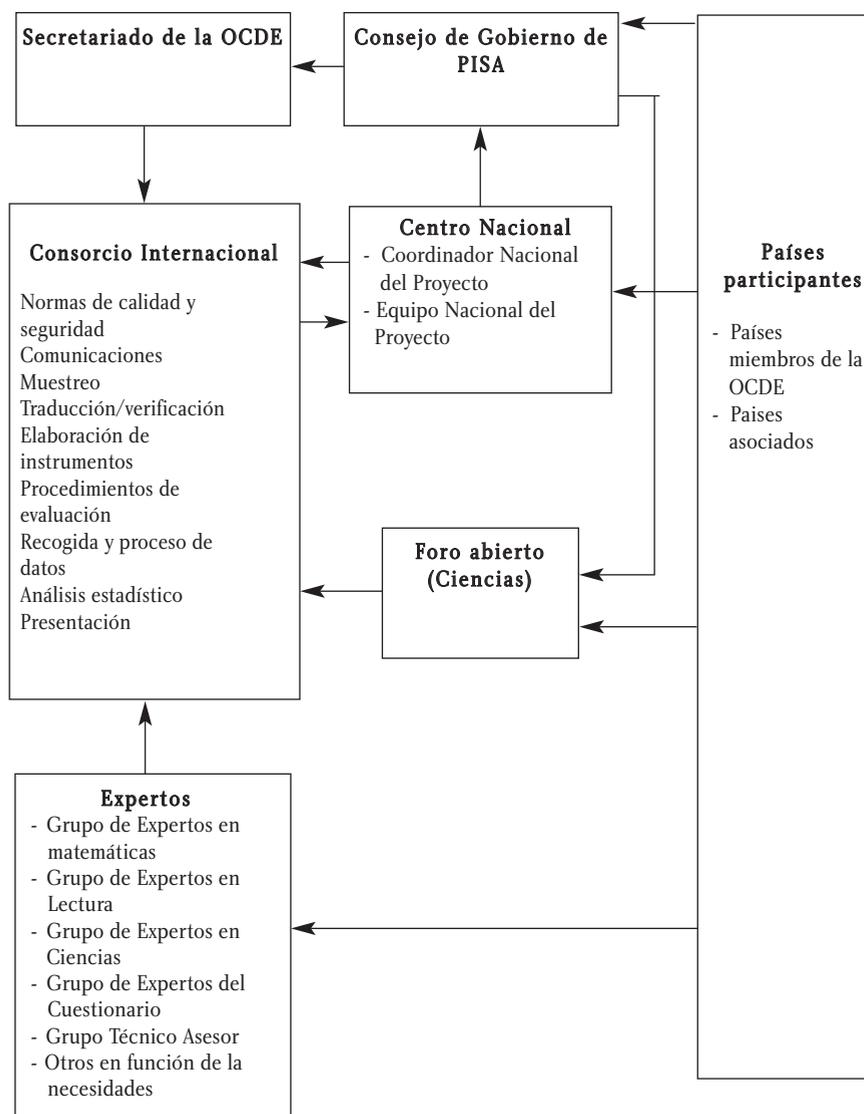
Para ello, PISA adopta una perspectiva «competencial» centrada en averiguar hasta qué punto los alumnos son capaces de usar los conocimientos y destrezas que han aprendido y practicado en la escuela cuando se ven ante situaciones y desafíos en los que esos conocimientos pueden resultar relevantes. Es decir, evalúa cómo los alumnos pueden hacer uso de sus destrezas lectoras para comprender e interpretar distintos tipos de material escrito con el que probablemente se van a encontrar al gestionar su vida diaria; de qué forma pueden utilizar sus conocimientos y destrezas matemáticas para resolver distintos tipos de desafíos y problemas relacionados con las matemáticas, y el modo en que los alumnos pueden hacer uso de sus conocimientos y destrezas científicas para comprender, interpretar y resolver distintos tipos de situaciones y desafíos científicos. PISA también permite en ocasiones la evaluación de competencias transversales, según lo estimen oportuno los países participantes. Por ejemplo, el ciclo del estudio de 2003 incluía una evaluación de competencias generales sobre la resolución de problemas. Además, el estudio PISA recopila información de los alumnos sobre distintos aspectos de su entorno familiar y escolar e información de los centros sobre distintos aspectos de su organización y oferta educativa. Esta información se recoge para facilitar un estudio pormenorizado de los factores, tanto de los países como entre ellos,

que están asociados con los distintos niveles de competencia lectora, matemática y científica de los alumnos de 15 años de cada país. Los análisis resultantes serán de interés para los responsables políticos de los países participantes y para los investigadores que buscan comprender mejor los factores que influyen en los resultados educativos.

CARACTERÍSTICAS ORGANIZATIVAS DE PISA

Se puede pensar PISA como en un proyecto que opera en una serie de niveles. Los distintos niveles y los principales participantes en el proyecto se ilustran en el gráfico I.

GRÁFICA I
Principales participantes en el Proyecto PISA



La organización del proyecto PISA corre a cargo de un Secretariado de la OCDE. El Consejo de Gobierno de PISA (PGB) es un comité de la OCDE formado por delegados y observadores de los países participantes, que supervisa y establece los parámetros de la política reuniéndose unas dos veces al año. Por tanto, la gestión de PISA se realiza a través de un proceso de colaboración en el que participan burócratas funcionarios de alto rango y responsables en materia de educación de los países participantes.

En cada país participante se establece un Centro Nacional y se nombra a un Coordinador Nacional del Proyecto (NPM) para coordinar todas las actividades a escala nacional. Normalmente, éste trabaja en estrecha colaboración con el miembro del Consejo de Gobierno de PISA del país, para establecer una perspectiva nacional sobre aspectos de la política y sobre cuestiones relacionadas con la aplicación del proyecto. Por lo general, el Centro Nacional cuenta con un pequeño equipo que trabaja en la elaboración y puesta en marcha del proyecto a escala nacional.

A escala internacional, el proyecto lo desarrolla un contratista designado por la OCDE para realizar dicho trabajo. Se selecciona a un contratista para cada ciclo de evaluación a través de un concurso público. Un consorcio internacional, dirigido por el Consejo Australiano de Investigación Educativa (ACER) ha sido el contratista elegido para cada uno de los tres primeros ciclos del estudio PISA. El contratista es el responsable de desarrollar todos los aspectos de la evaluación bajo la estricta supervisión del Secretariado. Esto incluye una amplia gama de actividades, algunas de las cuales se enumeran a continuación:

- el establecimiento de normas de calidad con relación a todos los aspectos del proyecto;
- el establecimiento de procedimientos para garantizar que todos los participantes cumplen correctamente esas normas de calidad, y de mecanismos para supervisar la calidad de la puesta en marcha del proyecto;
- el mantenimiento de comunicaciones abiertas y eficaces entre todos los participantes del proyecto;
- el trabajo en estrecha colaboración con el personal del Centro Nacional para recopilar todos los datos nacionales sobre cuestiones relativas a la elaboración y puesta en marcha del proyecto;
- el establecimiento de marcos de evaluación;
- el establecimiento de todos los instrumentos de evaluación y estudio;
- el establecimiento y puesta en marcha de métodos de muestreo;
- el establecimiento de procedimientos operativos para la administración de la prueba y de toda la documentación pertinente;
- la formación del personal más importante del Centro Nacional en cuestiones relativas a los requisitos para la puesta en marcha del estudio;

- el establecimiento de todos los procedimientos de recogida de datos; la recogida y el proceso de datos de las evaluaciones;
- el análisis de los resultados y la preparación de material de apoyo a la OCDE para elaborar los informes exigidos por ésta;
- el apoyo al Secretariado a través del análisis de los procedimientos y mejoras de planificación para los futuros ciclos del estudio.

Para llevar a cabo todas estas tareas el contratista debe, en primer lugar, reunir a un amplio equipo de personas con conocimientos especializados en varias áreas. Parte de estos conocimientos compete al personal disponible a través de los organismos asociados al Consorcio. Además, el Consorcio trabaja con varios grupos e individuos claves: una serie de grupos de expertos integrados por reconocidos expertos en áreas como la evaluación, la traducción, especialistas técnicos en muestreo, análisis estadístico y elaboración de cuestionarios. En caso necesario, el contratista puede recurrir a los servicios de otros profesionales.

Lo más importante es que el contratista trabaja en estrechísima colaboración con el personal empleado a escala nacional. La conexión entre los coordinadores nacionales del proyecto y el Consorcio Internacional es vital para el éxito del proyecto. Cada uno depende del otro para garantizar dicho éxito. Los centros nacionales proporcionan al Consorcio información sobre las condiciones y limitaciones que operan en cada país y, a su vez, éste responde con datos sobre los distintos elementos del proyecto que se están elaborando y con importantes recomendaciones sobre la mejor manera de llevarlo a cabo en cada país. El contratista proporciona a los centros nacionales información sobre los requisitos del proyecto, materiales preliminares para ser analizados a escala nacional, y respuesta, formación y materiales para facilitar la puesta en marcha del proyecto.

En ocasiones, se han empleado mecanismos de consulta, como el Foro de Ciencias. Se trata de un foro abierto que proporciona a los países participantes la oportunidad de designar a expertos nacionales que puedan representar directamente los intereses y las opiniones del país cuando se analizan ciertos aspectos técnicos específicos del proyecto. Este foro permite disponer de una base más amplia de información especializada de la que es posible obtener a través de los grupos de expertos del contratista internacional. En el caso del Foro de Ciencias, este grupo analizó las prioridades y los temas cuando se estaba conceptualizando el marco de esta área y ha proporcionado datos importantes para la elaboración del material del estudio relacionado con la evaluación de las ciencias y la evaluación de las actitudes de los alumnos hacia éstas. El miembro del Consejo de Gobierno de PISA de cada país designa a los delegados del foro. Un Foro de Matemáticas similar operó en PISA 2003 y ocasionalmente se ha convocado un Foro del Cuestionario.

¿QUIÉN PARTICIPA?

En el actual ciclo de evaluación PISA participan 57 países, incluidos los 30 de la OCDE y otras 27 naciones denominadas «países asociados». Todos ellos se enumeran en la Tabla tabla I.

TABLA I
Países participantes en PISA 2006

PAÍSES DE LA OCDE

Australia	Alemania	Luxemburgo	España
Austria	Grecia	México	Suecia
Bélgica	Hungría	Países Bajos	Suiza
Canadá	Islandia	Nueva Zelanda	Turquía
República Checa	Irlanda	Noruega	Reino Unido
Dinamarca	Italia	Polonia	EE. UU.
Finlandia	Japón	Portugal	
Francia	Corea	Eslovaquia	

PAÍSES ASOCIADOS

Argentina	Colombia	Letonia	Eslovenia
Azerbaiyán	Croacia	Lituania	Taiwán
Brasil	Estonia	Liechtenstein	Tailandia
Bulgaria	Indonesia	Qatar	Túnez
Chile	Israel	Rumania	Uruguay
China ¹ RAE de Hong Kong	Jordania	Federación Rusa	
China RAE de Macao	Kirguizistán	Serbia y Montenegro ²	

La presencia de estos países supone el empleo de más de 40 lenguas diferentes.

¹ China no participa como país único. Sin embargo, dos de sus regiones administrativas especiales (Hong Kong y Macao) lo hacen como si cada una fuese un país independiente.

² Cada una de las dos zonas de Serbia y Montenegro participan como si fuesen países independientes.

El resto del personal que tomó parte en el proyecto PISA 2003, como los miembros del Consejo de Gobierno de PISA, los coordinadores nacionales del proyecto, el personal del Secretariado de la OCDE, el personal del Consorcio Internacional, los miembros de los grupos de expertos u otros asesores, se incluye en el informe internacional de la OCDE (*OCDE 2004*, p. 474).

GESTIÓN Y PUESTA EN MARCHA DE PISA

El principal estudio de PISA se lleva a cabo cada tres años. Para el ciclo del estudio PISA 2003 el trabajo comenzó a finales de 2000 (antes de que hubiesen concluido las principales pruebas del estudio del anterior ciclo, y bastante antes de que ninguno de los datos más importantes de ese ciclo se hubiesen procesado y analizado). Las primeras actividades estuvieron relacionadas con el establecimiento de los conocimientos especializados exigidos al Consorcio Internacional y a los grupos de expertos del Consorcio.

El Consorcio Internacional comenzó el trabajo revisando el marco de las matemáticas, creando un nuevo marco de resolución de problemas y modificando, de forma sustancial, el marco del cuestionario a finales de 2000. Los grupos de expertos del consorcio, que incluyen a reconocidos y respetados académicos y profesionales de las principales áreas de conocimiento, orientaron el trabajo de desarrollo intelectual del marco de la política establecido por el Consejo de Gobierno de PISA y el Secretariado. Los países participantes dispusieron de una serie de oportunidades para proporcionar información y responder a los documentos según se iban elaborando. Los documentos marco resultantes constituirían posteriormente la base conceptual sobre la que se elaborarían las evaluaciones en esas áreas. La elaboración de los ítems de la prueba se inició una vez que las orientaciones de los marcos respectivos estuvieron lo suficientemente avanzadas.

Casi a principios de 2001, cuando se habían creado Centros Nacionales en la mayoría de los países participantes, se convocó una reunión de coordinadores nacionales del proyecto y el estudio PISA 2003 comenzó a prepararse activamente. Se invitó a los países a elaborar y a presentar ítems de la evaluación para su posible uso por parte del consorcio. Durante los meses siguientes, se continuaron elaborando ítems apartados de la prueba bajo la coordinación de ACER y con la participación de grupos para la elaboración de la prueba que pertenecían a organizaciones que formaban parte del consorcio internacional o que trabajaban en estrecha colaboración con él. Se elaboraron materiales preliminares que se sometieron a varias pruebas experimentales, se enviaron a los centros nacionales para estudiarlos, los revisaron los grupos de expertos de las áreas pertinentes los revisaron y, finalmente, se seleccionó y ultimó una serie de ítems para realizar una prueba piloto.

Se facilitaron a los países participantes las fuentes elaboradas por el contrastista internacional, tanto en inglés como en francés. Esto supuso que los países contasen con dos versiones equivalentes de referencia para elaborar sus versiones de la prueba y los instrumentos del cuestionario en la lengua local.

En 2002, se realizó una prueba piloto a gran escala en todos los países participantes. Esta prueba tenía dos objetivos. En primer lugar, se desarrollaron y comprobaron todos los procedimientos necesarios para el estudio definitivo, y los países experimentaron de forma directa la puesta en marcha de los mismos en un amplio número de centros. Esto incluía el muestreo de los centros, la negociación de la participación de los mismos, el muestreo de los alumnos, la negociación de la participación de éstos, la traducción de los materiales de la prueba a las correspondientes lenguas locales, la contratación y formación del personal local en los procedimientos exigidos para la administración de la prueba, la elaboración de todos los materiales de la prueba listos para su utilización en los centros, la puesta en marcha de los procedimientos de administración de la prueba en los centros seleccionados, la recogida y codificación de las respuestas de los alumnos, la recogida, depuración y procesamiento de los datos y la presentación de los mismos al contratista internacional para su análisis. En segundo lugar, los ítems de la prueba y del cuestionario seleccionados para la prueba piloto se aplicaron a un número considerable de alumnos y de centros de cada país, generando datos que se utilizaron posteriormente para determinar la calidad de la prueba y, por consiguiente, para contribuir a la elaboración de los ítems de la prueba final y del cuestionario, y a su selección para el estudio definitivo.

A lo largo de 2002 el contratista internacional recogió y procesó datos e información procedentes de la prueba piloto. Se analizaron y mejoraron los ítems de la prueba y del cuestionario a la luz de esos datos, completándose a finales de 2002 los instrumentos que habrían de utilizarse en el estudio definitivo. A continuación, se enviaron a todos los países participantes los principales instrumentos y materiales del estudio y, a su vez, se preparó al personal de cada Centro Nacional para el estudio definitivo que se llevó a cabo en 2003. Posteriormente, todos los datos recogidos en dicho estudio se analizaron en 2004 y se prepararon para la publicación de los resultados en diciembre de ese mismo año.

CARACTERÍSTICAS TÉCNICAS DE PISA

Las características técnicas de PISA tienen en cuenta distintos elementos. El diseño de la prueba y las características incorporadas a la prueba elaborada para PISA son rasgos fundamentales. El método de muestreo, que incluye los requisitos y procedimientos, tanto del muestreo de centros como del muestreo de alumnos, constituye una segunda área fundamental. Las características relacionadas con la naturaleza multilingüe de la prueba constituyen otra área que da origen a una serie importante de cuestiones técnicas, entre las que se encuentran las normas y procedimientos elaborados para garantizar la equivalencia de las versiones en los distintos idiomas que se utilizan en cada país participante y entre los mismos, a la vez que se tienen en cuenta los distintos contextos culturales de esos países. Los diferentes procedimientos de funcionamiento, entre los que se incluyen las disposiciones para la administración de la prueba, la recogida y el proceso de

datos, y los mecanismos para garantizar la calidad elaborados para asegurar que en todos los países se generan datos comparables, constituyen otra área centrada en los aspectos técnicos del proyecto. Naturalmente, a continuación, los detalles técnicos relativos al escalamiento y análisis de los datos y a su posterior presentación integran otro grupo importante de cuestiones. A su vez, se analizará brevemente cada una de estas áreas técnicas en los siguientes apartados. Se facilita mayor información respecto a cada una de estas áreas en el *Technical Report (Informe Técnico)* de PISA 2003 (OCDE, 2005).

Criterios rigurosos de calidad dentro de estas áreas se establecen, supervisan y garantizan a través de una serie de normas técnicas fijadas por el Consejo de Gobierno de PISA, y que constituyen el eje fundamental de la puesta en marcha del proyecto en cada país participante.

PLANIFICACIÓN Y DESARROLLO DE LA PRUEBA

Hasta el momento, PISA se ha llevado a cabo a través de pruebas que se complementan con papel y lápiz. Dentro del proyecto PISA se ha realizado una serie de estudios piloto en los que se han utilizado ordenadores. Sin embargo, las pruebas por ordenador aún no se han generalizado.

Los alumnos deben dedicar dos horas a la realización de la principal prueba de carácter «cognitivo» que cubre las áreas de Lectura, Ciencias, Matemáticas y, en PISA 2003, la Resolución de problemas. Además, los alumnos responden a un breve cuestionario diseñado para recopilar información complementaria relevante sobre sus características personales, opiniones, preferencias y aspiraciones, y algunas características sobre su entorno familiar y escolar. Este cuestionario está diseñado para realizarse en 20 ó 30 minutos. Por su parte, los directores de los centros docentes rellenan un breve cuestionario sobre aspectos más amplios del contexto escolar.

La parte de carácter cognitivo de la prueba debe cubrir adecuadamente cada área de la misma y generar datos relativos a los distintos *constructos* establecidos para cada uno de los marcos de la evaluación (OCDE, 2003). Esto incluye una serie de aspectos relacionados con el contenido de la materia de cada área y una serie de procesos de aprendizaje relevantes para cada área, debiendo emplearse distintos contextos para garantizar que ningún interés o experiencia particular esté con toda probabilidad injustamente sobrerrepresentada, de modo que la amplia variedad de experiencias de los alumnos esté suficientemente representada en la prueba. Las pruebas PISA están diseñadas para medir hasta qué punto los alumnos pueden utilizar los conocimientos y destrezas que han adquirido en la escuela al tratar de resolver los tipos de problemas a los que se enfrentarán en contextos no escolares. El énfasis no se pone en la evaluación de determinados resultados curriculares, sino en la aplicación de los conocimientos adquiridos en distintos contextos de la vida real.

Hasta la fecha, PISA ha adoptado un proceso de diseño de la prueba que, en cada ciclo del estudio, convierte a una de las áreas de la misma en el *área*

«principal» y, al resto, en áreas «secundarias». En el diseño de la prueba se ha asignado mayor cantidad de tiempo al área principal, para hacer posible una evaluación más minuciosa de aspectos del área en ese ciclo del estudio. La asignación del tiempo para las áreas secundarias sigue siendo suficiente para que de ahí puedan derivarse indicadores significativos del rendimiento. El área principal va rotando: la lectura lo fue en el ciclo del estudio PISA 2000, matemáticas en PISA 2003 y ciencias lo será en PISA 2006. Por tanto, en cada ciclo de la evaluación, se elabora una gran cantidad de material para la prueba con el fin de garantizar la adecuada cobertura y equilibrio en los distintos *constructos* y aspectos de cada área. Este material se distribuye en una serie de cuadernillos de acuerdo con un diseño de rotación de la prueba y los cuadernillos, a su vez, se distribuyen aleatoriamente entre los alumnos de la muestra.

En PISA 2003, los 167 ítems de la prueba del estudio definitivo se distribuyeron entre trece series de ítems a las que se denominó «*grupos de ítems*» (siete grupos de matemáticas y dos grupos en cada una de las otras áreas: lectura, ciencias y resolución de problemas), y donde cada grupo representaba 30 minutos del tiempo de la prueba. Los ítems se presentaron a los alumnos en trece cuadernillos, cada uno de ellos formado por cuatro grupos según el esquema que se muestra en la Tabla II. M1-M7 hace referencia a los grupos de matemáticas, R1 y R2 a los grupos de lectura, S1 y S2 a los de ciencias, y PS1 y PS2 a los de resolución de problemas. Cada grupo aparece exactamente una vez en cada una de las cuatro posiciones posibles dentro de un cuadernillo. Por tanto, cada ítem de la prueba aparecía en cuatro de los cuadernillos de la misma. Este diseño de rotación y de interconexiones permitía controlar cualquier consecuencia relativa al orden y aplicar técnicas de medición estándar a las respuestas dadas por los alumnos para calcular las dificultades de los ítems y las competencias de los alumnos.

TABLA II

Diseño de rotación de los grupos utilizado en la elaboración de los cuadernillos de la prueba para PISA 2003

Cuadernillo	Grupo 1	Grupo 2	Grupo 3	Grupo 4
1	M1	M2	M4	R1
2	M2	M3	M5	R2
3	M3	M4	M6	PS1
4	M4	M5	M7	PS2
5	M5	M6	S1	M1
6	M6	M7	S2	M2
7	M7	S1	R1	M3
8	S1	S2	R2	M4
9	S2	R1	PS1	M5
10	R1	R2	PS2	M6
11	R2	PS1	M1	M7
12	PS1	PS2	M2	S1
13	PS2	M1	M3	S2

Los ítems de la prueba se presentaron en varios formatos diferentes: elección múltiple, respuesta corta y respuesta larga. Los ítems de elección múltiple, o bien eran estándar, con un número limitado de respuestas (normalmente cuatro), de las que los alumnos debían elegir la correcta, o bien de elección múltiple compleja, con varias afirmaciones para cada una de las cuales los alumnos debían elegir una de varias respuestas posibles (verdadero/falso, correcto/incorrecto, etc.). Los ítems de respuesta corta incluían tanto ítems de respuesta construida cerrada, que generalmente exigían a los alumnos la elaboración de una respuesta dentro de unas restricciones muy limitadas (por ejemplo, los ítems de matemáticas que requieren una respuesta numérica), como los que exigían una palabra o una frase corta. Los ítems de respuesta corta eran similares a los de respuesta construida cerrada, aunque para éstos existía una mayor variedad de respuestas. Los ítems de respuesta construida abierta exigían una redacción más amplia, realizar un cálculo o mostrar un razonamiento lógico, y con frecuencia incluían alguna explicación o justificación.

En los anteriores ciclos de PISA, los organizadores de la prueba y los expertos en las áreas manifestaron con acierto que la variedad de procesos cognitivos que pueden mostrarse y aprovecharse es mucho mayor si se utilizan formatos abiertos que si se usan únicamente ítems de formato cerrado, como los de elección múltiple o los de respuesta corta, lo que confirmaron los comentarios naciona-

les que se realizaron sobre los ítems durante su fase de elaboración. El Consejo de Gobierno de PISA ha llegado a la conclusión de que los costes adicionales que implica la codificación y el procesamiento de las respuestas de estos ítems más abiertos están justificados por el mayor valor y riqueza de los datos que se derivan del uso de una gama más amplia de formatos de los ítems de la prueba.

Las respuestas de los alumnos a más de la mitad de los ítems del estudio definitivo PISA 2003 pudieron procesarse por ordenador. El resto, un total de 79 de los 167 ítems (es decir, el 47%), requirió la intervención de un codificador cualificado para procesar las respuestas de los alumnos.

Se escogió una batería típica de ítems de cuestionario para los cuestionarios de contexto. La finalidad de estos cuestionarios de contexto era identificar los factores sociales, culturales, económicos y educativos asociados con el rendimiento de los alumnos. Esto permitiría explorar las relaciones entre los resultados del rendimiento de éstos en las pruebas de carácter cognitivo y los distintos factores relativos a los alumnos y a los centros, y ver cómo estos factores podrían variar entre los distintos sistemas y países.

EL MUESTREO EN PISA

El muestreo en PISA se realiza en dos etapas, según un procedimiento diseñado para asignar a todos los alumnos de cada país participante, que cumplen los requisitos para ser elegidos, una probabilidad conocida de ser seleccionados para participar. En primer lugar, los centros con este tipo de alumnos se seleccionan al azar con una probabilidad proporcional a su tamaño. A continuación, se selecciona al azar a 35 de esos alumnos de cada centro de la muestra. Así, los estudiantes de las pruebas PISA son seleccionados al azar y representan realmente a la población de alumnos de 15 años escolarizados en cada país participante.

De cada país se toma como muestra un mínimo de 150 centros (o la totalidad de los mismos si son menos de 150 los que cuentan con alumnos que cumplen los requisitos para ser seleccionados), y más de 150, si existen centros pequeños en el país. El número fijado para la muestra de alumnos es de 35 por cada centro, lo que indica que se tomará como muestra un mínimo de 5.250 alumnos por país, con la expectativa de que se evalúe a un mínimo de 4.500. Si son menos de 35 alumnos los que están disponibles en un número bastante elevado de centros, entonces, se toman como muestra otros centros para garantizar un tamaño total mínimo adecuado de la muestra de alumnos.

Para determinar con exactitud los alumnos que iban a incluirse en el marco de muestreo, los países tuvieron que escoger un periodo de siete semanas entre el 1 de marzo y el 31 de agosto de 2003 (denominado *ventana de pruebas*) en el que administrarían la prueba. A continuación, se aplicaron normas claramente especificadas para determinar qué alumnos del país pertenecerían al intervalo de edad exigido y, por tanto, podrían ser seleccionados para realizar las pruebas durante ese periodo.

Asimismo, se aplicaron normas para garantizar la adecuada cobertura de la población que cumple los requisitos necesarios para ser seleccionada (lo que incluía normas estrictas sobre los centros y alumnos que podrían ser excluidos de forma legítima), la suficiente exactitud y precisión de los cálculos derivados de la muestra (lo que incluía normas estrictas sobre el tamaño exigido de la muestra) y unas tasas de respuesta aceptables tanto para los centros como para los alumnos (lo que incluía normas estrictas sobre las tasas de respuesta y los procedimientos para hacer uso de los centros suplentes allí donde fuese necesario alcanzar unas tasas de respuesta de los centros aceptables, y normas de decisión sobre la inclusión o exclusión de los datos de los alumnos en función de las tasas de respuesta de éstos).

TRADUCCIÓN E IDONEIDAD CULTURAL DEL MATERIAL DE PISA

PISA es el mayor estudio internacional de su género que jamás haya existido. PISA 2003 incluía la administración de la prueba en al menos 150 centros distintos de cada uno de los 42 países participantes, lo que implicaba 55 versiones nacionales distintas de los instrumentos de evaluación en 33 idiomas diferentes. Ante semejante empresa, la necesidad de garantizar la comparación del material de la prueba en todas las administraciones de la misma no ha sido una cuestión irrelevante. Para lograrlo, el primer paso fue garantizar que los propios materiales de la prueba fueran los adecuados para utilizar en contextos tan distintos culturalmente, y que las diferentes versiones empleadas en los distintos contextos fueran equivalentes.

El enfoque para garantizar la idoneidad cultural implica, en primer lugar, la utilización de una amplia variedad de materiales que representen las distintas experiencias y contextos culturales y, a continuación, el procesamiento y la depuración de dichos materiales para garantizar el correcto equilibrio de los distintos intereses y el buen funcionamiento de todos los materiales en el conjunto de países. Entre los mecanismos empleados para garantizar que los materiales elaborados para PISA son culturalmente adecuados se incluyen los siguientes:

- Los materiales de la prueba (ítems del cuestionario y de la prueba) se buscan entre el mayor número posible de fuentes, incluida la búsqueda de propuestas nacionales en todos los países participantes y la utilización de procedimientos para el desarrollo de la prueba que son supervisados por expertos internacionales en el campo pertinente, y gestionados por expertos de distintos países y contextos culturales.
- Se proporcionan diversas oportunidades a todos los países participantes para analizar y comentar el material en proceso de elaboración, que también es analizado por paneles de expertos internacionales en cada área de desarrollo.
- Se llevan a cabo actividades de laboratorio de carácter cognitivo con el material en proceso de elaboración, utilizando con la participación de alumnos reales en varios países diferentes.

- Se realiza una prueba piloto formal a gran escala con alumnos de todos los países participantes para comprobar el funcionamiento del material en proceso de elaboración.
- Los resultados del análisis estadístico de los datos de la prueba piloto se emplean para evaluar empíricamente el material tal y como se aplica en todos los países y para detectar ejemplos de ítems de la prueba cuyo comportamiento sea diferente en los distintos países.
- Se revisa todo el material en función de la información recibida de cada uno de estos diferentes mecanismos y se selecciona sólo el material que se ha demostrado que funciona, se selecciona para su utilización en la evaluación definitiva de PISA.
- Los coordinadores nacionales del proyecto analizan la decisión tomada sobre el material que finalmente se selecciona y el Consejo de Gobierno de PISA la refrenda.

Paralelamente a la elaboración del material culturalmente idóneo para utilizar en los instrumentos de la evaluación PISA, se preparan versiones fuente de todo el material, tanto en inglés como en francés, que son las precursoras de la elaboración de versiones nacionales equivalentes. A partir de estas dos versiones fuente, los centros nacionales de cada país participante elaboran sus propias versiones del material de la prueba, utilizando un proceso rigurosamente controlado de traducción y de verificación independiente, internacional y especializada de esas traducciones.

Se lleva a cabo una exhaustiva prueba piloto para comprobar los materiales traducidos de cada país y se emplean los análisis de los datos de dicha prueba para evaluar empíricamente la equivalencia de las distintas versiones lingüísticas y nacionales. Se revisan los materiales en función de los análisis de los datos de la prueba piloto y se lleva a cabo la selección final para el estudio definitivo de PISA, con el fin de garantizar que sólo se seleccionan los ítems de la prueba que funcionan plenamente en todas las versiones de los distintos países y lenguas. Al finalizar este proceso, cada versión nacional elaborada es lingüísticamente equivalente a las versiones fuente.

OPERACIONES DE CAMPO

El segundo paso más importante para garantizar la comparación de los resultados de la prueba en un conjunto tan heterogéneo de países y contextos es la estandarización de los procedimientos de administración de la misma. Se ha desarrollado y documentado una amplia gama de procedimientos para ayudar a los países participantes en las sesiones de administración de la prueba PISA, de forma que se facilite la generación de datos que sean comparables a escala internacional.

En el manual del Coordinador Nacional del proyecto se describen todos los procedimientos que cada Centro Nacional ha de establecer y poner en marcha, incluida la participación en varios procedimientos de consulta y análisis, la apli-

cación de los procedimientos de muestreo y de todos los procedimientos relativos a la preparación, elaboración y envío de los materiales de la prueba, la contratación y formación del personal responsable de la administración de la prueba y de la supervisión de la misma, la ayuda para poner en marcha procedimientos de supervisión control de la calidad, la contratación y formación del personal para codificar las respuestas de los alumnos, la gestión de la codificación de dichas respuestas y el registro de los datos de la respuesta de los alumnos, los procesos relativos a la recogida y elaboración de todos los datos de PISA para presentarlos al contratista internacional, y los posteriores procesos que faciliten el análisis de los datos y la presentación de los resultados.

Existen distintos manuales para cubrir procedimientos operativos concretos relacionados con el muestreo, la traducción, la administración de la prueba, la coordinación de los centros, la codificación de las respuestas de los alumnos, la gestión de los datos, y los relacionados con los programas informáticos especializados de recogida de datos utilizados en el proyecto. Además de esta amplia documentación, el contratista internacional mantiene varias reuniones con los coordinadores nacionales del proyecto y con otro personal clave de los centros nacionales. Uno de los principales objetivos de estas reuniones es proporcionar información y formación sobre las operaciones de campo y todos los demás aspectos relativos a la puesta en marcha del proyecto en cada país.

A través de la amplia documentación sobre las operaciones de campo, la formación y reuniones informativas previstas, y los contactos periódicos por correo electrónico o teléfono, el contratista internacional se asegura de que los complejos procedimientos de PISA puedan aplicarse sistemáticamente en todos los países participantes. Para comprobar el grado de aplicación sistemática de estos procedimientos se ponen en marcha distintos procedimientos de supervisión de la calidad que se describen en el siguiente apartado.

EL CONTROL DE LA CALIDAD

La puesta en marcha del proyecto PISA se desarrolla en torno a una serie de criterios de calidad relativos a los distintos aspectos del proyecto. Existen normas relacionadas con la definición de la población destinataria, el muestreo, el idioma de las pruebas, la elaboración de las mismas y de los manuales, la administración de la prueba, la legibilidad de la letra de los materiales de la prueba, la seguridad de éstos, etc.

Para cada área se establece un plan en el que se describe de qué forma cada país participante ha de cumplir dicha norma. A medida que el proyecto avanza, se van generando datos que muestran si se ha cumplido esa norma. Existen técnicas de control a través de las cuales los centros nacionales de PISA garantizan continuamente la calidad de la puesta en marcha del proyecto, y se aplican procedimientos externos de supervisión que permiten la intervención temprana por parte del contratista internacional en caso de que un país participante concreto parezca que no va a poder cumplir alguna de las normas, o la intervención del

contratista una vez que se ha determinado si los datos recogidos pueden utilizarse con total seguridad para los fines establecidos.

Por último, se desarrolla un proceso de evaluación a través del cual se llega a una resolución final sobre si los datos de un país participante están en condiciones de utilizarse y, por tanto, pueden incluirse en la base de datos internacional. Si se considera que no son adecuados para su utilización, entonces los datos de un país se excluyen de alguno o de todos los informes internacionales sobre los resultados de PISA. En el proceso de evaluación participa el personal de alto rango del Consorcio, miembros del grupo técnico asesor, incluido el consultor para el muestreo, y el Secretariado de la OCDE. Este grupo proporciona recomendaciones al Consejo de Gobierno de PISA.

A continuación se analizan brevemente dos ejemplos de control de la calidad: los relativos a los métodos de muestreo y los relacionados con los procedimientos de administración de la prueba.

Los métodos de muestreo se desarrollan y documentan a través de las negociaciones entre cada Centro Nacional y el personal de muestreo del contratista internacional, bajo la dirección del consultor internacional para el muestreo. La información sobre la observancia de esos métodos y de las normas de muestreo por parte de cada país se recogen a través de la documentación de los procedimientos de muestreo aplicados en cada país, y los datos principales se cotejan con otra información disponible, como son las estadísticas de población publicadas sobre la cohorte de PISA en cada país. Las tasas de respuesta alcanzadas, tanto en los centros como por los alumnos, proporcionan datos definitivos y explícitos sobre los resultados del muestreo, que se introducen directamente en el proceso de evaluación de los datos.

La calidad de los procedimientos de administración de la prueba en los distintos países participantes se supervisa a través de una serie de mecanismos que, en líneas generales, se pueden clasificar en dos estrategias complementarias. La primera estrategia consiste en garantizar que todos los países reciban las mismas instrucciones relativas a un conjunto estándar de procedimientos de administración de la prueba y la misma formación con relación a dichos procedimientos. Las instrucciones incluyen manuales detallados de funcionamiento que cubren distintos aspectos de la aplicación del proyecto en general y de la administración de la prueba en particular. La comprobación de las versiones nacionales de estos manuales forma parte del proceso de verificación internacional. La formación incluye sesiones dedicadas a aspectos relacionados con la aplicación del proyecto y de la prueba durante las reuniones de los coordinadores nacionales del proyecto y durante las sesiones de formación sobre las operaciones que se celebran antes de que comience la administración de la prueba. La segunda estrategia consiste en formular observaciones al azar de las sesiones de administración de la prueba en cada país, para garantizar la aplicación sistemática de los procedimientos estándar. El contratista internacional designa y forma a los supervisores de Calidad PISA en cada país y a continuación éstos visitan una serie de sesiones de administración de la prueba, sin previo

aviso, para realizar la pertinente observación. Cualquier centro de la muestra puede ser objeto de estas visitas. Los supervisores de Calidad observan y elaboran un informe sobre distintos aspectos de las sesiones de la prueba según un protocolo estandarizado y presentan sus informes al contratista internacional. Posteriormente, la información incluida en esos informes se utiliza como parte del proceso de evaluación de datos.

ESCALAMIENTO Y ANÁLISIS DE LOS DATOS

El escalamiento de los datos de PISA se sustenta sobre la premisa elemental de que existe algún rasgo latente o conjunto de rasgos latentes de interés (definidos en los marcos de la evaluación) que forman un continuo o escala, y que pueden elaborarse ítems de la prueba que incluyan ese rasgo latente (mediante la exigencia de distintas cantidades de ese rasgo), y que la cantidad del rasgo presente en los alumnos puede calcularse observando sus respuestas a los ítems de la prueba. Normalmente, se piensa en el continuo como en una línea, con «más» presencia del rasgo en una dirección y «menos» en la otra; también se piensa tiene en cuenta que los ítems de la prueba pueden situarse a lo largo de la línea en función de la cantidad de rasgo latente que exigen, y que los alumnos pueden también establecerse a lo largo de la línea en función de la cantidad de rasgo latente que presentan.

PISA utiliza un tipo de modelo de respuesta al ítem ampliamente aceptado para escalar los datos de los alumnos y obtener las distintas medidas de comparación que la OCDE elabora y difunde. Se trata de un «*modelo logit multinomial de coeficientes mixtos*», es decir, una generalización del modelo de Rasch. Básicamente utiliza las respuestas de los alumnos a un conjunto de ítems de la prueba para obtener al mismo tiempo cálculos aproximados de la «*dificultad*» de los ítems de la prueba y de la «*capacidad*» de los alumnos de la muestra, lo que permite, tanto a los ítems de la prueba como a los alumnos, situarse a lo largo del continuo del rasgo latente que se pretende medir. Para más información sobre el modelo utilizado y los métodos de escalamiento aplicados puede consultarse el *PISA 2003 Technical Report (Informe Técnico PISA 2003)* (OCDE, 2005).

En PISA, la aplicación de estos modelos también permite relacionar las respuestas observadas en los alumnos a los ítems de la prueba con diversas variables de contexto (género, situación socioeconómica, etc.) por medio de ecuaciones de regresión. Uno de los resultados de estos análisis son los cálculos aproximados de promedios de población y de otras estadísticas que permiten la comparación entre grupos de alumnos de las poblaciones de la muestra de PISA. Y son esas comparaciones entre grupos de alumnos, que relacionan el rendimiento en una o más áreas de la prueba de carácter cognitivo con diversas variables de contexto, las que permiten a los investigadores orientados hacia la política analizar los factores que pueden influir en el rendimiento de los alumnos y tener en cuenta cualquier implicación para la gestión de los centros u otras características organizativas de los mismos, prácticas de enseñanza-aprendizaje, etc.

El diseño de los instrumentos de la prueba PISA, junto con el escalamiento y análisis de los datos, también permite el seguimiento de las tendencias en los resultados de la prueba en todos los ciclos de la evaluación PISA. Asimismo, parte del material de la prueba de PISA 2000 se empleó en PISA 2003 y, de la misma manera, existe material de la prueba común a las evaluaciones PISA 2003 y PISA 2006. Esto permite obtener mediciones de las tendencias, a través de las cuales pueden supervisarse los cambios ocurridos en el tiempo. Podría esperarse que dichos cambios fuesen insignificantes de un ciclo a otro, sin embargo, puesto que PISA continúa a lo largo de otros ciclos, puede esperarse que los ajustes a los contextos políticos y a las prácticas docentes en los distintos países lleven a cambios más notorios en los resultados de la prueba PISA en periodos de tiempo más largos. En la primera prueba PISA (2000), la principal área de evaluación fue la lectura. Se espera que cualquier cambio en los resultados de lectura, relacionado con las modificaciones llevadas a cabo en la política y en la práctica como respuesta a los resultados de PISA 2000, afecte a los resultados de 2009, cuando de nuevo vuelva a ser la lectura la principal área de evaluación.

PRESENTACIÓN DE LOS RESULTADOS DE PISA

Tras la recogida de datos en cada ciclo de la evaluación PISA de los realizados hasta ahora, la OCDE ha elaborado un informe exhaustivo que recoge los resultados más relevantes desde una perspectiva internacional. Se han elaborado dos de estos informes, uno sobre los resultados de PISA 2000 (OCDE, 2001) y otro sobre los de PISA 2003 (OCDE, 2004).

Estos importantes informes internacionales han proporcionado una perspectiva general del proyecto PISA, detalles que ayudan a comprender las características fundamentales de la principal área de evaluación (lectura, en el informe de 2000, y matemáticas, en 2003) y a cómo deben interpretarse los resultados con relación al marco conceptual; las comparaciones entre países de los resultados en la principal área de evaluación y su relación con algunas de las variables clave de contexto de los alumnos; una perspectiva general de los resultados obtenidos en las áreas de evaluación secundarias; una perspectiva general de la información obtenida de las distintas variables de contexto recogidas con relación a los alumnos y a los centros, y las diferencias que presentan entre los países; comparaciones internacionales de aspectos relativos al entorno de aprendizaje y a la organización de la enseñanza; y varios debates sobre las implicaciones para la política de estos distintos aspectos del estudio.

En la presentación de los resultados relativos a la competencia, dentro de la principal área de evaluación, se ha enfatizado la descripción de los resultados de los alumnos en cada país con relación a las escalas y subescalas que proceden del correspondiente marco conceptual. En el caso de la competencia matemática se presentan los resultados de una escala de competencia matemática global, así como de subescalas basadas en las cuatro áreas «*de contenido*» del marco de matemáticas. Un aspecto fundamental para la descripción de los resultados de los alumnos es un conjunto de especificaciones de aquello que los alumnos

situados en distintos puntos a lo largo de la escala de competencia serían capaces de hacer normalmente. Las descripciones de la competencia matemática proporcionan una idea clara del modo en que los alumnos son capaces de utilizar las distintas competencias matemáticas descritas en el marco conceptual, explicando el aumento de la competencia matemática en relación con una capacidad cada vez mayor de los alumnos para demostrar y utilizar dichas competencias. El principal informe internacional de la OCDE concede especial importancia a los porcentajes relativos de alumnos de cada país que rinden a distintos niveles del espectro de competencias.

En la presentación de las variables de contexto (alumno, familia, centro y factores del sistema) se describen las diferencias existentes entre los países con relación a esas variables, y se analiza la relación de algunas de esas variables con los resultados cognitivos de los alumnos para tratar de describir de qué modo estos factores operan de manera diferente en los distintos países.

Además de los principales informes internacionales, la OCDE elabora o promueve una serie de informes adicionales. Tras el primer ciclo de la evaluación, se publicó un informe completo sobre los resultados de la evaluación de la lectura (OCDE, 2002). La OCDE también ha publicado otros informes sobre aspectos concretos de los resultados de la evaluación PISA 2000, como por ejemplo la participación de los alumnos, sus actitudes hacia el estudio, los factores escolares relacionados con la calidad y la equidad, etc. Asimismo, la OCDE publica una serie de documentos más técnicos entre los que se encuentran el informe técnico, un manual de la base de datos, los ítems de la muestra y los documentos marco. La OCDE también ha publicado o está preparando una serie similar de informes relativos a los resultados del ciclo del estudio PISA 2003. Se puede acceder fácilmente a ésta y otras publicaciones de PISA a través del sitio web de la OCDE (<http://www.pisa.oecd.org/>).

La mayoría de los países participantes en PISA elaboran sus propios informes nacionales prestando mayor atención a distintos resultados obtenidos en el país o, al menos, dando una perspectiva nacional más clara de los resultados y proporcionando, en muchos casos, análisis e interpretaciones más detalladas que tienen en cuenta distintos factores que operan en el país concreto.

CARACTERÍSTICAS DE LAS MATEMÁTICAS EN PISA

El campo que se evalúa en cada una de las áreas de la prueba PISA se define a través de una serie de marcos conceptuales publicados por la OCDE (OCDE, 2003). El marco de las matemáticas proporciona una definición y una explicación de la competencia matemática, y describe el área objeto de evaluación a través de tres componentes: las *situaciones* o *contextos* en los que se localizan los problemas; el contenido matemático que ha de usarse para resolverlos y que se organiza en cuatro grandes categorías fenomenológicas denominadas *ideas principales*; y los tipos de procesos matemáticos o *competencias* que los alumnos deben activar para conectar el mundo real, en el que se localizan los problemas, con las matemáticas y de ese modo resolverlos.

COMPETENCIA MATEMÁTICA

La competencia matemática se define como

«(...) la capacidad de un individuo para identificar y comprender el papel que desempeñan las matemáticas en el mundo, realizar juicios bien fundados y utilizar e implicarse con las matemáticas de modo que se satisfagan las necesidades de la vida de ese individuo como ciudadano constructivo, comprometido y reflexivo» (OCDE, 2004, p.24).

Por tanto, la evaluación de matemáticas en PISA no tiene como principal foco de interés los conocimientos curriculares, la adquisición de un conjunto de destrezas matemáticas básicas o el rendimiento correspondiente al curso escolar actual. Está diseñada para medir hasta qué punto los alumnos pueden hacer uso de sus conocimientos en situaciones reales de la vida diaria. Se interesa por el alcance de los conocimientos y de la comprensión de los alumnos y, más en concreto, por descubrir hasta dónde pueden activarse esos conocimientos en caso necesario. Por tanto, PISA se centra en la capacidad de los alumnos para analizar, razonar y comunicarse de manera eficaz según van planteando, resolviendo e interpretando problemas matemáticos en diversas situaciones que incorporan conceptos matemáticos cuantitativos, espaciales, probabilísticos o de otro tipo. El informe internacional de los resultados de PISA 2003 (OCDE, 2004) proporciona una base útil para dicho enfoque:

Durante gran parte del siglo pasado, el contenido de los currículos escolares de matemáticas y ciencias estuvo dominado por la necesidad de proporcionar las bases de la formación profesional de un reducido número de matemáticos, científicos e ingenieros. Sin embargo, el papel creciente de la ciencia, las matemáticas y la tecnología en la vida moderna hace que los objetivos de satisfacción personal, empleo y participación plena en la sociedad exijan cada vez más que todos los adultos, y no sólo los que aspiran a una carrera científica, sean competentes desde el punto de vista matemático, científico y técnico (p. 37).

CONTENIDO MATEMÁTICO—LAS IDEAS PRINCIPALES

L. A. Steen y otros (Steen, 1990) han influido en la definición del contenido matemático del área de evaluación del enfoque PISA. Se emplean cuatro categorías fenomenológicas que el marco para las matemáticas describe también como representantes de las principales tendencias en la evolución histórica de las matemáticas (OCDE 2003, p. 35). Las áreas de contenido seleccionadas reciben el nombre de *ideas principales* y se denominan *cantidad, espacio y forma, cambio y relaciones e incertidumbre*. Las categorías engloban numerosos fenómenos matemáticos y los modos en que se presentan en el mundo circundante. También incluyen los temas matemáticos normalmente presentes en los currículos escolares de matemáticas de los alumnos hasta la edad de 15 años.

SITUACIONES Y CONTEXTOS

Casi todos los ítems de matemáticas de PISA reflejan una situación del mundo real, aunque también se utilizan situaciones hipotéticas, así como un número muy reducido de contextos puramente matemáticos. Entre las situaciones en las que se establecen los ítems de la prueba PISA se encuentran aquellas cercanas a la experiencia inmediata del alumno, denominadas situaciones *personales*; otras ligeramente apartadas con las que cualquiera podría encontrarse como resultado de su trabajo u otra actividad diaria organizada, denominadas *educativas/ocupacionales*; algo más alejados se encuentran los contextos en los que están presentes el entorno y la comunidad del alumno, que reciben el nombre de situaciones *públicas*; y, por último, las llamadas situaciones *científicas*, que incluyen aplicaciones más abstractas y técnicas de las matemáticas, como los contextos puramente matemáticos que con frecuencia se estudian en las clases de matemáticas. El objetivo de PISA es utilizar contextos de problemas para los que los problemas y las soluciones deben tener un verdadero interés (al menos para alguien) y requerir la utilización de las matemáticas como parte de la solución. El marco de matemáticas de PISA denomina a estos contextos, contextos de problemas *auténticos*, lo que contrasta con aquellos en los que nadie estaría interesado realmente a no ser que fuese como excusa para utilizar un aspecto concreto de las matemáticas.

PROCESOS MATEMÁTICOS—LAS COMPETENCIAS

Las competencias matemáticas constituyen el núcleo de la alfabetización matemática y PISA les ha dado una importancia capital en la comprensión y descripción del desarrollo de la competencia matemática. El marco de matemáticas de PISA, basado en los trabajos de Mogens Niss y sus colegas daneses (Niss, 1999), describe ocho competencias: pensamiento y lógica; argumentación; comunicación; construcción de modelos, planteamiento y resolución de problemas; representación; lenguaje y operaciones formales y técnicas; y uso de soportes y herramientas. Estas competencias se activan de forma diferente y a distintos niveles a medida que los alumnos se implican en diferentes problemas, y los niveles de activación requeridos están íntimamente relacionados con el hecho de que un problema implique únicamente la *reproducción* del conocimiento practicado, forjando *conexiones* entre los distintos elementos del conocimiento, o con el hecho de asociar los distintos aspectos de un problema para encontrar una solución, o *reflexionar* sobre ellos y su contexto para buscar de manera creativa y estratégica dicha solución y, tal vez, para justificar y explicar los resultados. Es evidente que los alumnos poseen distintas capacidades para activar estas competencias a distintos niveles y son fundamentalmente esas diferencias en las capacidades lo que lleva a las variaciones observadas en el rendimiento de los alumnos en la evaluación PISA.

Una de las formas de presentar los resultados de PISA es a través del porcentaje de alumnos de cada país que rinden en cada uno de los seis niveles de com-

petencia descritos. Estos seis niveles se describen en el informe internacional de la OCDE (OCDE, 2004, p. 47 y siguientes) para el conjunto de la competencia matemática en PISA y, a continuación, algo más detalladamente para cada una de las cuatro áreas de contenido. Esas descripciones de competencia se basan en gran medida en la forma en que los alumnos en niveles cada vez más altos pueden activar las distintas competencias matemáticas a medida que aumenta su competencia. Por ejemplo, la competencia de *representación* es especialmente importante para los alumnos cuando relacionan sus conocimientos matemáticos con la comprensión de los contextos del problema. Esta competencia de *representación* supone descodificar y codificar representaciones de objetos matemáticos y cambiar entre distintas representaciones en caso de que resulte útil para exponer los aspectos de un problema. En el nivel más bajo, se observa que los alumnos manejan únicamente las representaciones más sencillas y directas (información expresada de forma explícita). En un nivel ligeramente superior, los alumnos pueden manejar información no expresada de forma tan directa (pueden identificar la información relevante en una representación que les resulte familiar, como un texto, diagrama o gráfico). En un nivel superior, los alumnos pueden hacer uso de distintas representaciones, interpretando información, por ejemplo, en un texto o diagrama o en ambos. En niveles aún más altos los alumnos pueden identificar y seleccionar información relevante de distintas representaciones, integrarla y, tal vez, relacionarla con seguridad con el mundo real. No obstante, las destrezas de representación de niveles más altos incluirían la utilización estratégica y planificada de distintas representaciones, la traducción flexible entre una serie de representaciones distintas y la capacidad de evaluar los puntos fuertes relativos de distintas representaciones.

ALGUNOS RESULTADOS DE PISA

En este último apartado se presenta una brevísima perspectiva general de algunos resultados y conclusiones de PISA 2003 que son de particular interés para el autor. En algunos casos, y cuando ha sido necesario, se han puesto de relieve los resultados de España.

ALGUNAS CONCLUSIONES RELATIVAS A LAS CARACTERÍSTICAS CONTEXTUALES

El informe internacional pone de relieve una serie de conclusiones relacionadas con las características contextuales de los alumnos, los centros y los sistemas escolares, que podrían ser importantes al tratar de interpretar los resultados de la prueba PISA de los distintos países o de distintos grupos de alumnos dentro de un país.

Por ejemplo, se adquieren algunos conocimientos de las características aportadas por los alumnos eficaces. Éstos suelen estar bastante motivados: tienen un alto interés por sus experiencias de aprendizaje y disfrutan de ellas, creen firmemente en los posibles futuros beneficios de la escolarización, suelen tener sentimientos positivos hacia su centro educativo y un sentido de pertenencia al mismo. Asimismo, los alumnos eficaces tienen normalmente fe en sus propias capacidades:

confían en su capacidad matemática y creen en su capacidad para superar desafíos. Los alumnos eficaces tienen también un alto grado de fuerza emocional y una notable ausencia de ansiedad respecto a las matemáticas. Los alumnos eficaces adoptan, normalmente, una serie de estrategias eficaces de aprendizaje, tales como las dirigidas a la memorización y práctica de sus conocimientos o las estrategias de elaboración de dichos conocimientos relacionando los nuevos con los antiguos. Asimismo, suelen desplegar todo un repertorio de estrategias de control, como por ejemplo la planificación, la adaptación al cambio, la comprobación de los objetivos y el aprendizaje autónomo.

Los datos de PISA proporcionan gran cantidad de información sobre el grado de las diferencias existentes entre los países en relación con la distribución de estas características aportadas por los alumnos. Los datos indican que los alumnos españoles tienen en promedio un sentido relativamente alto de pertenencia. Por otro lado, también suelen tener un nivel relativamente elevado de ansiedad frente a las matemáticas, lo cual es particularmente habitual en el caso de las chicas.

Si se tiene en cuenta la relación entre estas características y los niveles de competencia matemática, los datos de PISA indican que el interés y el placer están vinculados a un alto rendimiento en matemáticas, no tanto de manera directa, sino porque aumentan la probabilidad de que los alumnos desarrollen y hagan uso de otras estrategias eficaces de aprendizaje. La confianza en uno mismo también está relacionada con un rendimiento más alto, aunque en este caso el género tiene un peso considerable, pues las chicas tienden a expresar con mayor frecuencia más falta de confianza en sí mismas que los chicos. Además, la ansiedad frente a las matemáticas suele ser incompatible con un rendimiento alto en esta materia, mientras que la falta de ansiedad junto con el interés y el placer aumenta la probabilidad de que los resultados de rendimiento en matemáticas sean mejores. Los datos indican que los alumnos españoles se benefician más que los de muchos otros países del uso de distintas estrategias de control en su aprendizaje, las cuales pueden, hasta cierto punto, emplearse como una respuesta de compensación a los niveles relativamente elevados de ansiedad frente a las matemáticas.

Asimismo, existen pruebas de que algunos centros educativos son más eficaces que otros a la hora de promover actitudes positivas hacia las matemáticas y hacia otras características del aprendizaje asociadas con un mayor rendimiento en este área, y que la influencia de la escuela puede, tanto aumentar las características positivas de los alumnos, como disminuir el impacto de las negativas.

No obstante, un descubrimiento claro es la existencia de una mayor variación de estas características entre los alumnos de los centros que entre los propios centros, y que tanto estos últimos como el profesorado necesitan hacer frente a esta heterogeneidad.

Una conclusión fundamental que se obtiene de estas consideraciones es la importancia que para los centros tiene el abordar tanto las actitudes y conduc-

tas de aprendizaje como las materias académicas, y más aún, que estas actitudes y conductas de aprendizaje son importantes para todos los alumnos, no sólo para los que presentan un rendimiento más bajo. Esto afecta al funcionamiento de los centros y probablemente también a la gestión de la formación del profesorado y a su perfeccionamiento profesional continuo.

Si se analizan brevemente los factores contextuales a escala de centros, los datos de PISA muestran una enorme diferencia geográfica y regional en las medias relativas al agrupamiento de los alumnos. Parece claro que la calidad de los centros varía a la par que las variaciones de recursos y perfiles de las poblaciones escolares. Los datos de PISA llevarían a la conclusión de que el aumento de las diferencias entre los centros está asociado a la baja equidad y a los relativamente malos resultados globales, especialmente para los alumnos de entornos sociales desfavorecidos. España muestra una diferencia entre centros relativamente pequeña, en lo que parece ser una gran ventaja derivada de sus, hasta cierto punto, propuestas políticas de diferenciación de centros y programas. En España, el entorno social explica únicamente el 14% de las diferencias del rendimiento en matemáticas, presentando un gradiente social relativamente bajo. Es decir, la diferencia de rendimiento prevista es comparativamente baja entre los alumnos de entornos socioeconómicos altos y bajos.

A través de los estudios de contexto de PISA se obtiene información adicional de potencial interés y utilidad a la hora de facilitar la comprensión de los resultados de la evaluación en lectura, matemáticas y ciencias. Las características contextuales de los hogares de los alumnos pueden ser importantes; entre ellas se incluye la profesión de los padres, su nivel de estudios, el capital cultural de la familia, el estatus de inmigrante de los alumnos y la lengua hablada en casa. Se obtienen muchos factores relativos al entorno escolar que pueden ser importantes. El entorno escolar incluye características tales como la naturaleza de las expectativas de los profesores respecto a sus alumnos, la disposición de estos últimos para esforzarse en su educación, la existencia de un clima de disciplina y de unas relaciones alumno-profesor positivas. Las percepciones que los alumnos tienen del apoyo del profesorado influirán en sus actitudes hacia la escuela, como también influirá el clima de disciplina, incluidos los aspectos relativos al absentismo escolar, la frecuencia de conductas negativas por parte de los alumnos, los casos de intimidación y acoso escolar, el uso de drogas, la incidencia de los alumnos que faltan a clase, la existencia de ruidos y desorden, y el tiempo que normalmente tardan las clases en comenzar a trabajar. Asimismo, puede resultar interesante el nivel moral de alumnos y profesores. Los datos indican que los directores de los centros perciben un nivel moral y de compromiso bajo tanto por parte de los alumnos como por parte de los profesores. Entre el resto de factores escolares que recogen los datos de PISA se incluye el grado de participación de los padres, las prácticas de seguimiento del progreso de los alumnos y las consecuencias de las prácticas de evaluación, las prácticas para supervisar el rendimiento de los centros, las relativas a la formación del personal, a la gestión de los centros y a la supervisión del profesorado.

Pero, ¿qué historia subyace bajo estos factores? Se trata de una historia desigual y compleja, y no existe una relación sencilla entre estos factores y el rendimiento de los alumnos en las pruebas PISA. Está claro que el contexto cultural e histórico tiene una importancia vital y es probable que los métodos y las lecciones varíen enormemente de un país a otro. No cabe una solución simplista del tipo «adoptemos el modelo finés». Sin embargo, las comparaciones pueden facilitar el proceso de revisión y evaluación de la práctica actual. El debate de las implicaciones para la política en relación con el impacto de los factores contextuales a escala de centros que se incluye en el informe internacional (OCDE, 2004, pp. 191-203) constituye un resumen bastante útil sobre los diferentes enfoques de la política que podrían plantearse.

En el caso de España, podría afirmarse que el gradiente social relativamente horizontal que se observa implica que el objetivo en el que hay que centrarse debe ser el bajo rendimiento. Igualmente, en España, las diferencias relativamente pequeñas entre los centros pueden dar a entender que el objetivo hacia el que hay que dirigirse debe ser los alumnos con bajo rendimiento de todos los centros educativos. Y reconociendo que la mayoría de las diferencias se dan entre los alumnos de los centros, puede ser beneficiosa la búsqueda de estrategias dirigidas a reducir la segregación dentro de los centros educativos. Asimismo, también puede resultar útil un enfoque dirigido a ofrecer una mayor ayuda directa a los alumnos con bajo rendimiento.

ALGUNAS CONCLUSIONES RELATIVAS A LA COMPETENCIA MATEMÁTICA EN PISA

Los resultados de España en la evaluación PISA 2003 con relación a la competencia matemática muestran un nivel de rendimiento medio un tanto por debajo de la media de la OCDE. Esto es así para cada una de las áreas de matemáticas, lectura, ciencias y resolución de problemas.

Cataluña, Castilla y León, y el País Vasco disponen de datos representativos para sus territorios y los resultados generales varían de forma bastante notoria en función de la región. El rendimiento en los casos de Castilla y León y el País Vasco fue muy superior al de la media de España en todas las escalas matemáticas, y superior al de la media de la OCDE en algunas. Cataluña también se situó por encima de la media de España y, en general, se aproximó a la media de la OCDE. Respecto a las ciencias, Cataluña mostró un rendimiento relativamente alto y el País Vasco, relativamente bajo.

Las diferencias de género en matemáticas son menores en el caso de España que en el conjunto de la OCDE y de nuevo estas diferencias no son homogéneas en todas las regiones españolas: muy pequeñas en el caso del País Vasco y muy grandes en Cataluña.

Al comparar los resultados de PISA 2000 y PISA 2003 se han observado algunos cambios ocurridos a lo largo del tiempo. Las puntuaciones medias en lectura han descendido ligeramente, en especial entre los estudiantes de más bajo rendimiento. Asimismo, las puntuaciones en ciencias son peores entre los alum-

nos de niveles más bajos. Por otro lado, las puntuaciones de matemáticas en el área de *cambio y relaciones* han mejorado.

Los resultados de España relativos a la competencia matemática global y para cada una de las cuatro subescalas en particular se resumen en la tabla III.

TABLA III

Resumen de los resultados en matemáticas de España en PISA 2003

Global de Matemáticas	
Alumnos en el nivel más alto (nivel 6)	Total alumnos: España 1,4%; OCDE 4,0% Varones: España 1,9%; OCDE 5,1% Mujeres: España 1,0%; OCDE 2,9%
Alumnos por debajo del nivel 2	Total alumnos: España 23,0%; OCDE 21,4% Varones: España 22,5%; OCDE 20,7% Mujeres: España 23,5%; OCDE 22,2%
Diferencia de género	España: 9 puntos en la escala PISA a favor de los varones OCDE: 11 puntos en la escala PISA a favor de los varones
Orden de clasificación	22-24 (OCDE); 25-28 (total países) 20 en un puesto superior; 7 en el mismo puesto; 12 en un puesto inferior
Espacio y forma	
Alumnos en el nivel más alto (nivel 6)	Total alumnos: España 1,6%; OCDE 5,8% Varones: España 2,2%; OCDE 7,3% Mujeres: España 1,1%; OCDE 4,3%
Alumnos por debajo del nivel 2	Total alumnos: España 26,8%; OCDE 24,8% Varones: España 24,5%; OCDE 23,0% Mujeres: España 29,1%; OCDE 26,5%
Diferencia de género	España: 12 puntos en la escala PISA a favor de los varones OCDE: 9 puntos en la escala PISA a favor de los varones
Orden de clasificación	21-25 (OCDE); 25-29 (total países) 22 en un puesto superior; 7 en el mismo puesto; 10 en un puesto inferior

(continúa)

(continuación)

Cambio y relaciones	
Alumnos en el nivel más alto (nivel 6)	Total alumnos: España 2,0%; OCDE 5,3% Varones: España 2,5%; OCDE 6,6% Mujeres: España 1,5%; OCDE 4,1%
Alumnos por debajo del nivel 2	Total alumnos: España 26,2%; OCDE 23,2% Varones: España 25,7%; OCDE 22,6% Mujeres: España 26,7%; OCDE 24,0 %
Diferencia de género	España: 13 puntos en la escala PISA a favor de los varones OCDE: 6 puntos en la escala PISA a favor de los varones
Orden de clasificación	21-24 (OCDE); 25-29 (total países) 21 en un puesto superior; 8 en el mismo puesto; 10 en un puesto inferior
Cantidad	
Alumnos en el nivel más alto (nivel 6)	Total alumnos: España 2,6%; OCDE 4,0% Varones: España 3,2%; OCDE 4,7% Mujeres: España 2,1%; OCDE 3,3%
Alumnos por debajo del nivel 2	Total alumnos: España 22,1%; OCDE 21,3% Varones: España 22,3%; OCDE 21,1% Mujeres: España 22,0%; OCDE 21,6%
Diferencia de género	España: 5 puntos en la escala PISA a favor de los varones (no significativo) OCDE: 6 puntos en la escala PISA a favor de los varones
Orden de clasificación	20-23 (OCDE); 23-26 (total países) 21 en un puesto superior; 5 en el mismo puesto; 13 en un puesto inferior

(continúa)

(continuación)

Incertidumbre	
Alumnos en el nivel más alto (nivel 6)	Total alumnos: España 1,5%; OCDE 4,2% Varones: España 1,8%; OCDE 5,3% Mujeres: España 1,2%; OCDE 3,1%
Alumnos por debajo del nivel 2	Total alumnos: España 20,8%; OCDE 20,7% Varones: España 20,7%; OCDE 19,7% Mujeres: España 21,0%; OCDE 21,7%
Diferencia de género	España: 8 puntos en la escala PISA a favor de los varones OCDE: 13 puntos en la escala PISA a favor de los varones
Orden de clasificación	18-23 (OCDE); 21-26 (total países) 18 en un puesto superior; 7 en el mismo puesto; 14 en un puesto inferior

Lógicamente, estos resultados conjuntos proporcionan sólo una impresión general del actual estado de un país en relación con el resto y, posiblemente, algunas pistas muy generales sobre las prioridades de la política educativa. Lo que tiene un mayor interés potencial para los profesores y los responsables de la elaboración del currículo es la información más específica sobre cómo les fue a los alumnos españoles ante los ítems concretos del conjunto de ítems de matemáticas de PISA 2003. Aunque los resultados de España muestran un nivel de rendimiento inferior, en promedio, al del conjunto de la OCDE, el rendimiento de los alumnos españoles, en promedio, fue mejor que el promedio de la OCDE para unos 20 de los 87 ítems de la prueba de matemáticas. Resulta interesante señalar que en 11 de esos ítems, el rendimiento de las chicas fue mejor en comparación al de los chicos (como promedio); que en 6 de esos ítems el porcentaje de respuestas correctas fue similar para chicos y chicas; y que para tres de ellos los chicos superaron a las chicas. En el otro extremo, en cinco de los seis peores ítems para España las chicas respondieron mucho peor que los chicos.

El rendimiento medio de España en sus dos «mejores» ítems sólo fue superado por cuatro y cinco países respectivamente. Resulta interesante señalar que en ambos ítems los alumnos tenían que interpretar cuestiones relacionadas con la probabilidad. De los ítems liberados, ocho se incluyen entre aquellos en los que el rendimiento de los alumnos españoles fue mejor que la media de la OCDE, y dos entre los 6 en los que el rendimiento de España fue peor.

Todos ellos se enumeran en la tabla IV que se muestra a continuación. Los ítems liberados pueden consultarse en inglés y en francés a través del sitio web de la OCDE³.

TABLA IV

Ítem de matemáticas publicados en los que el rendimiento de los alumnos españoles fue alto o bajo

Ítems de matemáticas publicados en los que el rendimiento de los alumnos españoles fue relativamente alto:

- Caminar (M124Q01 y Q03)
- Cubos (M145Q01)
- Crecer (M150Q02)
- Exportaciones (M438Q01)
- Basura (M505Q01)
- Elecciones (M510Q01)
- Escalera (M547Q01)

Ítems de matemáticas publicados en los que el rendimiento de los alumnos españoles fue relativamente bajo:

- Carpintero (M266Q01)
- Exámenes de ciencias (M6468Q01)

Es probable que quienes mejor interpreten esta información y saquen las oportunas conclusiones sean los profesores y otros expertos en materia curricular que forman parte del sistema educativo español. Dichas personas pueden juzgar si estos resultados satisfacen las expectativas y son consistentes consecuentes con las prioridades de los currículos locales aplicados en los centros educativos españoles. Sin duda, dichos expertos podrán hacerse una idea sobre si el actual énfasis puesto en la resolución de problemas auténticos y en las aplicaciones al mundo real que está presente en el currículo español es el adecuado, o si los resultados de PISA tienen algo útil que decir acerca de esas prioridades. Dichas personas también serán capaces de determinar de qué manera los resultados de PISA probablemente influyan en los avances curriculares actuales de los centros educativos españoles.

³ En español se pueden consultar las publicaciones INECSE (2005): *PISA 2003. Pruebas de Matemáticas y de Solución de Problemas*, Madrid, MEC; y INECSE (2005): *Programa PISA. Pruebas de Comprensión Lectora*, Madrid, MEC. (Nota de la Redacción).

REFERENCIAS BIBLIOGRÁFICAS

- NISS, M. (1999): *Kompetencer og Uddannelsesbeskrivelse, Uddannelse, 9*, pp. 21-29, Dinamarca: , Ministerio de Educación.
- OCDE (2001): *Knowledge and Skills for Life – -First Results from PISA 2000*. París, Francia: OCDE.
- (2002): *Reading for Change – -Performance and Engagement across Countries*. París.
- (2003): *The PISA 2003 Assessment Framework – -Mathematics, Reading, Science and Problem Solving Knowledge and Skills*. París.
- (2004): *Learning for Tomorrow's World- – First Results from PISA 2003*. París, – (2005): *PISA 2003 Technical Report*. París.
- STEEN, L. A. (1990):*On the Shoulders of Giants: New Approaches to Numeracy*. National Academy Press, Washington, Distrito de Columbia.