

Fundamentos y cuestiones políticas subyacentes al desarrollo de PISA

Andreas Schleicher

Jefe de la División de Indicadores y Análisis del Directorio para la Educación de la OCDE

Resumen:

Tras una breve introducción sobre los orígenes de PISA en el marco de la OCDE y la voluntad política que la anima, el autor analiza en primer lugar los cuatro objetivos de PISA. Mediante los tests de rendimiento, el cuestionario dirigido a los padres de los alumnos, y el cuestionario de contexto socioeconómico, el análisis de PISA se ha centrado en la calidad de los resultados del aprendizaje, así como en la influencia del contexto socioeconómico de alumnos y centros y en la equidad en las oportunidades educativas. Los dos últimos objetivos, a saber, la eficacia y eficiencia de los procesos educativos, y el impacto de los resultados del aprendizaje en el bienestar social y económico, serán estudiados con mayor profundidad en el futuro.

A continuación, el autor examina aquellos aspectos del diseño técnico de PISA que han hecho posible satisfacer con éxito y alta fiabilidad sus objetivos: a) la población objeto de estudio: seleccionada según la edad, y con una tasa de exclusiones muy baja.; b) el marco de evaluación, inspirado por el innovador concepto de alfabetización introducido por PISA; c) los instrumentos de evaluación; y d) el desarrollo de escalas absolutas de rendimiento que permitan interpretar las medidas resultantes en términos políticos.

Los resultados de PISA han provocado un intenso debate, tanto en el ámbito nacional como internacional, en los países participantes. En consecuencia, muchos de ellos han emprendido análisis (Alemania) y reformas (Dinamarca) de sus políticas educativas.

Palabras clave: evaluación internacional, PISA, política educativa, características del diseño, impacto de resultados, alfabetización, administración.

Abstract: *Rationale and policy issues underlying the development of PISA*

After a brief introduction on the origins of PISA as a benchmark study of the OECD and the political will which favours it, the author analyses, in the first place, the four targets of the PISA project.

Through a set of performance tests, a questionnaire addressed to students' parents and a socioeconomic context questionnaire, the PISA approach focuses on the quality of learning outcomes as well as on equity in the distribution of learning opportunities. The last two objectives, i.e., effectiveness and efficiency of educational practices and the impact of learning outcomes on the economic and social welfare, will be dealt with in more detail in the future.

Next, the author inquires into those aspects derived from PISA technical approach which have made possible the successful and reliable attainment of its targets: First, the population to be studied, selected according to age and with a very low rate of exclusion. Second, the assessment framework inspired by the innovative concept of literacy introduced by PISA. It consists of specific areas defined on the basis of the content and the practices that students need to know as well as of the appropriate situations to apply them. Likewise, other competencies will also be evaluated in the future. Third, the assessment tools (items with different formats) developed by an international network of evaluation agencies and subject to inspections with the aim of getting rid of any cultural bias. Finally, the development of performance tools which allow the interpretation of the resulting measures in political terms.

PISA outcomes have given rise to an intense debate, both in national and international contexts, in the participant countries. As a consequence, many of them have undertaken the analysis (Germany) and subsequent reforms (Denmark) of their educational policies.

Key words: international assessment, PISA, educational policy, design characteristics, impact of the results, literacy, management.

LOS ORÍGENES DE PISA

Los 30 países miembros de la Organización para la Cooperación y el Desarrollo Económico (OCDE) colaboran en el desarrollo de líneas de actuación en prácticamente todas las áreas de la política pública excepto defensa (www.oecd.org). La labor en educación ha constituido parte del programa de la OCDE desde su fundación en 1961, y una parte de esa labor ha consistido en la recopilación y publicación de estadísticas relativas a los sistemas educativos nacionales. Sin embargo, a mediados de los años ochenta del siglo XX se constató con preocupación que los datos eran incompletos, y que los parámetros empleados para fuentes y análisis en la producción de datos no reunían calidad suficiente. También existía la impresión de que la educación estaba menos adelantada que otras áreas de trabajo de la OCDE en la construcción de una base de datos cuantitativa y relevante para su tarea en el ámbito político.

La respuesta fue crear el programa de Indicadores de Sistemas Educativos (INES), que dio comienzo en 1988 a través del Centro de la OCDE para la Investigación e Innovación Educativas (CERI, en sus siglas en inglés). Los países querían que los INES se desarrollaran basándose en estimaciones documentadas que determinasen cuáles de las estadísticas tenían importancia política y cómo éstas podían ser resumidas y presentadas en forma de indicadores que fueran útiles en el debate sobre qué política seguir. Griffith anotó:

El programa comenzó sin directrices sobre qué datos exactamente había que desarrollar. El mandato era mucho más general. La OCDE debía organizar a representantes de los países miembros para que trabajaran juntos en el desarrollo de un sistema de *indicadores educativos internacionales*, que sirviera a las necesidades de los responsables políticos interesados en comprender qué estaba ocurriendo en educación

fuera de sus propios países. (...) Pero hubo un tema que surgió ya en la etapa más temprana del INES: la atención estaba centrada sobre todo en la relevancia política de los indicadores y de la nueva información estadística. (Griffith, 2000, p. 2)

La labor se llevó a cabo a través de varias «redes», cada una estaba dirigida por un estado miembro, y en ellas se generaban innovaciones considerables. Primero se trabajó en el desarrollo de indicadores nacionales básicos de los recursos humanos y financieros destinados a educación, del acceso a la educación y de las pautas de promoción y, gradualmente, se emprendió asimismo la medición del rendimiento y de los resultados de la educación.

El resultado más significativo de este proceso ha sido el desarrollo de un ciclo de estudios del rendimiento estudiantil, dirigidos por la OCDE y realizados por el Programa para la Evaluación Internacional de Alumnos (PISA). Cuando el Consejo de la OCDE instituyó formalmente a PISA como programa en 1997, menos de la mitad de los entonces 29 países miembros se había decidido a tomar parte en él. Cuando se reunieron los primeros datos en PISA 2000, participaron 28 miembros, junto con otros cuatro en aquel momento y otros 11 que utilizaron los mismos tests en 2001. En PISA 2003 participaron todos los países miembros de entonces (30) y otros 11 más. Para PISA 2006 se han confirmado casi 60 participantes.

LOS OBJETIVOS DE PISA EN POLÍTICA EDUCATIVA

El propósito de PISA va mucho más allá de la simple supervisión del estado actual del aprendizaje por parte de los alumnos en los sistemas educativos nacionales. La información facilitada por PISA debía permitir a los responsables políticos observar qué factores están asociados al éxito educativo, y no limitarse a establecer comparaciones entre resultados aisladamente.

Los gobiernos estructuraron los resultados que se deseaban obtener de PISA desde una perspectiva política, en cuatro campos temáticos relativos a:

- la calidad de los resultados del aprendizaje;
- igualdad en los resultados del aprendizaje y equidad en las oportunidades educativas;
- la eficacia y eficiencia de los procesos educativos; y
- el impacto de los resultados del aprendizaje en el bienestar social y económico.

A continuación se detallan dichos resultados para cada uno de estos temas.

COMPARANDO LA CALIDAD DE LOS RESULTADOS DEL APRENDIZAJE

PISA se estableció principalmente para medir hasta qué punto los adultos jóvenes, que a los 15 años están a punto de terminar la educación obligatoria, se hallan bien preparados para afrontar los retos de las actuales sociedades del conocimiento. El estudio estaba orientado hacia adelante: se centraba en la capacidad de la gente joven para usar su conocimiento y sus destrezas para afrontar los retos de la vida real, más que en qué medida dominan un currículo

escolar específico. Esta orientación refleja un cambio en los objetivos y finalidades de los propios currículos, que están cada vez más centrados en lo que los alumnos saben hacer con lo que aprenden en los centros educativos, y no simplemente en si saben reproducir lo que han aprendido.

El interés de los países por las implicaciones políticas de PISA significa que la atención dedicada a medir la calidad de los resultados del aprendizaje va mucho más allá de una suerte de «carreras de caballos», en las que los países simplemente se clasifican por puestos según la puntuación media de los alumnos. En consecuencia, PISA no sólo revisa la calidad de los resultados del aprendizaje, sino que examina asimismo la distribución de dichos resultados para facilitar la comprensión de los atributos individuales, institucionales y del sistema que están asociados a diferencias de rendimiento entre individuos, instituciones y países.

Como primer paso, PISA presta considerable atención a la distribución de logros usando, en particular, escalas de rendimiento con las cuales se pueda interpretar de manera sustancial la distribución del rendimiento de los alumnos. En algunos países, la información que se desprende de las distribuciones es bastante diferente de la aportada por la media. España, por ejemplo, ocupa un puesto por debajo de la media de la OCDE en Matemáticas debido sobre todo a que posee una proporción mucho menor de alumnos que manejen Matemáticas del más alto nivel (1,4% comparado con una media de 4% en la OCDE), no porque posea una proporción de malos estudiantes mayor que la media (tabla 2.5a, OECD, 2004b).

En segundo lugar, PISA también busca distinguir, dentro de la variación en el rendimiento de los alumnos, entre dos componentes: variación dentro de un mismo centro, y variación entre centros. Tal comparación puede ayudar a valorar hasta qué punto la calidad de los resultados del aprendizaje varía entre los centros, y es por tanto un resultado constante y predecible del sistema educativo. Las comparaciones en estas áreas muestran de hecho que hay grandes diferencias entre los países en lo que respecta a cómo varía el rendimiento del alumno según los centros, con una variación en el rendimiento entre centros en algunos de los países mejor clasificados, que asciende a menos del 10% de la variación en el rendimiento conjunto de los alumnos (de manera que los padres en estos países pueden contar con estándares altos y constantes de rendimiento en todo el sistema educativo), mientras que, en otros países, más de la mitad de la variación media de la OCDE en el rendimiento se origina en los ámbitos del centro y/o del programa, a menudo combinada con un rendimiento conjunto moderado (tabla 4.1, OECD, 2004b). Por ejemplo, Canadá o Finlandia están entre los diez países de la OCDE que tienen la variación más pequeña respecto al rendimiento en Matemáticas entre centros, al tiempo que se encuentran entre los diez países de la OCDE con los mejores resultados de rendimiento conjunto, mientras que en Polonia y España sucede lo contrario.

A medida que discurra el ciclo de evaluación, PISA dedicará cada vez más atención a examinar cómo están cambiando las cosas. La evaluación de 2009 proporcionará el primer análisis completo de tendencias en competencia lectora durante el período 2000-2009, y permitirá comparar las mejoras en la calidad de los resultados educativos de cada país en las áreas evaluadas por PISA con las de otros

países. El período de nueve años se juzga apropiado para analizar el impacto de las políticas introducidas a raíz de PISA 2000.

Además de examinar las tendencias en los resultados del aprendizaje a lo largo del tiempo, PISA está trasladando parte del énfasis puesto en tomar como parámetro el rendimiento estudiantil por países, a la comparación del progreso en el aprendizaje dentro de los sistemas educativos, realizada mediante la evaluación de los conocimientos y destrezas de los alumnos en distintas etapas del sistema educativo. Dicho examen se llevará a cabo evaluando, de forma coordinada y en diferentes niveles de educación, los resultados en el aprendizaje y los factores contextuales asociados a ellos. Estos análisis no se limitarán a examinar cómo crece el rendimiento de los alumnos, sino que la evaluación del rendimiento a diversas edades ayudará a comprender cómo evolucionan la motivación y las características de los alumnos en diferentes países a medida que los estudiantes avanzan en el sistema educativo. En potencia, los análisis también podrían ayudar a comprender qué factores de la instrucción y del sistema están asociados a dicho aumento del rendimiento y a los cambios en la actitud de los alumnos hacia el aprendizaje.

Finalmente, PISA ha puesto en marcha un cuestionario para padres que aportará una opinión externa sobre cómo es, desde su punto de vista, la calidad de la enseñanza. Mediante este instrumento, los padres de los alumnos evaluados proporcionarán información sobre sus expectativas y aspiraciones respecto a los resultados que la educación de sus hijos conseguirá en el mercado laboral, su estrategia global respecto a la educación de sus hijos, cuánto planean invertir actualmente en el futuro educativo de sus hijos, la influencia de los progenitores en puntos críticos de la educación de un alumno (por ejemplo, la elección de materias), las actitudes de los padres hacia las diferentes materias y la eficacia que atribuyen a éstas para obtener resultados en el mercado laboral, y la implicación de los padres en el centro de enseñanza.

COMPARANDO LA IGUALDAD DE LOS RESULTADOS DEL APRENDIZAJE Y LA EQUIDAD DE LAS OPORTUNIDADES PARA APRENDER

PISA 2000 y PISA 2003 ya han dedicado una considerable atención a cuestiones de equidad, concretamente a cómo influye el contexto socioeconómico de alumnos y centros en los resultados del aprendizaje, y qué palancas políticas están asociadas con estos resultados. Dado que la diversidad socioeconómica en los países de la OCDE está en aumento (en España, por ejemplo, el porcentaje de alumnos de origen inmigrante se ha multiplicado por diez a lo largo de la última década), y dada la creciente desventaja en que se encuentran en la sociedad y en el mercado laboral los individuos con escasas competencias (OECD, 2005), el tema de la equidad continuará estando entre las prioridades de PISA en los próximos años.

La información sobre el contexto social de alumnos y centros se obtiene a través de un cuestionario que rellenan los alumnos junto con los tests de rendimiento. Usando la información proporcionada por los alumnos, se elabora un índice PISA del estatus económico, social y cultural a partir del más alto de los índices socioeconómicos del estatus ocupacional de la madre o del padre, del nivel educativo más alto de la madre o del padre, y de información sobre libros y otros recursos educativos y

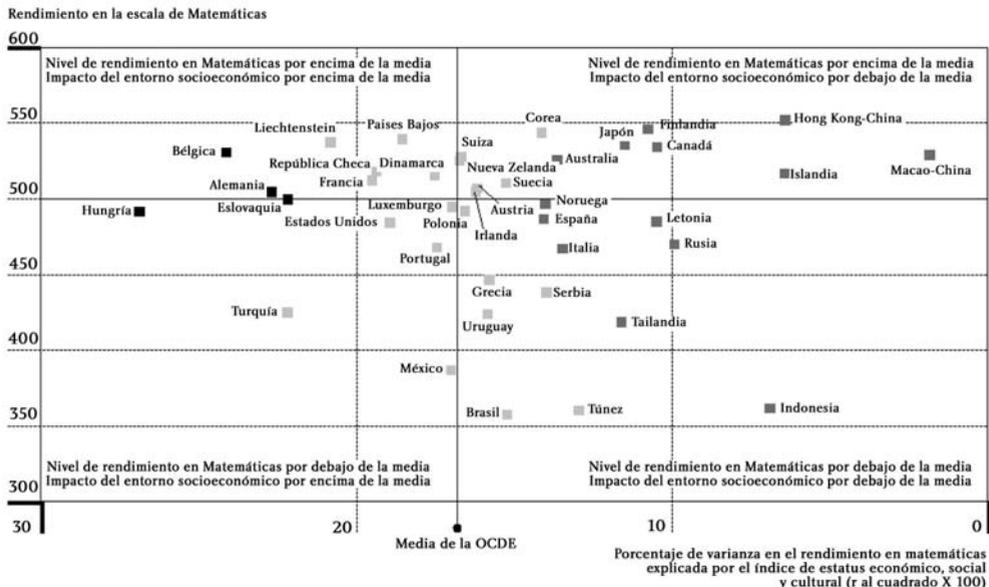
culturales presentes en el hogar. Al aplicar este índice al ámbito del centro, se obtiene un indicador de la composición social de los centros (OECD, 2004b, p.307).

Relacionar dichos indicadores socioeconómicos con el rendimiento logrado ofrece una vía para comprender hasta qué punto los sistemas educativos proporcionan oportunidades equitativas de aprendizaje. En todos los países existe una correlación positiva y una pendiente de regresión positiva que indican en qué medida, por término medio, lo ventajoso que sea el entorno social de los alumnos se asocia con el incremento en el nivel medio de rendimiento de las evaluaciones PISA. Sin embargo, ha resultado llamativo que la intensidad de esta relación varíe mucho en función del país del que se trate. En algunos países, incluido España, las diferencias de contexto social están relacionadas con diferencias de rendimiento educativo mucho menores que en otros Estados, lo cual sugiere que las cosas no han de ser forzosamente como son en otros países donde existe una relación muy acusada.

FIGURA I

Rendimiento en Matemáticas e impacto del entorno socioeconómico. Rendimiento medio de los países en la escala de Matemáticas PISA y relación entre el rendimiento y el índice de estatus económico, social y cultural

- Fuerza de la relación entre el rendimiento y el entorno socioeconómico por encima del impacto medio de la OCDE.
- Fuerza de la relación entre el rendimiento y el entorno socioeconómico que no se diferencia del impacto medio de la OCDE en un grado estadísticamente significativo.
- Fuerza de la relación entre el rendimiento y el entorno socioeconómico por debajo del impacto medio de la OCDE.



Nota: La media de la OCDE utilizada en esta figura es la media aritmética de todos los países de la OCDE.

Fuente: Base de datos OCDE PISA 2003. Tabla 4. 3a.

Gráfico tomado del *Informe PISA 2003. Aprender para el mundo del mañana* (corresponde a la figura 4.10 en OCDE, 2004b).

La figura I resume estos resultados. El eje horizontal representa aquella variación del rendimiento atribuible al contexto social de los alumnos. Los países donde esta relación es más acusada que para el conjunto de la OCDE, aparecen a la izquierda. Los países en los que la relación es menos marcada están a la derecha. Por lo tanto, se puede considerar al eje horizontal como la representación de la «equidad social» del sistema educativo. El eje vertical representa el rendimiento medio en Matemáticas por países. La figura está dividida en cuatro cuadrantes por una línea horizontal a la altura de 500, la calificación media en Matemáticas en el conjunto de la OCDE, y por una línea vertical que pasa por 16,8, el grado de relación en la OCDE entre el contexto social y el rendimiento de los alumnos. El cuadrante superior derecho se puede describir como «alta calidad, alta equidad» en relación con la OCDE en su conjunto, y los otros cuadrantes recibirían su denominación de la misma manera, como se ve en la figura I. El rasgo más importante de esta figura es que muestra que es posible conseguir al mismo tiempo una alta calidad y una alta equidad. No es necesario sacrificar a una por la otra, aunque está claro que hay países, como Bélgica, donde se obtiene una alta calidad como resultado medio pero donde existe también una conexión relativamente estrecha entre el contexto social de los alumnos y el nivel de sus competencias.

PISA ha llevado más allá el análisis de problemas relacionados con la equidad, separando aquéllos relativos a la heterogeneidad socioeconómica dentro de cada centro de aquéllos relativos a la segregación socioeconómica presente en el propio sistema escolar. Esto ha permitido examinar hasta qué punto los sistemas educativos moderan o refuerzan los factores del contexto socioeconómico. En países como Japón, los Países Bajos, la República Checa, Bélgica, Austria y Alemania, por ejemplo, se agrupa a los alumnos de alto rendimiento en centros de alto rendimiento, y éstos muestran una marcada tendencia a provenir de contextos sociales favorecidos. En cambio, los alumnos de bajo rendimiento se agrupan en centros de bajo rendimiento y tienden a proceder de entornos sociales desfavorecidos. La consecuencia es que hay muy poca relación entre el contexto social de los alumnos y su rendimiento en Matemáticas en un mismo centro, dada la restricción impuesta a ambas variables. En Finlandia y Canadá, por el contrario, el efecto del agrupamiento por centros es pequeño. En dichos países, la relación entre contexto social y rendimiento en Matemáticas es similar en los tres ámbitos: global, entre centros y dentro del centro. España también pertenece a este último grupo de países.

En algunos de los países en los que una proporción considerable de la varianza total corresponde a los centros, dicho fenómeno es una consecuencia de la política educativa. En Alemania, por ejemplo, los alumnos son distribuidos por centros de diferentes tipos a partir de los 10 años de edad, basándose en su rendimiento hasta ese momento y en una valoración sobre la conveniencia de un centro más académico o uno de formación profesional para la siguiente etapa. La intención deliberada de esta política es reducir la variación dentro de los centros reuniendo a alumnos relativamente similares, y aumentar la variación entre

centros, lo cual se verá reflejado en las diferencias entre los currículos de los centros. Por supuesto, en ningún país se distribuye deliberadamente a los alumnos por centros según su extracción social, pero la consecuencia de esta clasificación es que de hecho se distribuye también en función del contexto social. En muchos países la consecuencia es que los alumnos procedentes de un contexto social privilegiado son dirigidos a los centros académicos más prestigiosos, que proporcionan resultados educativos superiores (como indica su mayor rendimiento en las mediciones PISA), y los alumnos procedentes de un contexto social desfavorecido son dirigidos a los centros de formación profesional, menos prestigiosos, que proporcionan resultados educativos más pobres (como indica su menor rendimiento en las mediciones PISA). Por lo tanto, la organización de los centros refleja y reproduce al mismo tiempo las divisiones sociales existentes.

La relación que se planea establecer entre la evaluación PISA de los alumnos de 15 años y el rendimiento en un nivel primario de educación permitirá ampliar el potencial analítico de PISA en este contexto, de modo que PISA examinará también cómo evoluciona la distribución de resultados del aprendizaje en los diferentes niveles educativos; hasta qué punto tienen éxito los sistemas educativos a la hora de moderar el impacto que tiene en el rendimiento el capital económico, social y cultural de los alumnos y sus familias a medida que los alumnos se hacen mayores; y qué relación guardan los recorridos educativos y las estructuras institucionales de diferentes países con dichos resultados.

EFICACIA Y EFICIENCIA DE LOS SISTEMAS EDUCATIVOS

Una de las principales razones por las que los responsables políticos están interesados en las comparaciones internacionales del rendimiento de los alumnos es que nos ayudan a comprender las diferencias en el rendimiento de los centros y sistemas educativos, las cuales reflejan la eficacia y eficiencia de los procesos educativos. No obstante, a pesar de mostrar las importantes diferencias de rendimiento existentes entre sistemas educativos, las evaluaciones internacionales siguen teniendo grandes limitaciones a la hora de determinar cuáles son las actuaciones políticas asociadas a dichas diferencias.

Aunque los análisis en este campo se encuentran aún en su etapa inicial, PISA ha conseguido identificar una serie de características de los centros que pueden tener relevancia en los resultados del aprendizaje y en las diferencias de resultados entre centros. Tomando en conjunto las características de los alumnos, el contexto socioeconómico de alumnos y centros, la percepción del ambiente en el centro que tienen los alumnos y directores del mismo, los informes de los directores sobre las políticas y prácticas del centro y, finalmente, la evaluación de la disponibilidad y calidad de los recursos educativos, descrita en el informe internacional de PISA sobre resultados iniciales (OECD, 2004b), se explica el 54% de la variación en el rendimiento medio de los países de la OCDE, una media del 71% de la variación de rendimiento entre centros de un mismo país, y una media del 8% de la variación del rendimiento de los alumnos en un centro dado.

Aunque no siempre es posible medir con precisión el impacto que cada uno de estos factores tiene en el rendimiento de los alumnos, muchas de las diferencias que se han descubierto en un centro o entre centros ponen de manifiesto temas críticos para los responsables políticos. Por ejemplo, tanto directores como alumnos tienen opiniones muy dispares sobre la calidad del ambiente educativo de sus centros, ya los consideremos por países o por centros dentro de cada país. Se ha comprobado durante la investigación que estas diferencias afectan a la efectividad de la enseñanza y el aprendizaje, como también ocurre con las diferencias en la manera de administrar los centros. Por tanto, estos resultados muestran ante todo la necesidad de asegurar que todos los centros posean un ambiente que favorezca el aprendizaje, así como una cultura administrativa y unos recursos que sean compatibles con una enseñanza y aprendizaje efectivos.

Además de describir estos fenómenos de los centros educativos, PISA ha buscado la manera de medir el efecto que tienen en el rendimiento de alumnos y centros. Este efecto es de tres tipos.

El primero es un efecto independiente de otros factores del centro y del contexto socioeconómico. Por ejemplo, en el caso de la disciplina, ¿hasta qué punto pueden esperar los alumnos de un centro bien disciplinado obtener mejores resultados que en otro con poca disciplina, si la composición social de ambos centros es la misma, y si ambos son semejantes con respecto a política del centro, procesos y recursos? En estos términos, los resultados de PISA indican que el que se trate de un centro u otro supone una gran diferencia. Alumnos y centros tienden a rendir más en un ambiente caracterizado por la disciplina y por altos niveles de moral y compromiso en los alumnos. Por el contrario, los centros donde la relación alumno-profesor es pobre, tienden a rendir considerablemente menos. Por tanto, los centros pueden beneficiarse si ponen énfasis no sólo en las técnicas didácticas, sino también en la forma en que los profesores se relacionan con los alumnos. Además, los centros que ofrecen actividades relacionadas con Matemáticas también tienden a rendir más, como también ocurre en aquéllos que tienden a evitar la agrupación de alumnos según su capacidad, y en aquéllos donde los directores evalúan positivamente la adecuación de los profesores suplentes y de los recursos educativos.

Un segundo tipo de efecto tiene lugar cuando la composición socioeconómica de los centros contribuye a que aquéllos con determinadas características consigan mejores resultados. Esta faceta del efecto no tendría relevancia directa desde la perspectiva de los responsables políticos, ya que no sería aconsejable tratar de incrementar un factor debido a que parece estar asociado con el rendimiento, cuando de hecho es el contexto social el que tiene la influencia decisiva. No obstante, los padres que estén buscando qué centro elegir harían bien en prestar atención a tales factores, dado que a ellos les interesan los resultados globales del rendimiento en el centro, incluyendo cualquier efecto causado por la composición socioeconómica del mismo.

Un tercer tipo de efecto se da donde los factores socioeconómicos y del centro actúan juntos. Este análisis indica que los primeros refuerzan el impacto que

el ambiente tiene en el rendimiento del centro de manera importante, y esto debería ser relevante para responsables políticos que deseen asegurar que todos los centros cuentan con profesores comprometidos y un ambiente de orden, con independencia de su composición socioeconómica. Es posible que este efecto aparezca porque los alumnos de extracción privilegiada traen con ellos un nivel mayor de disciplina y una concepción más positiva de los valores escolares, o quizá porque las expectativas de los padres con respecto a una buena disciplina en clase y a un fuerte compromiso del profesorado son más altas en centros con una composición socioeconómica favorecida. A la inversa, centros desfavorecidos experimentarían una menor presión por parte de los padres a la hora de fomentar prácticas disciplinarias efectivas o asegurarse de reemplazar a profesores ausentes o desmotivados. De manera similar, el análisis ha puesto de manifiesto que una gran parte de la variación en el rendimiento entre centros es atribuible a políticas, prácticas y recursos del centro, junto con el aspecto socioeconómico. Los centros favorecidos en este último aspecto aplican probablemente políticas y prácticas más efectivas, quizá porque los mejores profesores eligen trabajar allí o porque tienen acceso a más y mejores recursos.

La intención de PISA es ahondar en estas cuestiones y hacer de la eficiencia y eficacia de los procesos educativos el tema «interdisciplinar» del estudio PISA 2012. Hay al menos dos razones por las que PISA 2012 es la ocasión más oportuna para abordar este tema. Futuras innovaciones harán posible relacionar los resultados observados con un *input* relevante y con variables de proceso en los distintos ámbitos de los alumnos, de los profesores y del centro, con un sondeo de la OCDE sobre el profesorado que se llevará a cabo de 2007 en adelante. Aparte de esto, PISA también quiere ampliar el panorama recopilando información, en el contexto del sistema, sobre ciertas palancas políticas clave que, en general, se considera que conducen al aumento de eficacia y eficiencia; y relacionando la información sobre políticas deseables con su puesta en práctica en el ámbito de los centros, junto con la opinión que les merece a padres, profesores y centros, y, por supuesto, con los resultados de aprendizaje que se hayan observado.

IMPACTO DE LOS RESULTADOS DEL APRENDIZAJE EN EL BIENESTAR SOCIAL Y ECONÓMICO

Se dispone de escasa documentación comparativa sobre el impacto que las competencias adquiridas durante la formación tienen en el bienestar económico y social del individuo y de la sociedad, o hasta qué punto las oportunidades de aprendizaje están conectadas de manera coherente entre las distintas instituciones educativas, o entre éstas y otras formas de aprendizaje.

Sin embargo, en el futuro, un vínculo más estrecho entre los estudios sobre los centros y el estudio planeado por la OCDE sobre las competencias de los adultos permitirá reexaminar las competencias, aspiraciones y compromiso de los jóvenes que están acabando su educación inicial y están adquiriendo responsabilidades en el trabajo y en la vida, junto con su impacto en el bienestar social y económico. Este último incluiría las consecuencias para el individuo, tales como

una integración satisfactoria en el mercado laboral, estatus y ganancias propias de un trabajador, participación en otros aprendizajes, en la vida cívica y en la educación a lo largo de todo el ciclo vital, además de resultados globales, como el fomento del crecimiento económico y de la participación social.

CARACTERÍSTICAS CLAVE DEL DISEÑO DE PISA

Afrontar la agenda de actuaciones perfilada más arriba implica elevadas exigencias en el diseño técnico de PISA. Para conducir el desarrollo de dicho diseño, han sido claves las siguientes características:

- su orientación política; los métodos de informes y diseño están determinados por la necesidad de los gobiernos de extraer lecciones políticas;
- el innovador concepto de «alfabetización», que se centra en la capacidad de los alumnos para aplicar sus conocimientos y destrezas en áreas temáticas clave, y para analizar, razonar y comunicarse de manera efectiva cuando plantean, resuelven e interpretan problemas en situaciones variadas;
- su relevancia para un aprendizaje que durará toda la vida, por lo cual PISA no se limita a evaluar las competencias curriculares y transversales de los alumnos, sino que también les pide informes sobre su propia motivación para aprender, lo que piensan sobre ellos mismos, y sus estrategias para el aprendizaje;
- su regularidad, lo cual permite a los países hacer seguimiento de su progreso a la hora de alcanzar objetivos clave de aprendizaje; y
- la amplitud de su cobertura geográfica y de su naturaleza colaboradora; los 48 países que han participado hasta ahora en un estudio PISA y los 11 países adicionales que se unirán a PISA 2006 representan un tercio de la población mundial y casi nueve décimas partes del producto interior bruto (PIB) del planeta.

Actualmente PISA es el programa internacional más integral y riguroso que existe para evaluar el rendimiento de los alumnos y para recopilar información sobre los factores estudiantiles, familiares e institucionales que puedan ayudar a explicar las variaciones en el rendimiento. Expertos de primer orden de los países participantes son los que deciden acerca del alcance y la naturaleza de las evaluaciones, y qué información de fondo debe reunirse; dichos expertos están dirigidos conjuntamente por sus gobiernos de acuerdo con intereses políticos comunes a todos (véase el Anexo). Se dedican esfuerzos y recursos considerables a lograr tanto equilibrio como amplitud cultural y lingüística en los materiales de la evaluación. Se aplican estrictos mecanismos que garantizan la calidad de la traducción, el muestreo y la recopilación de datos. La consecuencia es que los resultados de PISA son altamente válidos y fiables, y pueden mejorar significativamente la comprensión de los resultados educativos en los países más desarrollados del mundo, así como en un número creciente de países en vías de desarrollo económico.

Entre la visión de PISA y su puesta en práctica surgen muchos problemas técnicos a los que es necesario enfrentarse, incluyendo:

- la selección de la población objeto de estudio;
- el establecimiento de un marco de evaluación que defina lo que va a medirse y cómo;
- el desarrollo de los instrumentos de evaluación; y
- el desarrollo de escalas informativas que se puedan interpretar en términos políticos.

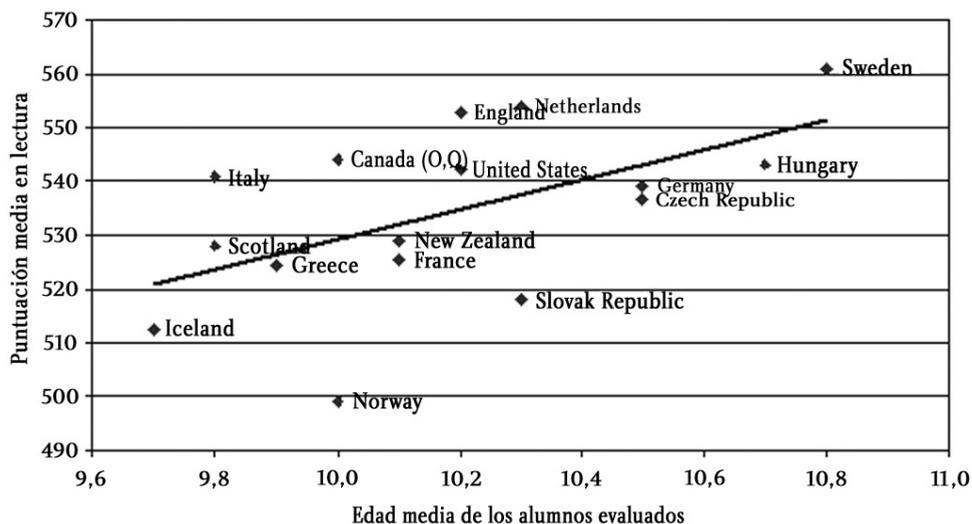
Estos temas se abordan a continuación con mayor detalle.

SELECCIÓN DE LA POBLACIÓN OBJETO DE ESTUDIO

Para asegurar que los resultados de diferentes países sean comparables, un estudio internacional necesita evaluar poblaciones que también lo sean. Estudios internacionales previos, como por ejemplo el TIMSS, definieron sus poblaciones en términos de grado escolar. Ello refleja la organización de los centros y permite un muestreo jerárquico según el centro y la clase (dentro de un mismo grado), y también según los estudiantes dentro de una misma clase, siempre que no todos los alumnos de una clase estén siendo estudiados. Se pueden analizar entonces tanto los efectos clase-profesor como los del centro. Sin embargo, las diferencias entre países en lo que respecta a la naturaleza y alcance de la educación y atención preescolar, la edad de entrada en la educación formal, y la estructura del sistema educativo, no permiten definir los grados escolares de manera que sean comparables internacionalmente. Por ejemplo, la figura II muestra la relación entre la puntuación media de competencia lectora por países y la edad media de los alumnos evaluados en el Estudio del Progreso en Competencia Lectora Internacional (PIRLS, en sus siglas en inglés), también basado en el grado, y organizado por la IEA para países de la OCDE participantes, excepto Turquía, que está excluida de este conjunto de datos. Más del 30% de la variación en la puntuación media de la competencia lectora es atribuible a las diferencias en la edad media de los estudiantes evaluados, lo cual reduce la posibilidad de establecer una relación entre las diferencias de logro nacional y las diferencias de política educativa.

FIGURA II

Relación entre puntuación media de competencia lectora y edad media de los alumnos evaluados en PIRLS 2001



Fuente: Mullis et al., 2003, Exhibit 1.1, p.26.

Para garantizar comparaciones internacionales válidas, PISA determina por tanto la población que va a evaluar refiriéndola a la edad de los alumnos, es decir, con referencia a una variable externa al sistema educativo. PISA estudia a alumnos cuyas edades oscilan entre los 15 años y 3 meses y los 16 años y 2 meses en el momento de la evaluación, con independencia del grado o del tipo de institución en los que se encuentren, o de si se trata de educación a tiempo completo o parcial. La utilización de esta edad en PISA, en distintos países y a lo largo del tiempo, permite comparar de forma coherente el rendimiento de los alumnos poco antes de que finalicen la educación obligatoria.

En consecuencia, PISA es capaz de realizar declaraciones sobre los conocimientos y destrezas de individuos que nacieron el mismo año y que a los 15 años de edad siguen escolarizados, pero que han tenido diferentes experiencias de tipo educativo tanto dentro como fuera del centro. El grado escolar en que se encuentren depende de la política del país en materia de entrada en la escolarización y promoción. Es más, en algunos países los alumnos evaluados por PISA representan distintos sistemas, trayectorias o corrientes de educación.

Se han establecido estrictos parámetros técnicos para determinar poblaciones meta nacionales. PISA excluye individuos de 15 años que ya no asisten a una institución educativa, y dada su población «meta» (adolescentes de 15 años que continúan asistiendo a clase) su cobertura es muy amplia comparada con otros estudios internacionales: relativamente pocos centros no son aptos para participar (por ejemplo, debido a su lejanía geográfica o a que los alumnos tienen necesidades especiales). En 24 de los 41 países participantes el porcentaje de exclusiones en el ámbito de los centros ascendía a menos del 1%, y en todos los países a

menos del 3%, excepto en México (3.6 %), Suiza (3.4 %), Reino Unido (3.4 %) y los países socios/adjuntos Letonia (3.8 %) y Serbia (5.3 %). En lo que concierne a la exclusión, dentro de un centro, de alumnos que se ajustan a ciertos criterios internacionales establecidos¹, la tasa de exclusión aumenta ligeramente. Sin embargo, se mantiene por debajo del 2% en 19 países participantes, del 4% en 29 países participantes, por debajo del 6% en todos los países salvo dos, y por debajo del 8% en todos los países (véase el Anexo A3 en OECD, 2004b). Este alto nivel de cobertura contribuye a que los resultados del estudio sean comparables. Por ejemplo, incluso asumiendo que los alumnos excluidos habrían obtenido sistemáticamente peores puntuaciones que los que sí participaron, y que esta relación es moderadamente acusada, una tasa de exclusión del orden de un 5% habría llevado seguramente a una sobreestimación en la puntuación media nacional de menos de cinco puntos. Por otra parte, las exclusiones eran inevitables en la mayoría de los casos. Por ejemplo, en Nueva Zelanda el 2,3% de los alumnos fueron excluidos porque habían recibido menos de un año de instrucción en inglés (a menudo debido a que eran estudiantes extranjeros que pagaban cuota) y por lo tanto no eran capaces de seguir las instrucciones que se impartían durante la evaluación.

El diseño y tamaño específicos de la muestra para cada país fueron calculados para maximizar la eficiencia del muestreo para estimaciones realizadas en el ámbito de los alumnos. En los países de la OCDE, el tamaño de las muestras abarcaba desde 3.350 alumnos en Islandia hasta 30.000 estudiantes en México. Esta selección de muestras fue supervisada de forma internacional y acompañada de rigurosos parámetros en la tasa de participación para garantizar que los resultados de PISA reflejaran las destrezas de los alumnos de 15 años de edad en los países participantes.

DESARROLLO DE ESPECIFICACIONES DE LOS TESTS

Los estudios internacionales de la IEA sobre rendimiento estudiantil comenzaron en 1959 con el Estudio Piloto en Doce Países, en el que se evaluó el rendimiento de los alumnos en Matemáticas, Comprensión Lectora, Geografía, Ciencias y Competencia no Verbal (www.iea.nl). En los muchos estudios que la IEA ha realizado a lo largo de los años, las especificaciones de los tests se basaban en un cuidadoso análisis de los currículos existentes en los países participantes,

⁽¹⁾ Estaba permitido a los países excluir hasta un 2,5% de la población meta deseable a nivel nacional (dentro de los centros) si esos alumnos reunían alguna de las siguientes condiciones:

- si, según la opinión profesional del director del centro o de otro miembro cualificado del personal, eran considerados como discapacitados psíquicos educables, o habían sido definidos como tales mediante test psicológico (incluyendo a alumnos que fueran emocional o mentalmente incapaces de seguir las instrucciones generales dadas en PISA);
- si se hallaban discapacitados física y permanentemente de tal forma que no pudieran tomar parte en la situación evaluativa de PISA (los alumnos discapacitados funcionales que podían responder habían de estar incluidos en el estudio); o
- si se trataba de hablantes no nativos que hubieran estudiado la lengua de la evaluación durante menos de un año.

Para más detalles, véase el Anexo A3 en OECD, 2004b.

para así poder concentrarse en aquellas enseñanzas comunes en todos los países y que, de esta forma, la validez de los tests fuera lo más equiparable posible, aunque esto dependiera asimismo de que el tiempo dedicado a las materias comunes dentro de los currículos de los diferentes países fuera comparable también.

En PISA, los países de la OCDE adoptaron un enfoque distinto. En lugar de comprobar si los alumnos dominan o no conocimientos y destrezas esenciales que puedan haber sido incluidos en los currículos que dieron forma a su escolarización, la evaluación se concentra en la capacidad de los alumnos de 15 años para reflexionar y utilizar las destrezas que hayan desarrollado (OECD, 1999, 2003). En el caso de ciencias, por ejemplo, el estudio se centra en si los alumnos saben identificar preguntas científicas, si saben qué se puede considerar como prueba en la que fundamentar sus respuestas, y si saben extraer conclusiones basadas en tales pruebas. Comisiones de expertos internacionales elaboraron los marcos del estudio para cada área: Lectura, Matemáticas y Ciencias en PISA 2000, y en PISA 2003 además, resolución de problemas.

Si bien el diseño de los estudios PISA no empieza por los currículos nacionales, dichas evaluaciones siguen guardando relación con los intereses de esos currículos. Ningún país admitiría que sus currículos no están centrados por completo en que sus alumnos de 15 años desarrollen su capacidad para utilizar las destrezas adquiridas en escenarios distintos de los definidos dentro del propio contenido curricular. La OCDE (2000, 2001, 2003 & 2004b) proporciona muestras de ítems de PISA .

El marco y la base conceptual de cada área de evaluación han sido desarrollados por expertos internacionales procedentes de los países participantes, y han sido aprobados, tras la correspondiente consulta, por los gobiernos de dichos países (OECD, 1999a y OECD, 2003e). El marco da comienzo con el concepto de alfabetización, que atañe a la capacidad de los alumnos para aplicar conocimiento y destrezas, y para analizar, razonar y comunicarse de forma efectiva cuando plantean, resuelven e interpretan problemas en situaciones diversas.

El concepto de alfabetización empleado en PISA es mucho más amplio que la idea tradicional de la capacidad de leer y escribir. A efectos de medida, se considera como un continuo, no como algo que un individuo tiene o no tiene. Puede que para determinados objetivos sea conveniente o necesario definir un punto en el continuo, por debajo del cual se consideren inadecuados los niveles de competencia, pero la variabilidad subyacente es importante. Una persona alfabetizada posee un abanico de competencias, y no existe una línea precisa de división entre una persona plenamente alfabetizada y una que no lo es.

La adquisición de estas capacidades es un proceso que dura toda la vida; no tiene lugar sólo en el centro educativo ni mediante la educación formal en exclusiva, sino también a través de la interacción con nuestros iguales, colegas, y otras comunidades más amplias. No se puede esperar que los quinceañeros hayan aprendido ya todo lo que van a necesitar saber como adultos, pero deberían tener una sólida base de conocimientos en áreas como la Lectura, las Matemáticas y las Ciencias. Para poder seguir aprendiendo en estos campos de

materias y aplicar lo aprendido en el mundo real, necesitan también comprender procesos y principios fundamentales, y saber utilizarlos con flexibilidad en diferentes situaciones. Ésta es la razón de que PISA evalúe la capacidad de realizar tareas relacionadas con la vida real, recurriendo a una comprensión amplia de conceptos clave, en lugar de limitar el estudio a la posesión de conocimiento específico de unas materias.

Las áreas de evaluación tratadas en PISA se definen en función de:

- el *contenido o estructura* de conocimiento que necesitan adquirir los alumnos en cada área de evaluación (familiaridad con conceptos matemáticos);
- los *procesos* que necesitan realizar (desarrollar un argumento matemático); y
- las *situaciones* en las que se enfrentan a problemas matemáticos y aplican sus destrezas y conocimientos relevantes (tomar decisiones relacionadas con la propia vida personal o comprender asuntos a escala mundial).

Una descripción detallada de lo tratado en las áreas de Matemáticas, Ciencias y Lectura se encuentra en OECD, 2003e. La tabla I resume el núcleo de la definición de cada área de competencia, y cómo se desarrollan las tres dimensiones en cada caso.

TABLA I
Resumen de las áreas evaluadas en PISA 2003

Área evaluada	Matemáticas	Ciencias	Lectura
Definición y rasgos distintivos	«La capacidad de identificar y comprender el papel que las Matemáticas tienen en el mundo, emitir juicios bien fundados y usar y manejar las Matemáticas de un modo que se ajuste a las necesidades vitales de la persona como ciudadana constructiva, responsable y reflexiva» (OECD, 2003e). Relacionado con un uso más amplio, funcional, de las Matemáticas, su manejo requiere la capacidad de reconocer y formular problemas matemáticos en diversas situaciones.	«La capacidad de usar conocimientos científicos, de identificar cuestiones científicas y de extraer conclusiones basadas en pruebas para poder comprender y ayudar a tomar decisiones sobre el mundo natural y los cambios introducidos en él mediante la actividad humana» (OECD, 2003e). Requiere la comprensión de conceptos científicos, y una capacidad para aplicar una perspectiva científica y pensar de forma científica sobre las pruebas.	«La capacidad de comprender, utilizar y reflexionar sobre textos escritos para así alcanzar las metas propias, de desarrollar el conocimiento y potencial de uno mismo, y de participar en la sociedad» (OECD, 2003e). Mucho más que decodificación y comprensión literal, la lectura supone comprensión profunda y reflexión, y la capacidad de utilizar la lectura para llevar a cabo los objetivos de uno mismo en la vida.
Dimensión Contenido	Grupos de áreas y conceptos de Matemáticas relevantes: <ul style="list-style-type: none"> • cantidad; • espacio y forma; • cambio y relaciones; e • incertidumbre. 	Áreas de conocimiento científico y conceptos, tales como: <ul style="list-style-type: none"> • biodiversidad; • fuerzas y movimiento; y • cambio fisiológico. 	La forma de los materiales de lectura: <ul style="list-style-type: none"> • los materiales continuos incluyen diferentes tipos de prosa, tales como narración, exposición, argumentación; y • los textos no continuos, como gráficas, fórmulas, listas.

(continúa)

TABLA I
Resumen de las áreas evaluadas en PISA 2003 (continuación)

Área evaluada	Matemáticas	Ciencias	Lectura
Dimensión Proceso	<p>«Grupos de competencia» definen las destrezas necesarias en Matemáticas:</p> <ul style="list-style-type: none"> • reproducción (operaciones matemáticas simples); • conexiones (relacionar ideas para solucionar problemas sencillos); y • reflexión (pensamiento matemático más amplio). <p>En general, están asociados con tareas de dificultad creciente, pero hay solapamiento en la clasificación de tareas en cada grupo.</p>	<p>La capacidad de usar conocimientos científicos y comprensión para adquirir pruebas, interpretarlas y actuar en consecuencia:</p> <ul style="list-style-type: none"> • describir, explicar y predecir fenómenos científicos; • entender investigaciones científicas; e • interpretar pruebas científicas y conclusiones. 	<p>Tipo de tarea o proceso de lectura:</p> <ul style="list-style-type: none"> • recuperar información; • interpretar textos; y • reflexión y evaluación de textos. <p>PISA centra su atención en la lectura para aprender, en lugar de en aprender a leer, y por consiguiente los alumnos no son evaluados en las destrezas lectoras más básicas.</p>
Dimensión Situación	<p>Las situaciones varían de acuerdo con su «distancia» respecto a las vidas de los individuos. Por orden de cercanía:</p> <ul style="list-style-type: none"> • personal; • educativa y ocupacional; • local y en una comunidad más amplia; y • científica. 	<p>El contexto de las Ciencias, con especial atención a sus usos en relación a:</p> <ul style="list-style-type: none"> • vida y salud; • la Tierra y su medio ambiente; y • tecnología. 	<p>El uso para el cual se elaboró el texto:</p> <ul style="list-style-type: none"> • privado (vg., una carta personal); • público (vg., un documento oficial) • ocupacional (vg., un informe); • educativo (vg., lectura asociada al centro educativo).

Además de valorar competencias en las tres áreas mencionadas, PISA se propone introducir progresivamente un examen «interdisciplinar» de dichas competencias. PISA 2000 ya comenzó interrogando a los alumnos acerca de su motivación y otros aspectos de su actitud hacia el aprendizaje, su familiaridad con los ordenadores y, bajo la denominación «aprendizaje autorregulado», diversos aspectos de sus estrategias para dirigir y controlar su propio aprendizaje. En PISA 2003 se desarrollaron estos elementos y se complementaron con una evaluación de conocimientos y destrezas para la resolución de problemas. En los estudios PISA subsiguientes, las competencias «intercurriculares», así como el uso de tecnologías de la información, tendrán un papel de creciente importancia.

DESARROLLO DE LOS TESTS Y EVALUACIÓN DE LOS MISMOS PARA SU ADECUACIÓN CULTURAL

En PISA se utilizan diversos tipos de ítems. Se incluyen ítems de tipo test por su eficiencia, pero también se hace amplio uso de preguntas que exigen del alumno que construya la respuesta, en lugar de que reconozca cuál es la correcta. En algunos ítems las respuestas se presentan en forma de dicotomías, como correcto/falso; en otros, la puntuación múltiple/numérica permite puntuar respuestas que son sólo parcialmente correctas.

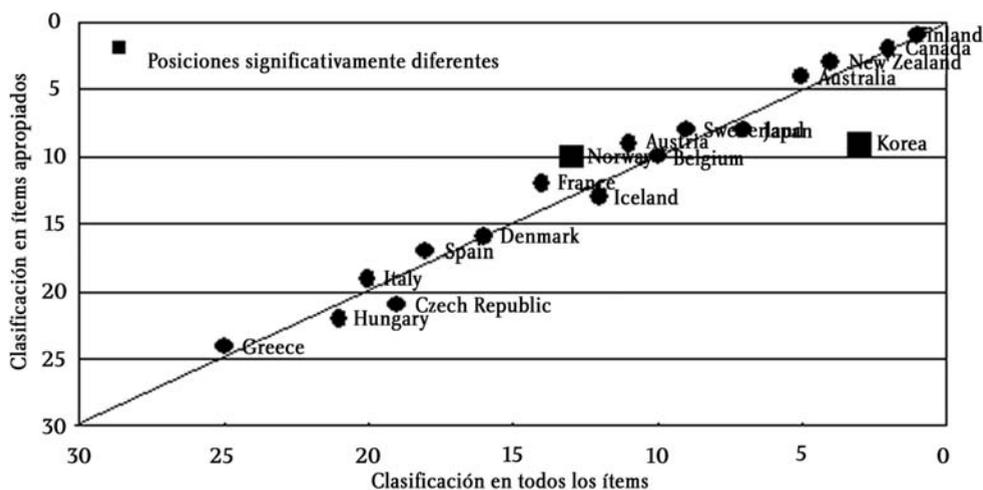
El desarrollo de los tests mismos fue una tarea de gran importancia que trascendía el desarrollo de los marcos de evaluación. En PISA, el trabajo detallado es realizado por un consorcio internacional de agencias de investigación y desarrollo, contratado para este fin, dirigido actualmente por el Consejo Australiano para la Investigación en Educación y que incluye socios en Japón, los Países Bajos y los Estados Unidos. Los países participantes y los miembros del consorcio aportan ítems de prueba. Todos los ítems son sometidos a una revisión para identificar el sesgo cultural u otras insuficiencias; los ítems ofensivos se retiran o se modifican. Todos los ítems que sobreviven a la revisión son puestos entonces a prueba en todos los países. Se lleva a cabo una segunda revisión a la búsqueda de sesgo cultural, mientras en distintos países está en marcha una prueba para cada ítem diferencial. Esencialmente, ésta sirve para identificar ítems que se comportan de distinta manera (en comparación con otros) en determinados países. Aquéllos con un comportamiento extraño son eliminados del fondo común, a partir del cual se construyen los tests finales.

En PISA 2000 fue posible hacer un análisis a posteriori del sesgo cultural porque, antes de que se hubiera realizado la prueba sobre el terreno y se hubiera aportado alguna prueba empírica de la dificultad que ofrecían los ítems en las distintas naciones, se invitó a todos los países a valorar, en una escala de cinco puntos, la adecuación de cada ítem para sus propios alumnos, tomando en consideración la similitud con el currículo, la relevancia cultural, el nivel de interés, la preocupación por la sensibilidad, los problemas de traducción y adaptación, y los problemas de codificación o marcado. Dieciocho de los países participantes facilitaron sus respuestas.

Una vez que se concluyó el análisis completo de los resultados de PISA 2000, el consorcio PISA repitió el análisis 18 veces, empleando sólo aquellos ítems a los cuales cada país había asignado un cuatro ó un cinco en el índice de adecuación. Una manera sencilla de observar los resultados es la comparación de, por un lado, el puesto obtenido por los países en todos los ítems, y por otro, el conseguido basándose exclusivamente en aquellos ítems puntuados por ellos con un cuatro ó un cinco. Esta comparación se muestra en la figura III. Los países por encima de la diagonal lograron un puesto más elevado en sus ítems «preferidos» que el obtenido en todos los ítems. Se podría esperar que todos los países se situaran por encima de la diagonal. De hecho, cinco se sitúan por debajo y cuatro sobre ella; en ambos casos poseen el mismo puesto en la clasificación. Sólo nueve se encuentran por encima. Sin embargo, estas diferencias no eran significativas para ningún país, exceptuando a dos de ellos. España, por ejemplo, obtiene prácticamente el mismo puesto en sus ítems «favoritos» que de forma global en la evaluación PISA, lo cual indica que un test distinto no habría cambiado mucho la posición global de España. Los dos países que presentaron diferencias significativas fueron Corea y Noruega: ésta tenía el puesto décimotercero en todos los ítems y el décimo en sus «favoritos»; Corea obtuvo, respectivamente, los puestos tercero y noveno. Aquello que Corea más valora y que, en su opinión, mejor conviene a sus necesidades, no resultaron ser las tareas en las que descollan sus alumnos de 15 años.

FIGURA III

Comparación por países entre las posiciones obtenidas en todos los ítems de Lectura de PISA 2000 y aquéllos juzgados como más apropiados.



Este análisis proporciona una interesante prueba empírica de que los materiales de la evaluación no tienen ningún sesgo cultural grave. Análisis posteriores en los que se emplearon sólo, por una parte, los ítems elaborados originalmente en inglés, y por otra, los ítems elaborados originalmente en una lengua distinta del inglés, no mostraron diferencias significativas en la clasificación nacional con respecto a aquellos análisis en que se utilizaron todos los ítems, lo cual prueba y corrobora la ausencia de sesgo cultural.

INTERPRETACIÓN DE LAS MEDIDAS RESULTANTES

A partir de los modelos psicométricos y los procedimientos analíticos utilizados en PISA, se pueden desarrollar escalas de rendimiento bien definidas, que permiten describir de forma sustancial las competencias de los individuos en diversos niveles de rendimiento. Para facilitar las descripciones y la interpretación de las escalas, se definen en éstas amplios niveles de rendimiento.

El desarrollo de dichas escalas es esencial para programas de evaluación que deseen hacer un seguimiento de los cambios a lo largo del tiempo. Si el rendimiento se dispone sólo como relación de un país con otro, en la clásica tradición donde la norma es la única referencia, entonces la mejora de uno no puede ser más que relativa, y depende por fuerza del descenso relativo de otro. Si todos mejoraran en el mismo sentido, no puede considerarse que haya mejorado ninguno, ya que ninguno habría mejorado en relación a los demás. Consignar el rendimiento con escalas bien definidas, y relacionar entonces las escalas empleadas en ocasiones sucesivas mediante algunos ítems comunes, con los tests utilizados en cada una de las ocasiones, permite hacer un seguimiento del cambio absolu-

to, preferible al relativo. La OCDE (2004b) informa sobre la incidencia de tales cambios entre PISA 2000 y PISA 2003, y muestra a Polonia como el único país que presenta resultados de mejoría en todas las escalas empleadas en ambas ocasiones. Para España, los resultados se mantienen esencialmente invariables tanto en términos absolutos como relativos.

EL IMPACTO DE PISA EN POLÍTICA EDUCATIVA

La presentación selectiva en el presente artículo de algunos de los análisis y conclusiones procedentes de las dos primeras rondas de PISA (2000 y 2003) ilustran el modo en que PISA contribuye al debate político. Se puede encontrar un resumen más completo en la OCDE (2004b).

La simple comparación entre los rendimientos de los países es probablemente el análisis menos interesante, aun cuando es el que antes capta la atención de los medios de comunicación y del público, y también el que de forma más inmediata provoca el deseo de actuar cuando los resultados son decepcionantes. Se pueden elevar las expectativas nacionales con la prueba de que se están logrando más cosas en otros lugares.

Los resultados de PISA 2000 provocaron un intenso debate en muchos países. Una respuesta inmediata del Ministerio Federal Alemán de Educación e Investigación fue encargar un estudio multilateral, dirigido por Eckhard Klieme, quien reclutó colaboradores en países con los que Alemania deseaba una comparación más detallada, esto es, Canadá, Inglaterra, Finlandia, Francia, los Países Bajos y Suecia. Su estudio conecta los resultados de PISA 2000 con pruebas cualitativas en mediciones tales como estrategias para la reforma e innovación educativa; cuestiones de administración y asignación de recursos; enfoques nacionales para la determinación de estándares, evaluación y seguimiento del sistema; la organización de sistemas de apoyo; el desarrollo profesional de los profesores; y enfoques para tratar las diferencias socioeconómicas en los contextos de los alumnos (OECD, 2004e).

Dinamarca estaba asimismo decepcionada con sus resultados en PISA 2000 porque, respecto al rendimiento, se encontraba tan sólo en la media de la OCDE, a pesar de haber gastado por cada alumno de 15 años más que cualquier otro país (salvo dos), y ciertamente más que en muchos otros que habían obtenido resultados mucho mejores. Dinamarca invitó a la OCDE a realizar una revisión general de su política bajo los auspicios de su programa de revisiones de políticas educativas nacionales. En dicha crítica, llevada a cabo por un grupo de expertos, se concluyó que el sistema carecía de cultura de evaluación, y se señaló que el nivel general de satisfacción con el sistema dentro del país había sido socavado no sólo, y de forma más sustancial, por los resultados de PISA 2000, sino que ya había sido desafiado por comparaciones internacionales anteriores (OECD, 2004a). Tras discutir el informe con el grupo de expertos y el Comité de Educación de la OCDE, el ministro danés indicó que el gobierno llevaría a la práctica las reformas propuestas.

Estos no son más que dos de los muchos análisis y reflexiones posteriores emprendidos por países que han participado en PISA hasta la fecha. En la página web de PISA se pueden encontrar todos los detalles de los informes nacionales y de los informes temáticos internacionales encargados por países conjuntamente a través del Consejo Directivo de PISA.

El seguimiento de los resultados de PISA por parte de los medios de comunicación ha sido extraordinario en muchos países. En la presentación de los resultados de PISA 2003 se prestó menos atención a las comparaciones entre países y al orden en la clasificación que con PISA 2000, aunque por supuesto seguía estando presente. Muchos reportajes dedicaron su atención a los análisis más complejos, como aquéllos que relacionan calidad y equidad. A éstos están dedicando también su atención las administraciones educativas, tanto en discusiones nacionales como internacionales. Cuando PISA 2006 aporte un tercer conjunto de datos para 28 de los 30 países de la OCDE², el análisis y la discusión se ampliarán, incluyendo en sus consideraciones las tendencias observables, y avanzarán más allá de los análisis de cambios ofrecidos por PISA 2000 y 2003 facilitando pruebas de la estabilidad de esos cambios.

ANEXO: EL DESARROLLO DE PISA Y SU PUESTA EN PRÁCTICA UN ESFUERZO CONJUNTO DE COLABORACIÓN

PISA es un esfuerzo conjunto de colaboración, que reúne a expertos científicos procedentes de los países participantes, dirigido conjuntamente por sus gobiernos, sobre la base de intereses compartidos que tienen la política como motor.

El Consejo Directivo de PISA, en el que está representado cada país participante, determina, en el contexto de los objetivos de la OCDE, las prioridades políticas de PISA, y supervisa la adhesión a ellas durante la puesta en práctica del programa. Ésta incluye fijar las prioridades para desarrollar los indicadores, para establecer los instrumentos de evaluación y para dar cuenta de los resultados.

Expertos de países participantes también toman parte en grupos de trabajo que se encargan de unir los objetivos políticos con la mejor pericia técnica disponible internacionalmente. Al participar en estos grupos de expertos, los países se aseguran de que los instrumentos son válidos en el ámbito internacional y que tienen en cuenta los contextos culturales y educativos existentes en los países miembros de la OCDE; de que los materiales de evaluación poseen grandes facultades de medición; y de que los instrumentos ponen énfasis en la autenticidad y validez educativas.

A través de los Administradores Nacionales del Proyecto, los países participantes llevan PISA a la práctica en el ámbito nacional según los procedimientos administrativos acordados. Los Administradores cumplen una función vital, pues garantizan la alta calidad del proceso de aplicación, además de verificar y evaluar los resultados, análisis, informes y publicaciones del estudio.

⁽²⁾ Turquía y la República Eslovaca dispondrán sólo de los datos de 2003 y 2006, ya que ninguno participó en PISA 2000. La República Eslovaca se unió a la OCDE en el año 2000.

El diseño y puesta en práctica de las evaluaciones PISA 2000, 2003 y 2006, dentro del marco establecido por el Consejo Directivo de PISA, es responsabilidad de un consorcio internacional, denominado Consorcio de PISA, dirigido por el Consejo Australiano para la Investigación en Educación (ACER en sus siglas en inglés). Otros socios de este consorcio son el Instituto Nacional de los Países Bajos para Mediciones Educativas (Citogroep), el Instituto Nacional para la Investigación en Educación de Japón (NIER), el Servicio de Pruebas Educativas (ETS en sus siglas en inglés) en los Estados Unidos, y WESTAT, igualmente en los Estados Unidos.

La Secretaría de la OCDE tiene la responsabilidad administrativa conjunta del programa, hace un seguimiento diario de la puesta en práctica, actúa como Secretaría del Consejo Directivo de PISA, construye el consenso entre países, y funciona como interlocutor entre el Consejo Directivo de PISA y el Consorcio Internacional encargado de llevar a la práctica las actividades. La Secretaría de la OCDE también elabora los indicadores y análisis, y prepara los informes internacionales y publicaciones en cooperación con el Consorcio de PISA y mediante consulta con los países miembros, tanto en el ámbito político (Consejo Directivo de PISA) como en el de la puesta en práctica (Administradores Nacionales del Proyecto).

REFERENCIAS BIBLIOGRÁFICAS

- GRIFFITH, J. E. (2000): (INES)—a story of progress: Report from external evaluator on progress since 1995 and remaining gaps in the development of the OECD education indicators. Paper prepared for the Fourth General Assembly of the OECD Education Indicators Programme. Tokyo, Japan.
- MULLIS, I. V. S.; MARTIN, M. O.; GONZALES, E. J.; KENNEDY, A. M. (2003): *PIRLS 2001 International Report: IEA's Study of Reading Literacy Achievement in Primary Schools*. Chestnut Hill, M. A., Boston College.
- OECD (1999): *Measuring student knowledge and skills: a new framework for assessment*. Paris.
- (2000): *Measuring student knowledge and skills: the PISA 2000 assessment of reading, mathematical and scientific literacy*. Paris.
 - (2001): *Knowledge and skills for life: First results from PISA 2000*. Paris.
 - (2003): *The PISA 2003 assessment framework—mathematics, reading, science and problem solving knowledge and skills*. Paris.
 - (2004a): *Denmark: lessons from PISA 2000*. Paris.
 - (2004b): *Learning for tomorrow's world: First results from PISA 2003*. Paris.
 - (2004c): *Messages from PISA 2000*. Paris.
 - (2004e): *What makes school systems perform? Seeing school systems through the prism of PISA*. Paris.
 - (2005): *Education at a Glance—OECD Indicators 2005*. Paris.