

¿Qué nos dice PISA sobre la educación de los jóvenes en Chile?

Nuevos análisis y perspectivas sobre los resultados en PISA 2006

El proyecto fue coordinado por una secretaría ejecutiva compuesta por:

Leonor Cariola, Gabriela Cares y Ema Lagos.

Con la colaboración de Catalina Covacevich y Johanna Gubler.



Unidad de Currículum y Evaluación
Ministerio de Educación



©Ministerio de Educación, Unidad de Curriculum y Evaluación, SIMCE Derechos Reservados
www.simce.cl

¿Qué nos dice PISA sobre la educación de los jóvenes en Chile?

Primera Edición de 1000 ejemplares, Santiago, Agosto de 2009

Inscripción en Cámara Chilena del Libro código I S B N 978-956-292-235-7

Registro Propiedad Intelectual Inscripción N°182700

Corrección de estilo por Ana María Amengual

Diseño por Laboratorio de Marketing

Impreso en Chile por Gráfica 7 Ltda.

Agradecimientos

La Secretaría Ejecutiva del proyecto desea agradecer la colaboración de todas las personas que han hecho posible el desarrollo del Proyecto “Análisis en Profundidad de los resultados de PISA 2006”, que culmina con esta publicación. No podremos mencionarlos a todos, pero sí queremos consignar que sin el apoyo del Comité Técnico y Editorial, este trabajo no habría sido posible. Su aporte con comentarios y orientaciones a los proyectos permitió retroalimentar a los autores de manera sustantiva.

Asimismo queremos agradecer a dichos autores que con entusiasmo formaron parte de este proyecto que, como es natural en estos casos, les exigió capacidad de adaptación a los requisitos de formatos y plazos. También valoramos el esfuerzo de aquellos que, por diversas razones, no pudieron llegar a la etapa final.

Hacemos un reconocimiento a todos nuestros compañeros y compañeras del SIMCE, y a los miembros de la Coordinación UCE, del Ministerio de Educación, que se comprometieron con el proyecto y que directa o indirectamente lo apoyaron. Especiales reconocimientos a Pedro Montt, Coordinador de la UCE, quien logró salvar todo tipo de barreras para obtener financiamiento y nos apoyó desde el Comité Técnico. Debemos destacar el impulso inicial que dio Lorena Meckes al proyecto, mientras era Coordinadora del SIMCE. En ese momento, Emiliana Vega, del Banco Mundial, también estimuló la concretización del proyecto. En particular, agradecemos las facilidades otorgadas por Juan Bravo, quien sucedió a Lorena y ha sido el Coordinador del SIMCE mientras el proyecto estuvo en ejecución.

La Secretaría Ejecutiva valora y agradece a los otros miembros de Estudios Internacionales, Catalina Covacevich y Johanna Gubler, quienes constituyeron un apoyo indispensable durante la edición y diagramación del libro. También a Patricia Chaos, encargada de la gestión, quien hizo posible lo que parecía imposible, a lo largo de todo el proyecto. Manifestamos agradecimiento a las secretarías del SIMCE, que han colaborado en todo momento. Claudia Bizama estuvo permanentemente enfrentada a los requerimientos del proyecto y encabeza esta lista por ser la secretaria oficial de Estudios Internacionales en el SIMCE. Junto a ella y en orden alfabético: Liliana Espinosa, M.Rita Ocaranza y Ginette Prieto.

Secretaría Ejecutiva

Leonor Cariola
Gabriela Cares
Ema Lagos

Comité técnico y editorial

Jorge Aedo	Guillermo Fuentes	Lorena Meckes	Dagmar Raczyński
Ernesto Alabarce	Jacqueline Gysling	César Muñoz	Pilar Romaguera
Leonor Cariola	Jorge Manzi	Agustina Paglayan	Loreto Sazo
Giuliana Espinosa	Claudia Matus	Harry Patrinos	Xavier Vanni
Carolina Flores	Pedro Montt	Dino Plaza	

Presentación

El presente documento es fruto de un esfuerzo nacional encabezado por el Ministerio de Educación que busca dar a conocer los análisis en profundidad y los principales hallazgos y lecciones para la política educativa, que dejan los resultados de la prueba PISA 2006, en la que Chile participó junto a otras 56 naciones, que representan un tercio de la matrícula mundial de alumnos de 15 años.

Cada vez que se liberan resultados de estudios internacionales que evalúan el aprendizaje de los estudiantes, desde el Ministerio se hace un esfuerzo por publicar un reporte nacional que presente la información más relevante para Chile¹. Sin embargo, no parece suficiente, toda vez que los análisis hechos en el país sobre la base de estos datos internacionales han sido escasos².

Los resultados internacionales de PISA 2006 fueron entregados por la OCDE en diciembre de 2007³. Para esa fecha, el SIMCE preparó un resumen ejecutivo con los principales resultados de Chile, indicando que se requerían análisis más complejos y de mayor profundidad para obtener evidencia que pudiera guiar acciones futuras en la política educacional del país. En consecuencia, se desarrolló un proyecto específico con esta intención.

1 Tal es el caso de las siguientes publicaciones de la Unidad de Currículum y Evaluación: “Competencias para la vida. Resultados de los estudiantes chilenos en el estudio pisa 2000”; “El Estudio Internacional de Educación Cívica 1999. Informe de Resultados Nacionales Estudio Educación Cívica 14 años”; “Estudio Internacional de Educación Cívica 1999. Resultados de interés para el subsector orientación del 2° ciclo de Enseñanza Básica”; “Educación Cívica y el Ejercicio de la ciudadanía. Los estudiantes chilenos de 4° medio en el Estudio Internacional de Educación Cívica”; “Resumen Ejecutivo: Informe Nacional de Chile Educación Cívica y el Ejercicio de la Ciudadanía (4° Medio)”; “Chile y el Aprendizaje de Matemáticas y Ciencias según TIMSS. Resultados de los estudiantes chilenos de 8° básico en el Estudio Internacional de Tendencias en Matemáticas y Ciencias 2003”

2 La búsqueda en internet, además de las publicaciones SIMCE, arroja dos artículos para TIMSS y algunos comentarios en la prensa. Para PISA, aparecen distintas publicaciones del Ministerio e información en la prensa respecto de los resultados.

3 OCDE. PISA 2006. Competences for the World of Tomorrow

Descripción del *Proyecto de Análisis en Profundidad de PISA 2006*

Una vez conocidos los resultados de PISA 2006, el Ministerio de Educación, a través de su Unidad de Currículum y Evaluación, más específicamente del Equipo de Estudios Internacionales del SIMCE, se propuso la tarea de realizar un estudio en profundidad de esos resultados con un doble propósito. Por una parte, para abrir un espacio en el que equipos de investigadores de las más diversas sensibilidades y visiones tuvieran la oportunidad de realizar un trabajo colaborativo que maximizara el aprovechamiento de los resultados PISA 2006, a través de análisis secundarios de sus datos. Por otra parte, para contribuir al posicionamiento de la evidencia que se deriva de esta evaluación de aprendizajes y a la toma de decisiones informadas en el diseño de la política educacional.

Con la opción de hacer una convocatoria amplia, se quiso incorporar al análisis de estos datos, diversas perspectivas y propuestas metodológicas, enfoques y disciplinas, así como la acumulación de conocimiento y capacidades de investigación en esta dimensión estratégica para la política educativa.

A esta iniciativa se la denominó *Proyecto de Análisis en Profundidad de PISA 2006* y se inició a mediados de 2008, con una amplia invitación a centros de estudios y de investigación que se ocuparan de temas educacionales, así como a facultades de educación de diversas Universidades. En esta convocatoria también se incluyó a los equipos internos del SIMCE, bajo el supuesto de que sus conocimientos técnicos los ayudarían a analizar desde una mirada distinta las pruebas PISA y SIMCE, y considerando que el carácter confidencial de los instrumentos solo permitía el acceso a estos profesionales.

Esta publicación es producto de ese esfuerzo conjunto entre el Ministerio de Educación e investigadores que asumieron el desafío y concordaron en llevar a cabo el proyecto.

Con una asistencia cercana a 50 personas, en junio de 2008 se hizo una reunión explicativa del proyecto y de los resultados de PISA en Chile y en el resto de los países, señalando algunas posibilidades y límites de potenciales estudios. En esa oportunidad, se pusieron a disposición los resultados generales, el total de información disponible, así como las restricciones técnicas que los datos y los análisis podían tener. El proyecto fue coordinado por una Secretaría Ejecutiva, compuesta por tres miembros del Equipo de Estudios Internacionales del SIMCE, que contó con la valiosa y desinteresada cooperación de un Comité Técnico. Este estuvo compuesto por un connotado grupo de investigadores y expertos que accedieron

a apoyar con sus comentarios críticos durante todo el tiempo que duró el trabajo. Especial mención se debe hacer a expertos del Banco Mundial —que colaboraron muy activamente en esta iniciativa— y al apoyo logístico y organizativo de la OEI (Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura).

Después de la reunión informativa, el desarrollo del proyecto tuvo los siguientes hitos:

- Taller de tres días (15, 17 y 18 de julio de 2008) para el uso de las bases de datos PISA. Dirigido por la Dra. Claudia Matus, Asesora Estadística del SIMCE, al que asistieron alrededor de 25 personas. El taller entregó las herramientas estadísticas para poder realizar los análisis con la confiabilidad y validez necesaria.
- Taller de presentación de proyectos, en el cual se expusieron y comentaron 15 proyectos. Del 25 al 28 de agosto, 2008. Con posterioridad al taller, los comentarios se enviaron por escrito y se sostuvo un permanente intercambio de investigadores con la Secretaría Técnica, hasta considerar que el proyecto era viable.
- Entrega escrita, por parte de los investigadores, de las metodologías empleadas y de sus principales resultados, al Comité Técnico. Dicho Comité envió sus comentarios por escrito. (Noviembre de 2008).
- Jornadas de retroalimentación (diciembre 9 al 12 de 2008), que tuvieron como objetivo: retroalimentar las investigaciones de cada equipo; informar a los equipos sobre la viabilidad de publicación de sus investigaciones; orientar a los investigadores para llegar a una publicación de calidad, y obtener información para elaborar las conclusiones generales de la publicación.
- En marzo de 2009, los investigadores debían entregar sus artículos definitivos. Cada uno de ellos recibió comentarios escritos de a lo menos tres miembros del Comité Técnico y de la Secretaría Ejecutiva. Luego de incorporados esos comentarios, los artículos entraron en proceso de edición y diagramación.

En todo este proceso el Comité Técnico y la Secretaría Ejecutiva del proyecto hicieron comentarios y advertencias. Inicialmente, respecto del planteamiento del problema y la metodología propuesta. Con el avance del proyecto, los comentarios estuvieron dirigidos a aspectos sustantivos, así como a la organización de los artículos y a la fundamentación

de las conclusiones. El espíritu con que se abordó esta tarea no fue de imposición, sino de colaboración, para aprovechar al máximo la información que se desprendía de los análisis. Es por esto que con toda propiedad, se puede afirmar que tanto el mérito como la responsabilidad, sobre lo dicho en los artículos, corresponde exclusivamente a los autores.

Organización de la publicación

La presente publicación se inicia con un capítulo que cumple un doble propósito: (a) describir las características generales de PISA, la forma como se organiza y administra en Chile y elementos relevantes para la adecuada comprensión de los artículos y (b) exponer los resultados generales obtenidos por los estudiantes chilenos, en la aplicación 2006 -en las distintas áreas de aprendizaje: Ciencias, Lectura y Matemática- así como antecedentes relevantes sobre el marco de referencia de la prueba PISA de Ciencia, ya que ésta fue la principal área evaluada ese año.

Luego, se incluyen los artículos que componen la publicación, los cuales tienen en común el uso de datos de PISA 2006 y la inclusión de los resultados de Chile en sus análisis. Pese a este vínculo de base, los artículos son diversos: en algunos casos se enfocan especialmente en los estudiantes chilenos y en otros, comparan los resultados de los estudiantes chilenos con los de otros países de la OCDE y/o de la Región o de los países participantes en general. Asimismo, los análisis se realizan desde distintas perspectivas disciplinarias y usando diversas metodologías. Esta diversidad, es posible que también redunde en una diversidad de públicos objetivos: responsables de políticas educacionales, docentes, expertos en evaluación e investigadores educacionales. La variedad surgió por el especial interés de fomentar investigaciones y de no restringir las iniciativas de los autores. Esperamos que esta característica convoque a un mayor número de lectores, que si solo estuviera dirigido a un grupo específico de especialistas.

Dada la situación descrita, la agrupación en tres secciones, que se hizo de los artículos, puede considerarse bastante arbitraria, pero el criterio utilizado fue que compartieran los focos o el objetivo último de los análisis.

Así, en la primera sección, se incluyen los artículos relativos a la búsqueda de factores asociados que expliquen los rendimientos de los estudiantes -evidentemente sin atribuir causalidad, pero buscando orientar la política hacia donde podría ser más efectiva, ya sea a nivel de sistema o de escuela. En esta sección, el tema de la equidad no está ausente, ya

que la igualdad de oportunidades es un indicador relevante en la calidad de la educación impartida por el sistema nacional. En efecto, dado que Chile aparece como uno de los países en que los aspectos socioeconómicos estarían más relacionados con las diferencias en los rendimientos de los alumnos, la búsqueda de otros factores justamente permite encontrar variables más fáciles de alterar que el nivel de ingreso o la desigualdad económica.

En esta primera sección, dos artículos se refieren al desempeño en Ciencias. En “*¿Cómo las escuelas chilenas pueden mejorar el aprendizaje en ciencias?*” (de Ernesto Treviño, Francisca Donoso y Macarena Bonhomme) se hace un análisis jerárquico de regresión múltiple⁴ que considera variables a nivel del estudiante y de la escuela. En el otro artículo referido al desempeño de Ciencias, se indaga sobre el uso de computadores y su asociación con el desempeño, mediante un análisis descriptivo y econométrico (“*Evidencia sobre uso de tecnologías y su correlación con desempeño en PISA- Ciencias 2006*”, de Martha Kluttig, Claudia Peirano y Constanza Vergara). También en esta primera sección, el equipo CIAE de la Universidad de Chile, analizó el desempeño de Matemática y Lectura, a través de la descomposición de la varianza, para determinar si el menor rendimiento de Chile -en relación con España, Uruguay y Polonia- obedecía a las diferencias en recursos y características de los estudiantes o las escuelas, o más bien a la eficiencia en el uso y aprovechamiento de tales recursos y características. Los dos artículos del equipo CIAE (el de Matemática y el de Lectura) comparten la introducción, los antecedentes y la metodología, pero los resultados no son los mismos, por lo que se deben considerar como complementarios.

En la segunda sección, se presentan cuatro artículos escritos por profesionales que trabajan en el SIMCE, y que abordan las características y objetivos propios de la prueba PISA y el sentido que ésta puede tener para los profesores de aula y para la formación docente inicial. Estos artículos incluyen análisis menos cuantitativos y más cercanos a la evaluación. En uno de ellos, “*Análisis comparativo de la estructura interna y objetivos de evaluación de las pruebas de lectura SIMCE y PISA*” (de Cristián Donoso y Pablo Lima), se comparan aspectos de formato de las preguntas incluidas en la prueba, como también conocimientos y habilidades evaluadas, de las pruebas PISA y SIMCE, con el objeto de determinar cuáles son las verdaderas diferencias entre ellas y comprobar si, desde un punto de vista cualitativo, se puede hipotetizar respecto de las razones de las diferencias de tendencia entre los resultados de PISA y SIMCE, en el período 2001-2006. En el artículo “*Análisis comparativo de los*

⁴ Se refiere a una metodología que pretende determinar cuánto cambia una variable dependiente por cada unidad de variación de la independiente, suponiendo que los restantes factores en juego no se alteran. Las variables pueden corresponder al individuo o a agrupaciones mayores.

resultados chilenos en las pruebas de Matemática SIMCE 2° Medio 2006 y PISA 2006” (de Luis Alfaro y Raúl Gormaz), se intenta explicar la mayor dificultad de los estudiantes chilenos para responder a la prueba PISA, comparando los contenidos evaluados en ella y el currículum nacional. Como esta no parece ser la razón, analizan y comparan las características de las preguntas en ambas pruebas, para establecer si existen formatos, contextos o contenidos de algunas preguntas que aumenten la dificultad para los estudiantes chilenos. En otro artículo *“Resultados en Ciencias de los estudiantes chilenos en la Prueba PISA 2006: Una mirada a sus competencias”* (de Johanna Gubler y Alexis Williamson), se indaga sobre los aspectos evaluados en las preguntas de la prueba PISA que pudieran presentar mayores dificultades o fortalezas para los estudiantes chilenos, tomando como referencia los resultados de los estudiantes del promedio de países OCDE. Finalmente, un artículo diferente, tanto por su metodología como por sus objetivos, es el de Gabriela Cares quien fue asistente de la Coordinación de Estudios Internacionales en el SIMCE y miembro de la Secretaría Ejecutiva del proyecto. Este se titula *“Percepciones de los estudiantes de pedagogía sobre PISA”* y constituye un análisis que se basa en focus groups con estudiantes de pedagogía que están finalizando su carrera, para ver cuán familiarizados están con las pruebas externas y estandarizadas, con las pruebas internacionales e incluso con el SIMCE, con el objeto de identificar maneras de aprovechar la información de PISA en las prácticas de aula.

La tercera y última sección incluye tres artículos, cuyo principal énfasis es “evaluar la evaluación” desde un punto de vista metodológico. Es sabido que hay partidarios y detractores de la evaluación estandarizada y externa del aprendizaje, pero los comentarios y críticas están dirigidos a los supuestos efectos o consecuencias de la evaluación y no a la generación de los datos. El ámbito técnico-métrico permanece reducido al campo, aparentemente inexpugnable, de los especialistas. De allí que parece conveniente que se expliquen los riesgos y seguridades, las certezas e incertezas de los resultados que arroja la medición. Con este enfoque, en esta sección se incluye un artículo, *“Rendimiento en la prueba PISA: ¿Es posible entender los alcances y límites de las comparaciones entre países?”* (de Jorge González y Ernesto San Martín), que cuestiona metodológicamente la comparabilidad de los análisis que utilizan modelos jerárquicos entre países, en circunstancias que en cualquier modelo, resulta imposible considerar todas las variables explicativas y que, en general, dichos modelos solo explican los puntajes cercanos a la media. Este artículo plantea desafíos importantes a tener en cuenta en el momento en que se quieran hacer recomendaciones de política, basándose en evidencia proporcionada por este tipo de análisis. También en esta sección se incluye el artículo *“Distribución de puntajes SIMCE en PISA 2006 y SIMCE*

2006 en perspectiva comparada” (de Catalina Fernández, miembro del Equipo de Análisis de Resultados del SIMCE), en el que se valida la representatividad de la muestra PISA de 2º Medio (distinta a la de 15 años). Para ello se hacen diversas comparaciones de resultados SIMCE, utilizando el universo de la prueba censal y la muestra de estudiantes incluidos en la aplicación PISA de 2º Medio. Este análisis permite aclarar una posible duda respecto a la representatividad de la muestra PISA. Finalmente se incluye el artículo titulado *“Evaluación de la eficiencia de los sistemas educativos nacionales en su objetivo de proveer calidad y equidad”* (de Claudio Thieme, Víctor Giménez y Diego Prior), que incluye el análisis de 54 países, con el objeto de explicar las diferencias de eficiencia entre los mismo y, con ello, , elaborar una tipología de países según los tipos de problemas que deben enfrentar para alcanzar el máximo de eficiencia (gestión, recursos o entorno). Este artículo fue incluido en esta sección porque utiliza una metodología novedosa para establecer la eficiencia de los sistemas educacionales y presenta una definición de calidad educacional que incluye tanto aprendizajes, como equidad. La calidad educacional, así entendida, es muy apreciada en términos conceptuales, pero rara vez se operacionaliza para utilizarla como variable dependiente en estudios empíricos.

Luego de las tres secciones antes señaladas, se presentan conclusiones generales y posibles proyecciones de políticas.

Este trabajo busca marcar un hito y abrir una tradición, convirtiendo los futuros estudios sobre resultados de Chile en evaluaciones internacionales, en un espacio de convocatoria nacional y ampliada a las más diversas miradas existentes en el campo de la investigación educativa. Se espera que, desde ahora, estos estudios nacionales sean un esfuerzo participativo y de Estado, por sobre lo que podría hacer un gobierno. Ello adquiere especial importancia, en un momento de nuestra historia educativa en que se avencinan cambios institucionales relevantes como es la creación de una Agencia de la Calidad de la Educación, ente autónomo del Ministerio de Educación, que pasará a formar parte de un sistema institucional diseñado para asegurar la calidad de la educación, en el cual las pruebas estandarizadas nacionales e internacionales son una pieza fundamental. Se espera así, que esta nueva institución, dé continuidad a este esfuerzo de Estado.

Pedro Montt Leiva
Coordinador Nacional
Unidad de Currículum y Evaluación
Ministerio de Educación de Chile

Tabla de contenidos

Capítulo 1: Participación de Chile en PISA y Resultados en 2006	15
Capítulo 2: Factores Asociados	45
Evidencia sobre el uso de tecnologías y su correlación con el desempeño en Pisa-Ciencias 2006. <i>Martha Kluttig, Claudia Peirano, Constanza Vergara</i>	47
¿Cómo las escuelas chilenas pueden mejorar el aprendizaje en Ciencias? <i>Ernesto Treviño, Francisco Donoso, Macarena Bonhomme</i>	71
¿Qué explica las diferencias de resultados PISA Matemática entre Chile y algunos países de la OCDE y América Latina? <i>Juan Pablo Valenzuela, Cristián Bellei, Alejandro Sevilla, Alejandra Osses</i>	105
¿Qué explica las diferencias de resultados PISA Lectura entre Chile y algunos países de la OCDE y América Latina? <i>Cristián Bellei, Juan Pablo Valenzuela, Alejandro Sevilla, Alejandra Osses</i>	149
Capítulo 3: Característica de la Prueba	171
Percepciones de los estudiantes de pedagogía sobre PISA. <i>Gabriela Cares</i>	173
Resultados de los estudiantes chilenos en la prueba PISA Ciencias 2006: una mirada a sus competencias. <i>Johanna Gubler, Alexis Williamson</i>	197
Análisis comparativo de los resultados chilenos en las pruebas de Matemática SIMCE y PISA. <i>Luis Alfaro, Raúl Gormaz</i>	239
Análisis comparativo de la estructura interna y objetivos de evaluación de las pruebas de Lectura SIMCE y PISA. <i>Cristián Donoso, Pablo Lima</i>	261
Capítulo 4: Metodología	287
Evaluación de la eficiencia de los sistemas educativos nacionales en su objetivo de proveer calidad y equidad. <i>Claudio Thieme, Víctor Giménez</i>	289
Distribución de puntajes SIMCE en PISA y SIMCE 2006, en perspectiva comparada. <i>Catalina Fernández</i>	315
Rendimiento en la prueba PISA: ¿es posible entender los alcances y límites de las comparaciones entre países. <i>Jorge González, Ernesto San Martín</i>	327
Capítulo 5: Conclusiones	353



Participación de Chile en PISA y Resultados en 2006

1



Introducción

En la primera parte de este capítulo se entregan algunos antecedentes e informaciones generales sobre la prueba PISA que se consideran necesarios para la lectura de los artículos que componen esta publicación, tales como el sentido de la participación de Chile en este tipo de evaluaciones internacionales; las características generales de la aplicación 2006, y la organización e implementación de la prueba en Chile. En la segunda parte, se describen los resultados globales obtenidos por Chile y a nivel internacional y se explica con algún detalle la composición de la prueba en general y sus tres áreas de evaluación.

1. Antecedentes generales sobre PISA

1.1 La importancia de la participación de Chile en las evaluaciones internacionales de aprendizaje

Los estudios internacionales de evaluación de aprendizajes de los estudiantes tienen larga data en nuestro país. Ya en 1971, Chile participó en el estudio de las seis asignaturas de la IEA¹ y desde 1997, se ha incorporado decididamente a las mediciones de Latinoamérica (a través de LLECE² 1998 y 2004) y del resto del mundo, mediante su participación en TIMSS³ (1999, 2003 y 2011), en los ciclos del Estudio de Educación Cívica de la IEA (1998/2000 y 2009), y en PISA (2000, 2006 y 2009) de la OCDE⁴, prueba que en Chile se implementa desde el Ministerio de Educación, a través del SIMCE-UCE⁵.

Participar en estos estudios constituye un desafío, no solo por el hecho de someter al sistema de evaluación nacional al escrutinio de entidades externas, internacionalmente reconocidas y validadas por su capacidad técnica; sino por cuanto el ingreso en sí mismo, requiere del cumplimiento de una serie de condiciones. Por ejemplo, la IEA exige ser miembro, lo que supone postular y defender la solicitud ante la Asamblea General, y en PISA, se debe enviar una solicitud al PISA Governing Board, que es el órgano de gobierno de PISA, quien analiza si las condiciones y experiencia del país permiten suponer que cumplirá con los estándares establecidos para el programa, y luego de las aplicaciones, se debe demostrar que estos se han cumplido.

La participación de Chile en este tipo de evaluaciones, se fundamenta en un interés múltiple.

Por una parte, Chile aspira a pertenecer a las economías desarrolladas del mundo, para lo cual resulta crucial contrastar los resultados de aprendizaje de sus estudiantes y a la educación que estos reciben, en un contexto más amplio que el meramente nacional. Ello permite obtener una visión más abarcadora y rigurosa de las reales posibilidades de estos jóvenes para insertarse en un mundo globalizado como el actual. En efecto, mediante la aplicación de una batería de instrumentos (pruebas y cuestionarios) así como de procedimientos estandarizados y detallados con precisión, los estudios internacionales someten a todos los países a la misma medición y generan, por tanto, datos válidos, confiables y comparables, que permiten establecer relaciones entre los logros de aprendizaje de los estudiantes de un determinado país con los de otras naciones y conjuntos de naciones.

Por otra parte, mientras las evaluaciones nacionales contrastan los aprendizajes de los estudiantes chilenos con el currículo nacional, las internacionales lo hacen con estándares internacionales y en función de otros países participantes, lo que permite obtener insumos para enriquecer, tanto el diseño curricular, como

1 International Association for the Evaluation of Educational Achievement.

2 Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación.

3 Trends in International Mathematics and Science Study.

4 Organización para la Cooperación y el Desarrollo Económico.

5 SIMCE es el Sistema de Medición para la Calidad de la Educación, del Ministerio de Educación de Chile, dependiente de la Unidad de Currículum y Evaluación (UCE).

las evaluaciones nacionales. De hecho, en Chile, la participación en estos estudios ha tenido un impacto directo en la mejora de los procedimientos de la evaluación nacional SIMCE.

Finalmente, las evaluaciones internacionales incluyen el estudio de variables asociadas a los logros, de modo que las diferencias y semejanzas observadas en el contexto educacional de países con distintos resultados de aprendizaje, permiten orientar las políticas educacionales individuales. Por ejemplo, en Chile, los resultados de TIMSS condujeron a evaluar la necesidad de mejorar el currículum de Educación Básica en Ciencias y Matemática y de proyectar los niveles de logro nacionales. Al desarrollar las descripciones de dichos niveles, a su vez, se consideraron los estándares y resultados, tanto de PISA como de TIMSS.

1.2 La estructura organizacional de PISA

A nivel internacional, PISA cuenta con una serie de órganos colegiados que se ocupan de que este proyecto cumpla los estándares de calidad y con las metas que se ha fijado y que tiene comprometidas con la OCDE. Entre estos órganos, se cuentan el Órgano de Gobierno PISA (PGB), en el que participan representantes de todos los países; el Grupo de Desarrollo Estratégico, y los Grupos de Consejeros Técnicos por áreas evaluadas.

Para el diseño e implementación de la prueba PISA, la OCDE realiza una licitación del proyecto. A esta se presentan distintos organismos de evaluación, conformando normalmente consorcios que actúan como un solo cuerpo, conducido por un equipo coordinador. Hasta ahora, PISA ha sido siempre coordinado por el Consejo para la Investigación en Educación de Australia (ACER). En PISA 2006, el consorcio estuvo compuesto, además de ACER, por el Instituto Nacional

para la Medición en Educación de Holanda (CITO), el Instituto de Evaluación Educacional en Estados Unidos (ETS), el Instituto Nacional para la Investigación en Política Educacional de Japón (NIER) y la compañía norteamericana dedicada a las estadísticas y muestreo, WESTAT. Este consorcio coordinó el desarrollo de la prueba y todos los instrumentos, definió los procedimientos, supervisó el trabajo y produjo las bases de datos con los resultados. Estos procesos los lidera el consorcio, pero el apoyo de los centros y de los coordinadores nacionales de cada país es indispensable para el funcionamiento del proyecto.

Chile participó por primera vez en PISA 2000, inaugurando, junto con otros diez países, la participación en este estudio, de naciones no miembros de la OCDE. En 2003, nuestro país optó por no participar, pero se reincorporó en PISA 2006 y se encuentra en los procesos para la aplicación 2009.

Para la participación en cada uno de los ciclos de PISA, se requiere desde el Estado la inversión de recursos humanos y financieros, que alcanzan un total aproximado de 790 mil euros, incluyendo cuotas de participación, aplicación, viajes para reuniones y honorarios durante los 4 años que dura el estudio. Asimismo, se requieren conocimientos y experiencia, para cumplir con los estándares internacionales.

Como se señaló al comienzo, en Chile, PISA está a cargo del Equipo de Estudios Internacionales del SIMCE (perteneciente a la Unidad de Currículum y Evaluación del Ministerio de Educación⁶). Esto ha sido intencionado así, para facilitar y promover la retroalimentación al sistema nacional de evaluación. Además, esta organización facilita el desarrollo de los estudios internacionales ya que cuenta con expertos de evaluación que prestan apoyo en los distintos procesos.

6 Hay países en que la organización y gestión de PISA está fuera de los ministerios de educación.

1.3 Las principales características de las pruebas PISA

PISA (*Project for International Student Assessment*) es un proyecto de la Organización para la Cooperación y el Desarrollo Económico (OCDE) que funciona desde el año 2000 y que evalúa a los alumnos de 15 años que estén cursando un grado entre 7° y 12°, en el entendido de que están próximos a concluir su educación obligatoria.

PISA busca determinar en qué medida los estudiantes han adquirido las competencias que les permiten enfrentar los retos de la actual sociedad del conocimiento. Para ello, cada tres años, evalúa las áreas de Lectura, Matemática y Ciencias, en términos de la alfabetización de los estudiantes, entendida como la “capacidad de los alumnos para aplicar conocimientos y habilidades, y para analizar, razonar y comunicarse con eficacia cuando plantean, resuelven e interpretan problemas relacionados con distintas situaciones”⁷.

La evaluación de cada una de las tres áreas comprende: definiciones y características distintivas del área; contenidos del conocimiento, competencias o destrezas⁸ requeridas, y contextos o situaciones, como se observa en la Figura 1. En este sentido, una de las características más destacadas de las pruebas PISA es que todos los conocimientos evaluados son puestos en situaciones concretas de la vida real, lo que exige poner en juego un conjunto de conocimientos y las capacidades para realizar procesos con ese conocimiento, en distintos contextos.

7 OCDE (2008). Informe PISA 2006. Competencias científicas para el mundo del mañana. Santiago, Chile: Santillana, p.22.

8 En este contexto, “destreza” debe entenderse como “habilidad” o capacidad para desarrollar una tarea específica.

Cada ciclo PISA profundiza en un área determinada. En PISA 2000, el área principal fue Lectura; en 2003, Matemática y en el año 2006, Ciencias. En esa aplicación, además de medir conocimientos y habilidades científicas en distintas áreas de contenido, incluyó algunas preguntas que medían las actitudes de los estudiantes respecto de su interés y apoyo a la investigación científica, en los mismos temas sobre los cuales se medían conocimientos. En concreto, PISA 2006 recabó evidencia sobre si los estudiantes:

- tienen conocimiento científico y lo utilizan para identificar cuestiones científicas, adquirir nuevos conocimientos, explicar fenómenos científicos y extraer conclusiones basándose en evidencia acerca de problemas relacionados con las ciencias;
- comprenden las características de las ciencias como forma de conocimiento e investigación;
- son conscientes de que las ciencias y la tecnología conforman nuestro medio material, intelectual y cultural, y
- se comprometen, como ciudadanos reflexivos, en problemas e ideas relacionados con las ciencias⁹.

La prueba PISA destina más tiempo y 60% de las preguntas al área principal del ciclo. Con esta mayor información sobre la competencia de los estudiantes, es posible reportar resultados por subáreas. En PISA 2006, se construyeron siete subescalas específicas de Ciencias, además de una escala general para cada área evaluada. Para cada escala y subescala se desarrollaron niveles de logros, asociados a rangos de puntajes, que describen las competencias demostradas por los estudiantes.

9 OCDE, 2008, op. cit, p.23.

Figura 1.
Resumen de las áreas de evaluación de PISA 2006

	Ciencias	Lectura	Matemáticas
Definición y características distintivas	<p>El grado en que un individuo:</p> <ul style="list-style-type: none"> Tiene conocimiento científico y lo utiliza para identificar cuestiones, adquirir nuevos conocimientos, explicar fenómenos científicos y extraer conclusiones basándose en pruebas acerca de problemas relacionados con las ciencias. Comprende las características de la ciencia como forma de conocimiento e investigación. Es consciente de que la ciencia y la tecnología conforman nuestro medio material, intelectual y cultural. Se compromete como ciudadano reflexivo en problemas e ideas relacionados con las ciencias. <p>La <i>competencia científica</i> requiere comprensión de conceptos científicos, capacidad para aplicar un punto de vista científico y pensar sobre las pruebas de una manera científica.</p>	<p>La capacidad de un individuo de comprender, utilizar y analizar textos escritos para conseguir los objetivos propios, desarrollar el conocimiento y el potencial y participar en la sociedad.</p> <p>Además de descodificación y comprensión literal, la <i>competencia lectora</i> implica lectura, interpretación y reflexión, y la capacidad de usar la lectura para cumplir las metas de cada cual en la vida.</p> <p>PISA se centra en leer para aprender más que en aprender a leer, de aquí que no se evalúa a los alumnos sobre las destrezas de lectura más básicas.</p>	<p>La capacidad de un individuo de identificar y comprender el papel de las matemáticas en el mundo actual, emitir juicios bien fundamentados y utilizar las matemáticas y comprometerse con ellas de manera que puedan satisfacer las necesidades de la vida del individuo como ciudadano constructivo, comprometido y reflexivo.</p> <p>La <i>competencia matemática</i> se refiere a un uso más amplio y funcional de las matemáticas: un compromiso con las matemáticas requiere la capacidad de reconocer y formular problemas matemáticos en distintas situaciones.</p>
Contenido del conocimiento	<p><i>Conocimiento de la ciencia:</i></p> <ul style="list-style-type: none"> Sistemas físicos. Sistemas vivos. Sistemas terrestres y espaciales. Sistemas tecnológicos. <p><i>Conocimiento acerca de la ciencia:</i></p> <ul style="list-style-type: none"> Investigación científica. Explicaciones científicas. 	<p>La forma de los materiales de lectura:</p> <ul style="list-style-type: none"> <i>Textos continuos</i>, que incluyen diferentes tipos de prosa, como narración, exposición y argumentación. <i>Textos no continuos</i>, como gráficos, formularios o listas. 	<p>Grupos de áreas y conceptos matemáticos importantes:</p> <ul style="list-style-type: none"> <i>Cantidad.</i> <i>Espacio y forma.</i> <i>Cambio y relaciones.</i> <i>Incertidumbre.</i>
Destrezas implicadas	<p>Tipo de tarea o proceso científico:</p> <ul style="list-style-type: none"> Identificar cuestiones científicas. Explicar fenómenos científicos. Utilizar pruebas científicas. 	<p>Tipo de tarea o proceso de lectura:</p> <ul style="list-style-type: none"> Recuperar información. Interpretar textos. Reflexionar sobre los textos y evaluarlos. 	<p>Los grupos de competencias definen las habilidades necesarias para las matemáticas.</p> <ul style="list-style-type: none"> Reproducción (operaciones matemáticas sencillas). Conexión (conectar ideas para resolver problemas sencillos). Reflexión (pensamiento matemático de mayor amplitud).
Contexto y situaciones	<p>El área de aplicación de la ciencia, centrándose en sus usos relacionados con situaciones personales, sociales y globales, tales como:</p> <ul style="list-style-type: none"> Salud. Recursos naturales. Medio ambiente. Riesgo. Fronteras de la ciencia y la tecnología. 	<p>El uso para el que se elaboró el texto:</p> <ul style="list-style-type: none"> <i>Privado</i> (por ejemplo, una carta personal). <i>Público</i> (por ejemplo, un documento oficial). <i>Profesional</i> (por ejemplo, un informe). <i>Educativo</i> (por ejemplo, lecturas escolares). 	<p>El área de aplicación de las matemáticas, centrándose en sus usos relacionados con situaciones personales, sociales y globales, tales como:</p> <ul style="list-style-type: none"> <i>Personales.</i> <i>Educativas y profesionales.</i> <i>Públicas.</i> <i>Científicas.</i>

Fuente: OCDE, 2008, *op. cit.*, Tabla 1.

Las pruebas de cada área son comparables entre los distintos ciclos, a condición de que haya habido un ciclo previo que evalúe en profundidad el área que se quiere comparar. Por ejemplo, Chile participó en el ciclo 2000, donde el área principal fue Lectura y, por tanto, puede comparar los resultados 2006 de Lectura con los del ciclo 2000. Dado que no participó en 2003, donde PISA puso el énfasis en Matemática, no tienen con qué ciclo comparar su rendimiento en esta área. Como en 2006, fue el primer ciclo que profundizó en Ciencias, los resultados de esta área no tienen referente para hacer comparaciones en el tiempo.

Junto con las pruebas que evalúan competencias, PISA aplica cuestionarios a estudiantes y directores, que recogen información sobre variables de contexto asociadas al rendimiento. En los cuestionarios respondidos por los estudiantes, se indaga principalmente en aspectos socioeconómicos de sus familias, en sus actitudes hacia las ciencias, la lectura o la matemática (dependiendo del énfasis de la aplicación) y en la enseñanza de la disciplina en sus escuelas. Por su parte, en los cuestionarios para los directores, se recaba información sobre aspectos administrativos y pedagógicos y acerca de las percepciones de los directivos sobre la cantidad de recursos materiales y profesionales de sus escuelas.¹⁰ A partir de estos cuestionarios se recogen datos muy relevantes, que sirvieron de base para muchos de los análisis que se presentan en esta publicación.

1.4 Los procesos involucrados en las aplicaciones PISA

Para las tres áreas evaluadas en PISA —Matemática, Lectura y Ciencias— existe un marco de referencia, en el que se describen las “líneas maestras de la evaluación, definidas en función de los contenidos que deben aprender los alumnos, los procesos que han de ser capaces de ejecutar y los contextos en los que

deben aplicar sus conocimientos y habilidades”¹¹, para desempeñarse en la sociedad del futuro. Acorde con este marco, se diseñan pruebas mediante las cuales los estudiantes pueden demostrar en qué medida son capaces de resolver distintos tipos de problemas.

1.4.1 Elaboración de las preguntas

En el diseño de las preguntas, participan los comités de expertos de cada área y se reciben aportes de los países, a través de los coordinadores nacionales.

Cada pregunta elaborada es evaluada y visada en primer lugar por los comités de expertos del consorcio. Luego, los expertos de los países participantes envían su opinión. Posteriormente, las preguntas se aplican en pequeños estudios pilotos en algunos países, pero se concluye su validación en la aplicación experimental que se realiza en todos los países participantes, cuyo objetivo es probar una gran cantidad de preguntas para evaluar su comportamiento estadístico y psicométrico, así como las traducciones a los distintos idiomas. Sobre la base de los resultados de esta aplicación experimental, se seleccionan las preguntas que mejor responden a las exigencias psicométricas y disciplinarias.

Respecto de la traducción, cabe señalar que los instrumentos utilizados en cada país, deben tener idéntica forma y contener información que sea equivalente a las versiones originales que se entregan en inglés y francés¹². Cada país traduce y adapta los instrumentos con el apoyo de traductores y expertos disciplinarios en cada centro nacional. Esto permite asegurar que la traducción no afecte la dificultad de las preguntas ni altere su corrección disciplinaria. Posteriormente, el consorcio verifica la traducción de todos los países, utilizando árbitros expertos en todos los idiomas.

¹⁰ Los cuestionarios aplicados en PISA 2006, se encuentran disponibles en <http://www.simce.cl/index.php?id=440>

¹¹ OCDE (2006). PISA 2006. Marco de la evaluación. Conocimientos y habilidades en Ciencias, Matemática y Lectura. París, Francia.

¹² Estos son los dos idiomas oficiales de la OCDE y los instrumentos se entregan en versiones para ellos.

En la prueba definitiva de PISA 2006, se aplicó un total de 184 preguntas. De ellas 28 fueron de Lectura, 48 de Matemática y 108 de Ciencias¹³. El conjunto de preguntas se distribuye en formas o cuadernillos distintos (trece, en PISA 2006), cuya composición es comparable (formas complementarias). Cada estudiante contesta una sola forma de prueba, lo que asegura que trabaje un tiempo razonable en las distintas preguntas. En la conformación de los cuadernillos, se resguarda que cada pregunta tenga un número suficiente de respuestas, como para obtener un dato estadísticamente válido. Este sistema de formas complementarias hace posible evaluar una mayor cantidad de contenidos.

Como se verá en los estudios incluidos en la segunda sección de esta publicación, las preguntas de PISA son muy diversas entre sí y se organizan en unidades, cada una de las cuales está formada por un “estímulo” que plantea una situación a partir de la cual se desprenden las preguntas. El estímulo puede ser un texto, un diagrama, una imagen, etc.

De acuerdo con su formato¹⁴, las preguntas se pueden clasificar en:

- Selección múltiple simple. Los estudiantes deben seleccionar la correcta, de un total de cuatro alternativas.
- Selección múltiple compleja. Los estudiantes deben seleccionar varias alternativas para dar respuesta a distintas parte de una misma pregunta.
- Respuesta abierta breve. Los estudiantes deben desarrollar una respuesta, generalmente simple y que requiere poca elaboración, en un contexto en que existen varias posibilidades de respuestas correctas.

- Respuesta abierta compleja. Los estudiantes deben desarrollar una respuesta que requiere bastante elaboración y puede implicar una explicación, una argumentación o la descripción del proceso para llegar a la respuesta correcta. En este tipo de preguntas generalmente también se distinguen respuestas parcialmente correctas.

Las dos categorías de preguntas de selección múltiple corresponden a preguntas cerradas, contrastando con las dos últimas que son de desarrollo.

1.4.2 Selección de la muestra

Cada aplicación PISA se realiza sobre un muestra representativa nacional de los estudiantes de 15 años, que se selecciona en dos etapas. En un primer momento, cada país envía al consorcio una base de datos con su marco muestral, es decir, con el total de establecimientos educacionales con estudiantes “elegibles” (esto es, que tendrán 15 años y cursarán grados entre 7° y 12°, al momento de la aplicación). De este universo de establecimientos, el consorcio selecciona al azar, un subconjunto representativo, en función de las variables de caracterización identificadas por cada país¹⁵. Cada establecimiento seleccionado cuenta con dos alternativas de reemplazo.

A los establecimientos que aceptan participar, se les solicita la lista completa de sus estudiantes “elegibles”. Entre ellos, se escoge al azar un grupo de al menos 35, para que respondan la prueba. En caso de que el establecimiento cuente con menos de 35 estudiantes de esa edad, se solicita la participación de todos ellos. Los estándares PISA establecen que debe rendir la prueba al menos 80% de los estudiantes de la muestra.

13 En el momento de la calibración, post aplicación, cinco preguntas de Ciencias, del total aplicado, se eliminaron por presentar problemas métricos en una gran cantidad de países. Por lo tanto, el total considerado para la construcción de las escalas de puntajes de Ciencias fue de 103 preguntas.

14 Véanse ejemplos de preguntas en el Anexo, figuras A.1 y A.2.

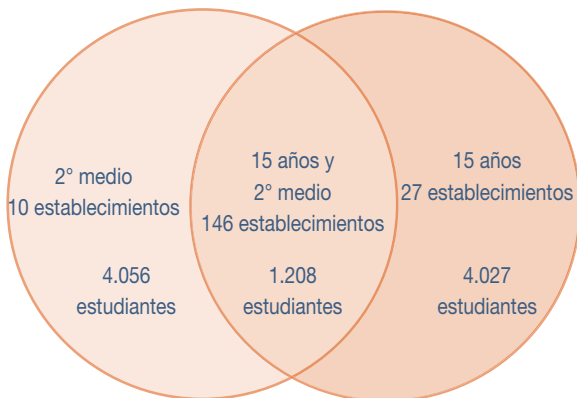
15 En Chile, las variables para caracterizar a los establecimientos en 2006, fueron: tamaño del establecimiento (distinguiendo tres categorías: con 1 a 17 estudiantes de 15 años; con 18 a 34 estudiantes y con 35 o más estudiantes); dependencia administrativa (municipal, particular subvencionado y particular pagado); niveles que imparte (sólo Enseñanza Básica, sólo Enseñanza Media y Enseñanza Básica y Media), y modalidad de enseñanza (técnico profesional y/o científico humanista).

Como se observa en la Figura 2, la muestra PISA 2006 de Chile consideró un total de 173¹⁶ escuelas y de 5.235 estudiantes de 15 años, que habían nacido entre el 1 de mayo de 1990 y el 30 de abril de 1991. En cada establecimiento se seleccionó un grupo de aproximadamente 40 estudiantes o a todos, cuando había menos estudiantes elegibles.

Adicionalmente, Chile incluyó la opción internacional de una muestra de grado, seleccionándose un curso completo de estudiantes de 2° Año Medio, en cada uno de los establecimientos participantes, con excepción de aquellos que no tenían Educación Secundaria. La muestra de 2° Año Medio quedó constituida por 156 establecimientos¹⁷ y un total de 5.264 estudiantes, 23% de los cuales (1.208 estudiantes), también formaban parte de la muestra de 15 años. El resto, 77% (4.056 alumnos) eran estudiantes de 2° Año Medio que no fueron seleccionados para la muestra de 15 años o que no tenían esa edad.

Figura 2.

Establecimientos y estudiantes chilenos de la muestra nacional PISA 2006 y de la muestra de 2° Año Medio



16 En la práctica se aplicó a 183 establecimientos, pero diez de ellos no cumplieron con las asistencias requeridas para representar al establecimiento en la muestra de 15 años y, por lo tanto, fueron excluidos.

17 En esta muestra de 2° medio permanecieron los 10 establecimientos excluidos de la muestra de 15 años, ya que sí contaban con la asistencia requerida en el curso seleccionado.

Para la muestra de estudiantes de 2° medio fue posible comparar su rendimiento relativo en las pruebas PISA y SIMCE 2006. Estos análisis consideraron solo a los estudiantes que efectivamente participaron en estas dos pruebas y que tenían puntajes en ambas, lo que correspondió a un total de 4.772 alumnos.

1.4.3 Corrección de preguntas de respuesta abierta

Las preguntas de respuesta abierta requieren una corrección que no puede ser mecanizada. Para corregirlas se cuenta con pautas elaboradas por los comités de expertos del consorcio, en las que se establecen detallados criterios de corrección, equivalentes para todos los países. El proceso de corrección se realiza en cada país, con una metodología compartida y con correctores entrenados y supervisados, de acuerdo con las indicaciones entregadas en los manuales correspondientes. La consistencia¹⁸ en la corrección de estas preguntas en cada país, es validada internacionalmente.

1.4.4 Aplicación de los instrumentos

La aplicación de PISA es completamente estandarizada, todos sus procesos están regidos por manuales de uso internacional y sujetos a distintos mecanismos de control de calidad. Las condiciones de aplicación deben ser idénticas para no poner en riesgo la comparabilidad de los resultados. Asimismo, se trabaja rigurosamente en la elaboración e implementación de protocolos sobre el manejo de la información y la administración de las pruebas.

La prueba se aplica en dos sesiones continuas, con una duración total de dos horas. Posteriormente, los estudiantes completan el cuestionario. Las sesiones son dirigidas por examinadores especialmente capacitados

18 La consistencia se refiere al grado en que dos correctores independientes, asignan un mismo puntaje a una misma respuesta.

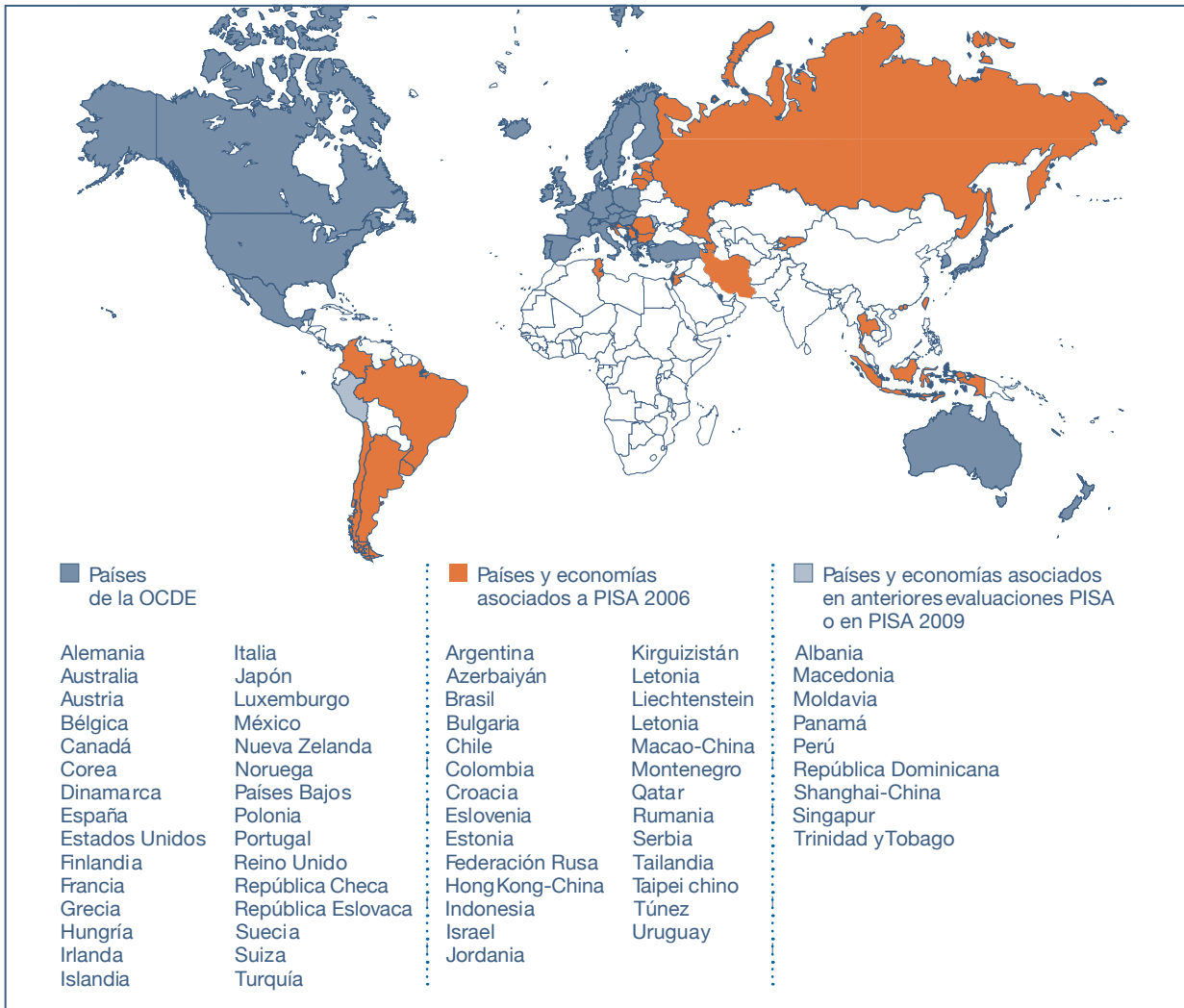
y supervisados. Los materiales para responder, la disposición de las salas, las instrucciones y los tiempos de descanso, son iguales para todos los estudiantes de los países participantes.

Para la aplicación de las pruebas, se acuerda una fecha y hora con los establecimientos, dentro de un período que ha sido previamente acordado. La aplicación en Chile de PISA 2006, se realizó entre el 21 de agosto y el 7 de septiembre de ese año.

1.5 Países participantes en PISA 2006

En PISA 2006, participaron alrededor de 400 mil estudiantes de 57 países (véase la Figura 3), representando un universo de 20 millones de estudiantes de 15 años. La extensión geográfica que cubre el proyecto, permite reflejar lo que sucede en un tercio de la población estudiantil de 15 años en el mundo.

Figura 3.
Países participantes en PISA 2006



2. Resultados generales de Chile en PISA 2006

2.1 Ciencias

2.1.1 Aspectos evaluados en PISA 2006 - Ciencias

Como se señaló en la sección anterior, las tareas y preguntas que desarrolló PISA 2006 para la evaluación de las ciencias consideraron cuatro aspectos relacionados entre sí, a saber:

- *Los conocimientos y las características distintivas de las ciencias.*
- *Las competencias que los alumnos han de aplicar (por ejemplo, cuando desarrollan un proceso científico determinado).*
- *Los contextos en los que los alumnos se enfrentan a problemas científicos y aplican los conocimientos y habilidades necesarios (por ejemplo, la toma de decisiones sobre su vida privada o la comprensión de la situación mundial).*
- *La actitud y disposición de los alumnos hacia las ciencias*¹⁹.

Entre los primeros, se distinguen los conocimientos relativos a conceptos y teorías y los conocimientos sobre la ciencia.

Los **conocimientos relativos a conceptos y teorías** fundamentales de la ciencia, que los estudiantes debían utilizar para mostrar sus competencias científicas en PISA 2006 se clasificaron en:

- **Sistemas físicos.** Conocimientos y conceptos relativos a la estructura, propiedades y cambios químicos de la materia, movimientos y fuerzas, la energía y su transformación, y las interacciones de la energía y la materia.
- **Seres vivos.** Conocimientos sobre las células, los

sistemas, aparatos y estructura del cuerpo humano, y sobre distintas poblaciones, ecosistemas y la biósfera.

- **La Tierra y el espacio.** Conocimientos respecto de la estructura de los sistemas del planeta y su energía, así como de la historia de la Tierra y su lugar en el espacio.
- **Sistemas tecnológicos.** Conocimientos y conceptos relativos al papel de la tecnología científica, las relaciones entre ciencia y tecnología, y los conceptos y principios de la tecnología.

Los **conocimientos sobre la ciencia** se refieren a la naturaleza de esta, como actividad humana y sobre sus potencialidades y limitaciones, distinguiendo:

- **Investigación científica.** Conocimientos sobre el origen y propósito de las investigaciones, los experimentos y los tipos de datos, así como sobre aspectos relativos a las mediciones, sus procedimientos e instrumentos y las características de los resultados.
- **Explicaciones científicas.** Conocimientos relativos a los tipos de explicaciones (hipótesis, teorías, modelos, leyes), la formación de nuevas explicaciones, las reglas que deben cumplir las explicaciones científicas, los resultados y las nuevas interrogantes que surgen tras la producción de nuevo conocimiento.

Las **competencias científicas** evaluadas por PISA 2006 fueron:

- **Identificar problemas científicos.** Capacidad para identificar qué preguntas son posibles de responder por la ciencia, cuáles son los términos claves para buscar información científica y cuáles son los rasgos fundamentales de la investigación científica.
- **Explicar fenómenos científicos.** Capacidad para aplicar el conocimiento científico en situaciones específicas y la habilidad para identificar, describir e interpretar fenómenos y para predecir cambios.
- **Utilizar evidencia científica.** Capacidad para interpretar

19 OCDE, 2006, op. cit.

pruebas científicas y para elaborar y comunicar conclusiones a partir de esas evidencias. Habilidad para identificar supuestos y pruebas que subyacen a las conclusiones y reflexionar sobre las implicancias del conocimiento científico y tecnológico.

Los contextos en que los estudiantes debían demostrar sus competencias y aplicar sus conocimientos se referían a situaciones concretas. Dado que PISA pone el acento en medir conocimientos y habilidades “para la vida”, las preguntas enfrentaban a los estudiantes con situaciones científicas o tecnológicas que podían encontrar en distintos ámbitos de su vida cotidiana. Entre éstas: salud, recursos naturales, calidad del medio ambiente, riesgos y la relación entre ciencia y

tecnología. Estas situaciones fueron relacionadas con tres contextos mayores: personal (la persona misma, su familia, sus pares), social (la comunidad) y global (el mundo o el planeta). Por ejemplo, se plantearon preguntas sobre el tema de la salud humana, en el contexto personal (mi salud, mi nutrición), en el social (vacunaciones, control de enfermedades) y en el global (epidemias).

En la Tabla 1, se muestran las competencias y conocimientos medidos en PISA 2006 en el área de Ciencias. En cada celda se indica el número de preguntas clasificadas de acuerdo a la competencia y el conocimiento medido²⁰.

Tabla 1.
Número de preguntas, según competencias y conocimientos científicos medidos en PISA 2006²¹

Competencias	Conocimiento de conceptos científicos				Conocimiento sobre la ciencia	
	Sistemas físicos	Seres Vivos	La tierra y el espacio	Sistemas tecnológicos	Investigación científica	Explicaciones científicas
Identificar problemas científicos	-	-	-	-	23	-
Explicar fenómenos científicos	15	21	11	2	-	-
Utilizar evidencia científica	2	1	-	6	1	21

Fuente: Elaboración propia a partir de Matriz de Clasificación de los Ítems OCDE, 2009. PISA 2006 Technical Report, página 28 y Listado de los Ítems puntuados (Documentación interna del proyecto).

20 Cada pregunta de la prueba corresponde a una competencia y a un solo tipo de conocimiento.

21 Se incluyen solo los 103 preguntas que fueron utilizados para calcular los puntajes a nivel internacional.

La prueba estructurada de esta manera permitió elaborar una Escala General de Ciencias, que considera la totalidad de las preguntas puntuadas. Asimismo, con los subconjuntos de preguntas fue posible elaborar una serie de subescalas, para aspectos específicos de la alfabetización científica. Estas subescalas son²²:

- Competencia Identificar problemas científicos.
- Competencia Explicar fenómenos científicos.
- Competencia Utilizar evidencia científica.
- Conocimiento científico. Sistemas físicos.
- Conocimiento científico. Seres vivos.
- Conocimiento científico. La Tierra y el espacio.
- Conocimiento sobre la ciencia.

2.1.2 Resultados en la Escala General y subescalas de Ciencias

Los resultados de los estudiantes se resumen en la Escala General de Ciencias, que para el conjunto de países de la OCDE, se fijó en un promedio de 500 puntos.

Entre los 57 países participantes, hubo 39 con un promedio significativamente superior al de Chile, 15 con promedios significativamente más bajos y dos con puntajes similares (véase el Anexo, Tabla A.1). Sin embargo, en el contexto regional, el desempeño de los jóvenes chilenos fue relativamente mejor, como se observa en la Tabla 2.

Tabla 2.
Puntajes en Escala General de Ciencias y subescalas de competencias científicas.
Chile, Latinoamérica y la OCDE.

Países	Escala General de Ciencias	Identificar problemas científicos	Explicar fenómenos científicos	Utilizar evidencia científica
Chile	438	444	432	440
Uruguay	428 ▼	429 ▼	423 •	429 •
México	410 ▼	421 ▼	406 ▼	402 ▼
Argentina	391 ▼	395 ▼	386 ▼	385 ▼
Brasil	390 ▼	398 ▼	390 ▼	378 ▼
Colombia	388 ▼	402 ▼	379 ▼	383 ▼
Promedio Latinoamericano	408 ▼	415 ▼	403 ▼	403 ▼
OCDE	500 ▲	499 ▲	500 ▲	499 ▲

▼ : Tiene un puntaje significativamente inferior a Chile.

▲ : Tiene un puntaje significativamente superior a Chile.

• : Tiene un puntaje que no es significativamente distinto de Chile.

Fuente: UCE (Unidad de Currículum y Evaluación) (2007). PISA 2006: Rendimientos de estudiantes de 15 años en Ciencias, Lectura y Matemática. Extraído desde <http://www.simce.cl/index.php?id=440>

22 No se desarrolló una subescala de conocimiento científico relativa a "sistemas tecnológicos", porque el número de preguntas aplicadas fue insuficiente. En relación con el "conocimiento sobre las ciencias", se elaboró una sola subescala.

2.1.3 Niveles de desempeño en Ciencias

Los niveles de desempeño PISA, se definen con el propósito de describir las competencias científicas que demuestran los estudiantes en distintos niveles de la escala. Los puntajes de PISA en Ciencias se agrupan en seis niveles de desempeño, donde el nivel 6 representa los puntajes más altos y el 1, los más bajos. De acuerdo a su puntaje, los estudiantes son clasificados en un determinado nivel de desempeño y de esta manera se puede obtener el porcentaje de estudiantes que, es capaz de realizar las tareas que se describen en cada caso. Los niveles de desempeño se consideran inclusivos, en términos de que los

estudiantes que alcanzan un nivel determinado, son capaces de realizar todas o la mayoría de las tareas que corresponden a niveles inferiores.

Sobre la base de los resultados de PISA 2006, se puede afirmar que el porcentaje de estudiantes chilenos clasificado en los niveles inferiores (nivel 2 o menos) fue más bajo que el de los demás países latinoamericanos, y el porcentaje de estudiantes clasificado en niveles superiores (4, 5 y 6) fue más alto. Sin embargo, en comparación con el promedio del grupo OCDE, la situación fue inversa: muchos más estudiantes fueron clasificados en los niveles inferiores, y muchos menos en los superiores (véase la Tabla 3).

Tabla 3.
Niveles de desempeño en la Escala General de Ciencias
y porcentaje de estudiantes chilenos y promedio OCDE, en cada nivel

Niveles	Límite Inferior en escala	¿Qué son capaces de hacer los estudiantes?	% en Chile	% en OCDE
6	707,9	Los estudiantes pueden, de manera consistente, identificar, explicar y aplicar conocimientos científicos y conocimientos sobre la ciencia en una variedad de situaciones complejas de la vida. Son capaces de justificar sus decisiones utilizando evidencia proveniente de diversas fuentes de información y de explicaciones. Demuestran, de manera clara y consistente, un pensamiento y razonamiento científico avanzado y la capacidad de usar su comprensión para respaldar la búsqueda de soluciones a situaciones científicas y tecnológicas poco habituales. Pueden usar conocimiento científico y argumentar para respaldar recomendaciones y decisiones sobre situaciones personales, sociales, o globales.	0,1	1,3
5	633,3	Los estudiantes pueden identificar los componentes científicos de muchas situaciones complejas de la vida y aplicar conceptos científicos como también conocimiento sobre la ciencia a estas situaciones, y comparar, seleccionar y evaluar evidencia científica apropiada para responder a situaciones de vida. Además, poseen habilidades de indagación bien desarrolladas, establecen adecuadamente relaciones entre conocimientos y aportan su comprensión lúcida y relevante a diversas situaciones. Pueden elaborar explicaciones fundadas en evidencia y desarrollar argumentos basados en su análisis crítico.	1,8	7,7

(Continuación en página siguiente)

(Continuación Tabla 3)

4	558,7	Los estudiantes pueden enfrentar exitosamente situaciones y problemas que puedan involucrar fenómenos explícitos y que les exigen hacer inferencias acerca del rol de la ciencia o la tecnología. Pueden seleccionar e integrar explicaciones de diferentes disciplinas científicas o tecnológicas y relacionarlas directamente con aspectos de la vida. También, reflexionar sobre sus acciones y comunicar decisiones usando conocimiento y evidencia científica.	8,4	20,3
3	484,1	Los estudiantes pueden identificar problemas científicos claramente descritos en una variedad de contextos. Pueden seleccionar hechos y conocimientos para explicar fenómenos y aplicar modelos simples o estrategias de investigación. Pueden interpretar y usar conceptos científicos de diferentes disciplinas y aplicarlos directamente. Pueden desarrollar argumentos breves a partir de hechos y tomar decisiones basadas en conocimiento científico.	20,1	27,4
2	409,5	Los estudiantes poseen el conocimiento científico adecuado para dar explicaciones posibles en contextos habituales o para establecer conclusiones basadas en investigaciones simples. Son capaces de realizar razonamiento directo y de hacer interpretaciones literales de los resultados de una investigación científica o de la resolución de un problema tecnológico.	29,9	24,0
1	334,9	Los estudiantes tienen un conocimiento científico limitado que sólo pueden aplicar a pocas situaciones que les resulten muy habituales. Pueden presentar explicaciones científicas que son obvias y que se desprenden explícitamente de la evidencia dada.	26,7	14,1
Bajo 1			13,1	5,2

Fuente: Descripciones: UCE 2007, PISA 2006: Rendimientos de estudiantes de 15 años en Ciencias, Lectura y Matemática. Extraída de http://www.simce.cl/fileadmin/Documentos_y_archivos_SIMCE/PISA2006/PISA_2006.pdf, Pag. 25.

Porcentajes: OCDE, 2008, *op. cit.*

2.2 Lectura

2.2.1 Aspectos evaluados en PISA 2006 - Lectura

La alfabetización en Lectura es definida como “la capacidad de comprender, utilizar y analizar textos escritos para alcanzar los objetivos del que lee, desarrollar sus conocimientos y posibilidades, y participar en la sociedad”²³. De esta manera, la alfabetización en Lectura en PISA, implica leer, interpretar y reflexionar, así como la habilidad de usar

la lectura para alcanzar los objetivos del lector en la vida real. El foco de PISA está en leer para aprender más que en aprender a leer.

La forma de evaluar la competencia lectora en PISA incluye tres dimensiones: el formato del material escrito, el tipo de tareas de lectura y el contexto, que se refiere al uso para el cual el texto fue construido. El formato del texto se refiere a si el estímulo presenta textos continuos (distintos tipos de prosa) o discontinuos (gráficos, formularios, listas). El tipo de tarea o los procesos involucrados en la lectura, que distingue PISA son: extraer información, interpretar textos, reflexionar y evaluar los textos. El contexto de la lectura, es decir,

²³ OCDE, 2008, *op. cit.*, p.294.

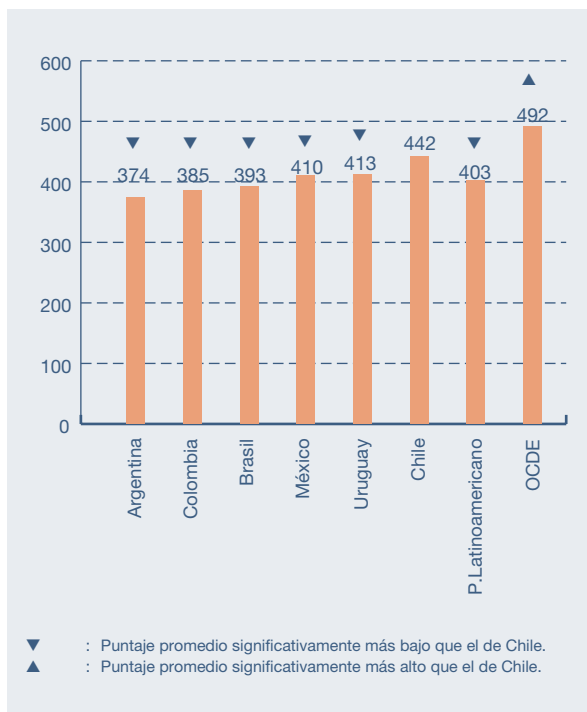
el uso que se le da puede ser privado, público, laboral o educacional.

2.2.2 Resultados en la Escala de Lectura²⁴

Entre los 56 países que aplicaron la prueba de Lectura de PISA 2006²⁵, tres obtuvieron un promedio semejante al de Chile (Turquía, Federación Rusa e Israel), 36 países obtuvieron promedios significativamente superiores al de Chile, y 16, significativamente más bajos.

Al igual que en Ciencias, en Lectura la situación de Chile en relación a Latinoamérica es más positiva que en relación al total de países participantes, como se observa en el Gráfico 1.

Gráfico 1: Puntaje promedio en la Escala de Lectura. Chile, Latinoamérica y la OCDE



Fuente: UCE, 2007, *op.cit.*

²⁴ Véase detalle en el Anexo, Tabla A.2.

²⁵ Por un problema técnico, Estados Unidos no tiene puntaje en Lectura, por tanto, en esta área se puede comparar el puntaje de solo 56 países.

2.2.3 Niveles de desempeño para Lectura

Los puntajes de la prueba PISA de Lectura, se agrupan en cinco niveles de desempeño, donde el nivel 5 representa los puntajes más altos y el 1 los más bajos. Los estudiantes que obtienen menos de 334.8 puntos en la escala de Lectura se clasifican bajo el nivel 1.

En la prueba PISA 2006, el nivel que agrupó a la mayoría de los estudiantes chilenos fue el nivel 2 (28%) y 63.6% logró ese nivel o uno superior. Esto indica que la mayoría de los alumnos, son al menos capaces de localizar uno o más fragmentos con información en un texto, aun cuando haya otra información que compita o si la información relevante no está suficientemente destacada. Pueden identificar la idea general de un texto, comprendiendo las relaciones entre las distintas partes de este o haciendo inferencias simples, y son capaces de evaluar la forma y el contenido de un texto, de acuerdo con su conocimiento previo y sus experiencias personales. Sin embargo, se debe notar que más de un tercio de los estudiantes chilenos tiene una comprensión lectora muy rudimentaria como la descrita en el nivel 1 o incluso inferior a ella (ver Tabla 4). Esto contrasta con los estudiantes de la OCDE donde un quinto de ellos alcanza solo este nivel o bajo éste.

Tabla 4.
Niveles de desempeño en la Escala de Lectura
y porcentaje de estudiantes chilenos y promedio OCDE, en cada nivel

Niveles	Límite Inferior en escala	¿Qué son capaces de hacer los estudiantes?	% en Chile	% en OCDE
5	625,6	Los estudiantes pueden ubicar y posiblemente ordenar en secuencias o combinar múltiples fragmentos de información que no se visualiza fácilmente en el texto, parte de la cual puede estar fuera del cuerpo principal del mismo. Inferir qué información en el texto es relevante para desarrollar una tarea. Pueden discriminar satisfactoriamente entre mucha información principal que compite entre sí y construir el significado de sutilezas del lenguaje o demostrar una comprensión completa y detallada de un texto. Son capaces de evaluar críticamente y formular hipótesis, basándose en conocimiento especializado, manejar conceptos contrarios a las expectativas y lograr una profunda comprensión de textos largos o complejos. Pueden analizar textos continuos, cuya estructura discursiva no es obvia o no está claramente delineada, con el objetivo de encontrar la relación entre partes específicas del texto con su tema o propósito implícitos. En textos discontinuos, los estudiantes pueden identificar patrones entre muchos fragmentos de información presentados de manera extensa y detallada, refiriéndose algunas veces a información externa. El lector puede necesitar darse cuenta que la total comprensión de un fragmento del texto requiere referirse a una parte separada del mismo documento, como un pie de página.	3,5	8,6
4	552,9	Son capaces de ubicar y posiblemente ordenar en secuencias o combinar múltiples fragmentos de información que no se visualiza fácilmente, cada uno de los cuales puede necesitar satisfacer múltiples criterios, en un texto de contenido o forma familiar. Inferir qué información en el texto es relevante para desarrollar una tarea. Usar una inferencia de alto nivel, basada en el texto, para comprender y aplicar categorías en un contexto no familiar, y para construir el significado de una sección del texto, considerando el texto como un todo. Manejar ambigüedades, ideas que son contrarias a lo esperado o formuladas negativamente. Usar conocimiento formal o general para formular hipótesis o evaluar críticamente un texto. Mostrar una comprensión precisa de textos complejos o largos. En textos continuos, establecer relaciones lingüísticas o temáticas entre varios párrafos, frecuentemente en ausencia de marcas discursivas claras, con el objetivo de localizar, interpretar o evaluar información no destacada, o para inferir sentido psicológico o metafísico. En textos discontinuos, explorar un texto largo y detallado con el propósito de encontrar información relevante, frecuentemente con poca o ninguna ayuda de organizadores como rótulos o formatos especiales, para localizar diversos fragmentos de información que deban ser comparados o combinados.	11,0	20,7

(Continuación en página siguiente)

(Continuación Tabla 4)

3	480,2	Los estudiantes pueden ubicar y en ocasiones reconocer la relación entre fragmentos de información, cada uno de los cuales puede requerir satisfacer múltiples criterios. Discriminar satisfactoriamente entre información relevante que compite entre sí. Integrar varias partes de un texto para identificar la idea principal, comprender una relación o construir el significado de una palabra u oración. Comparar, contrastar o categorizar tomando en cuenta diversos criterios. Son capaces de establecer conexiones o comparaciones, dar explicaciones o evaluar una característica de un texto; también, demostrar una comprensión detallada del texto en relación con conocimiento familiar, cotidiano o basándose en conocimiento menos común. En textos continuos, pueden usar convenciones sobre la organización de textos, cuando están presentes, y establecer relaciones lógicas –implícitas o explícitas– como relaciones causa-efecto entre oraciones o párrafos, para localizar, interpretar o evaluar información. En textos discontinuos, considerar una forma de presentación a la luz de una segunda, separar documentos o formas de presentación, posiblemente en un formato distinto, o combinar diversos fragmentos de información espacial, verbal y numérica en un gráfico o mapa para extraer conclusiones acerca de la información representada.	21,1	27,8
2	407,5	Los estudiantes son capaces de ubicar uno o más fragmentos de información, cada uno de los cuales puede requerir satisfacer múltiples criterios. Discriminar satisfactoriamente entre información que compite entre sí. Identificar la idea principal en un texto, comprender relaciones, establecer o aplicar categorías simples, o construir significado dentro de una parte limitada del texto, cuando la información no está destacada y se requieren inferencias simples. Hacer comparaciones o conexiones entre el texto y conocimiento extratextual, o explicar una característica del texto basándose en experiencias y actitudes personales. En textos continuos, establecer relaciones lógicas y lingüísticas dentro de un párrafo, para localizar o interpretar información; o sintetizar información entre textos, o de partes de un texto, para inferir el propósito del autor. En textos discontinuos, demostrar una noción de la estructura subyacente a una presentación visual, por ejemplo, un diagrama de árbol o un cuadro, o combinar dos fragmentos de información de un gráfico o un cuadro.	28,0	22,7
1	334,8	Pueden ubicar uno o más fragmentos independientes de información explícita que comúnmente requieren satisfacer un solo criterio, con poca o ninguna información que compite entre sí en un texto. Reconocer el tema principal o el propósito del autor en un texto sobre un tema que le resulta familiar, cuando la información en el texto está destacada. Establecer una relación simple entre la información del texto y el conocimiento de la vida cotidiana. En textos continuos, pueden usar redundancia, encabezados de párrafos o convenciones comunes de formato, para formarse una impresión de la idea principal del texto, o para localizar información explícita dentro de una sección breve del mismo. En textos discontinuos, son capaces de enfocar fragmentos discretos de información, usualmente en una sola forma de presentación, por ejemplo, un mapa simple, un gráfico de líneas o uno de barras con sólo una pequeña cantidad de información de manera directa, y en el que la mayor parte del texto verbal está limitado a una pequeña cantidad de palabras o frases.	21,5	12,7
Bajo 1			14,8	7,4

Fuente: Descripciones: UCE 2007, PISA 2006: Rendimientos de estudiantes de 15 años en Ciencias, Lectura y Matemática. Extraída de: http://www.simce.cl/fileadmin/Documentos_y_archivos_SIMCE/PISA2006/PISA_2006.pdf, Páginas 33-34. Porcentajes: Porcentajes: OCDE, 2008, *op. cit.*

2.2.4 Comparación de los puntajes de Lectura 2000-2006²⁶

Los estudiantes chilenos que participaron en PISA 2006 obtuvieron un promedio en Lectura significativamente más alto que el obtenido por aquellos que rindieron PISA 2000²⁷. Asimismo, en comparación con el conjunto de 36 países que participaron en ambas mediciones, Chile fue el que más aumentó su puntaje (33 puntos).

Como se observa en la Tabla 5, solamente siete países, cinco de los cuales no pertenecen a la OCDE, subieron su promedio en Lectura, entre ellos Chile. Por otra parte, 15 mantuvieron su promedio sin variación —13 de los cuales pertenecen a la OCDE y dos no (Brasil e Israel)— y 14 bajaron su promedio —nueve de la OCDE y 5 que no pertenecen.

Tabla 5.
Comparación puntajes en escala de Lectura,
PISA 2006-PISA 2000.

Situación	Frecuencia	Países
Aumentaron	7	Corea, Polonia, Chile, Hong Kong (China), Indonesia, Letonia, Liechtestein
Se mantuvieron	15	Austria, Bélgica, Canadá, República Checa, Dinamarca, Finlandia, Alemania, Hungría, Irlanda, Nueva Zelanda, Portugal, Suecia, Suiza, Brasil, Israel
Bajaron	14	Australia, Francia, Grecia, Islandia, Italia, Japón, México, Noruega, España, Argentina, Bulgaria, Rumania, Federación Rusa, Tailandia

Fuente: UCE, 2007, *op.cit.*

²⁶ Aunque la medición en Chile se realizó el año 2001, los datos corresponden al primer ciclo PISA, realizado en 2000 y son perfectamente comparables con los de los países que aplicaron ese año.

²⁷ Como ya se señaló, esta es la única área con resultados comparables entre ciclos para Chile.

2.3 Matemática

2.3.1 Aspectos evaluados en PISA 2006 - Matemática

PISA define la alfabetización Matemática como la “capacidad del individuo para identificar y entender la función que desempeñan las matemáticas en el mundo, emitir juicios fundados y utilizar y relacionarse con las matemáticas de forma que se puedan satisfacer las necesidades de la vida de los individuos como ciudadanos constructivos, comprometidos y reflexivos”²⁸.

La alfabetización matemática se relaciona con un uso más amplio y funcional de la matemática, que incluye la habilidad de reconocer y formular problemas matemáticos en variados contextos.

La alfabetización Matemática en PISA fue medida desde tres dimensiones: el **contenido** al que se refieren los problemas y las preguntas planteadas (Cantidad, Espacio y forma, Cambio y Relaciones e Incerteza), los **procesos** que los estudiantes necesitan activar para conectar los fenómenos que observan con la matemática (Reproducción, Conexión y Reflexión), y las situaciones o **contextos** en que son propuestos los problemas utilizados en los estímulos (Personal, Educativo y ocupacional, Público y Científico).

2.3.2 Resultados en la Escala de Matemática²⁹

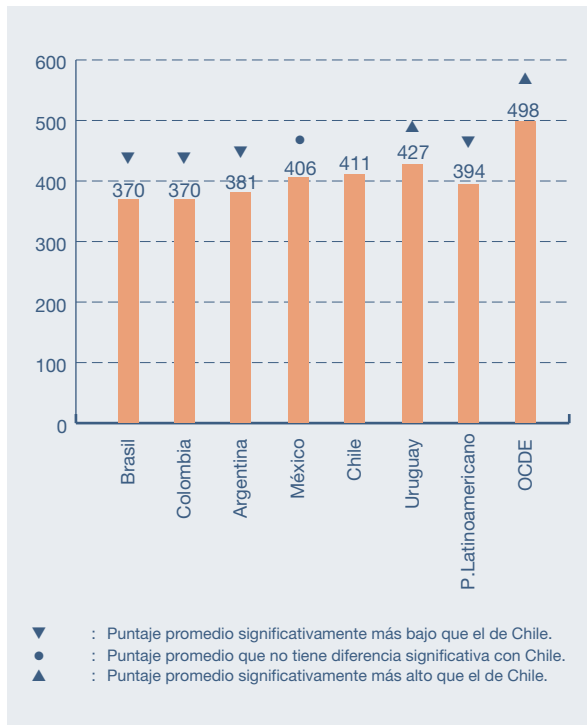
Entre los 57 países participantes en PISA 2006, 42 obtuvieron un promedio significativamente superior a Chile en la escala de Matemática, nueve estuvieron por debajo y cinco obtuvieron un puntaje similar, como se observa en el Gráfico 2. Claramente esta es el área en que los estudiantes chilenos obtienen

²⁸ OCDE, 2008, *op. cit.*, p. 314.

²⁹ Véase detalle en el Anexo, Tabla A.3.

los peores resultados en términos relativos. Hay más países con promedios significativamente superiores a los de Chile y la distancia con el promedio de la OCDE es de 0.89 desviaciones estándar, mientras en el caso de Ciencias es 0.60 y en Lectura, 0.50. Esta situación es similar en la comparación con los países latinoamericanos: Uruguay supera a Chile y México no se diferencia significativamente.

Gráfico 2: Puntaje en la Escala de Matemática. Chile, Latinoamérica y la OCDE



Fuente: UCE, 2007, op.cit.

2.3.3 Niveles de desempeño para Matemática

Los puntajes de PISA Matemática se clasifican en seis niveles de desempeño, donde el nivel 6 agrupa a los estudiantes con puntajes más altos y el 1, a aquellos con puntajes más bajos.

En PISA 2006, cerca de 55% de los estudiantes chilenos integran el nivel 1 y bajo 1, mientras solo un 21,3% de los estudiantes OCDE alcanza solo estos niveles. Este es un dato preocupante, porque significa que más de la mitad de los estudiantes no ha desarrollado competencias para enfrentar situaciones problemáticas de la vida que impliquen el uso de la Matemática. Su razonamiento matemático sólo se aplica a contextos muy familiares; podrían usar procedimientos rutinarios, siguiendo instrucciones, pero sin ser capaces de proponer modos alternativos de resolver los problemas. En la Tabla 6, se pueden observar las descripciones de los niveles de aprendizaje y la distribución de los estudiantes de Chile y de la OCDE en ellos.

Tabla 6.
Niveles de desempeño en la Escala de Matemática
y porcentaje de estudiantes chilenos y promedio OCDE, en cada nivel

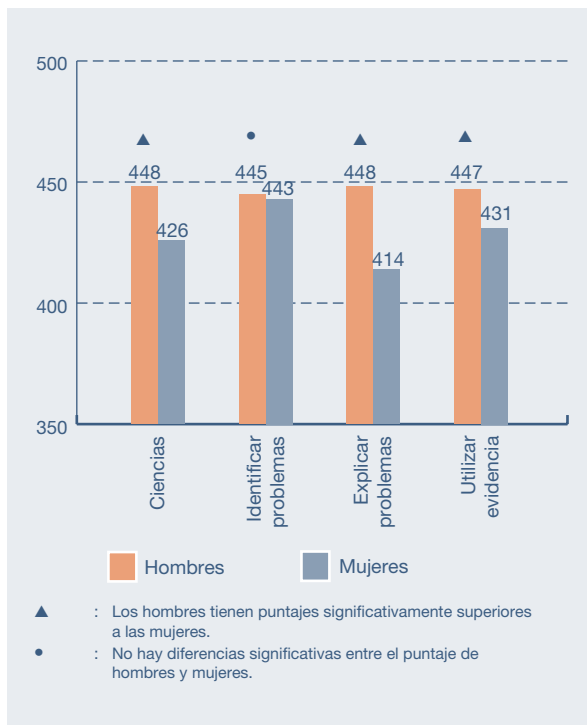
Niveles	Límite Inferior en escala	¿Qué son capaces de hacer los estudiantes?	% en Chile	% en OCDE
6	707,9	Los estudiantes conceptualizan, generalizan y usan información basada en sus investigaciones y en el modelamiento de situaciones problemáticas complejas. Pueden relacionar diferentes fuentes de información y representaciones y hacer traducciones entre ellas de manera flexible. Logran un razonamiento y pensamiento matemático avanzado, y pueden aplicarlo, junto con el dominio de las operaciones y relaciones matemáticas simbólicas y formales, en el desarrollo de nuevas aproximaciones y estrategias para enfrentar situaciones novedosas. Pueden formular y comunicar con precisión sus acciones y reflexiones relacionadas con sus descubrimientos, interpretaciones, argumentos, y la adecuación de éstas respecto de las situaciones originales.	0,1	3,3
5	633,3	Pueden desarrollar y trabajar con modelos para situaciones complejas, identificando sus limitaciones y especificando sus supuestos; seleccionar, comparar y evaluar estrategias de resolución de problemas adecuadas para abordar problemas complejos referidos a estos modelos. Pueden trabajar de manera estratégica usando habilidades de pensamiento y razonamiento amplias y bien desarrolladas, representaciones vinculadas adecuadamente, caracterizaciones simbólicas y formales, y comprensión profunda de estas situaciones. Son capaces de reflexionar sobre sus acciones y formular y comunicar sus interpretaciones y razonamiento.	1,3	10,0
4	558,7	Los estudiantes pueden trabajar de manera eficiente con modelos explícitos de situaciones concretas complejas que involucren condicionantes o la necesidad de reconocer supuestos. Pueden seleccionar e integrar diferentes representaciones, incluyendo las simbólicas, relacionándolas directamente con situaciones del mundo real. Son capaces de usar habilidades bien desarrolladas y razonar con flexibilidad y cierta profundización en estos contextos. Ellos pueden elaborar y comunicar explicaciones y razonamientos basándose en sus propias interpretaciones, argumentos y acciones.	5,6	19,1
3	484,1	Pueden ejecutar procedimientos claramente descritos, incluyendo los que requieren decisiones secuenciales. Ellos pueden seleccionar y aplicar estrategias simples de resolución de problemas, interpretar y usar representaciones basadas en diferentes fuentes de información, y razonar directamente a partir de ellas. Son capaces de elaborar comunicaciones breves para reportar sus interpretaciones, resultados y razonamiento.	13,9	24,3
2	409,5	Son aptos para interpretar y reconocer situaciones en contextos que requieren sólo inferencia directa. Ellos pueden extraer información relevante de una fuente de información a la vez y hacer uso de una manera de representación. Pueden utilizar algoritmos, fórmulas, procedimientos o convenciones básicas. Son capaces de razonar directamente y hacer interpretaciones literales de los resultados.	23,9	21,9
1	334,9	Los estudiantes pueden responder preguntas que involucren contextos familiares, en los cuales toda la información relevante está presente y las preguntas se encuentran claramente definidas. Son capaces de identificar información y llevar a cabo procedimientos rutinarios, siguiendo instrucciones directas en situaciones explícitas. Pueden realizar acciones que son obvias y se desprenden directamente de los estímulos presentados.	26,9	13,6
Bajo 1			28,2	7,7

Fuente: Descripciones: UCE 2007, PISA 2006: Rendimientos de estudiantes de 15 años en Ciencias, Lectura y Matemática. Extraída de: http://www.simce.cl/fileadmin/Documentos_y_archivos_SIMCE/PISA2006/PISA_2006.pdf. Página 40. Porcentajes: OCDE, 2008, *op. cit.*

2.4 Aspectos críticos de los resultados

Antes de terminar este capítulo introductorio, cabe destacar dos aspectos que inciden en la equidad de los logros de aprendizaje. Chile es uno de los países donde las mujeres tienen mayores brechas que las desfavorecen respecto de los hombres. Esto es especialmente serio en el caso de Ciencias, área en la que la mayoría de los países no muestran diferencias significativas por género. En el Gráfico 3 se muestran las diferencias por género de los estudiantes chilenos, en la Escala General de Ciencias y en las subescalas de competencias científicas.

Gráfico 3: Diferencias por género en la Escala General de Ciencias y subescalas de competencias científicas. Chile

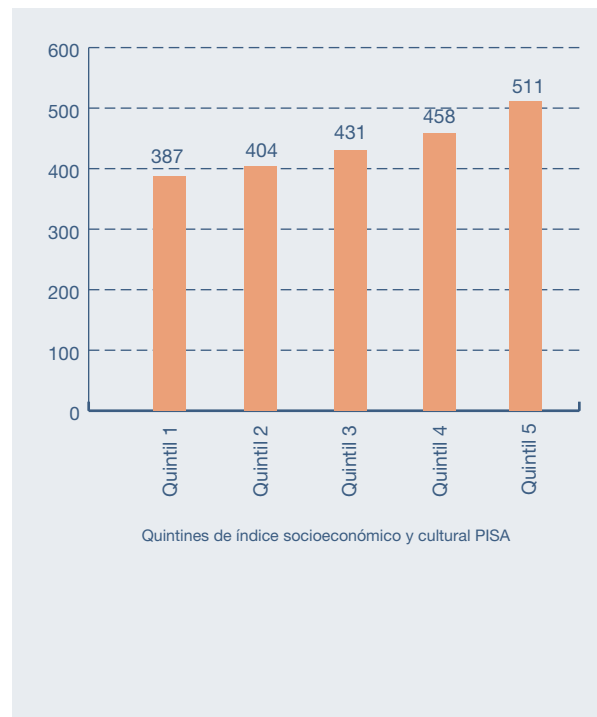


Fuente: UCE, 2007, *op.cit.*

Un segundo aspecto preocupante es la falta de equidad en los logros de aprendizaje considerando el nivel socioeconómico. En el Gráfico 4, se muestran los resultados promedio en Ciencias de los estudiantes chilenos, agrupados en quintiles, según el índice socioeconómico y cultural de PISA. Como se puede observar, la diferencia entre el primer y último quintil es de 124 puntos PISA, que equivalen a 1.24 desviaciones estándar.

Estos aspectos de inequidad merecen la atención de la política educacional y los artículos de esta publicación, los abordan desde distintas perspectivas.

Gráfico 4: Puntajes en la Escala General de Ciencias para quintiles del Índice Socioeconómico y Cultural PISA. Chile.

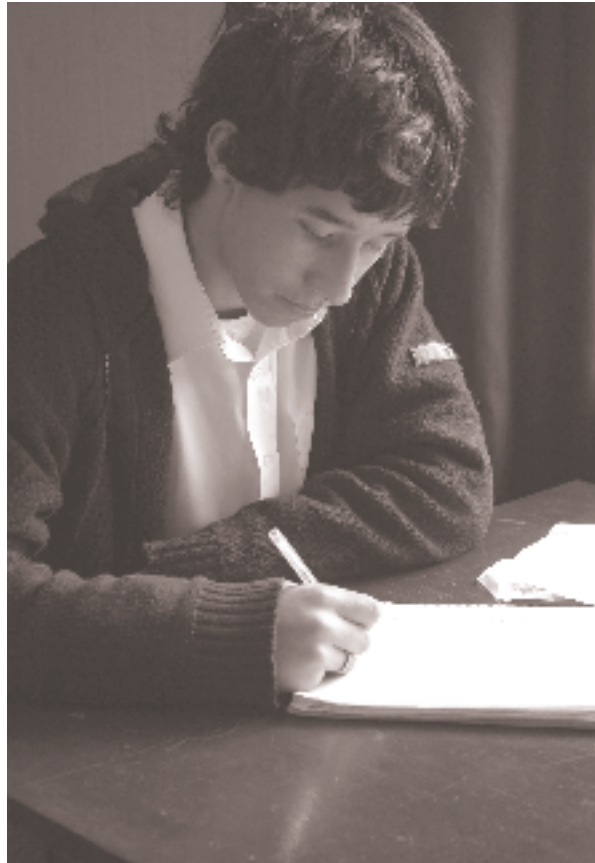


Fuente: UCE, 2007, *op.cit.*

3. Conclusiones

Chile se encuentra participando actualmente en PISA 2009, cuya aplicación definitiva coincidirá con la aparición de esta publicación. El foco de ese ciclo es Lectura y será interesante analizar si se confirma o no la tendencia al alza que se encontró en 2006. Este será el primer ciclo en que se podrán comparar los resultados de una misma área principal, en dos ciclos, con nueve años de por medio. Además del interés que reviste esta comparación en detalle, en esta oportunidad, al igual que en 2006, se podrán comparar los resultados PISA con los de la prueba SIMCE, ya que ambas se aplicaron en 2008 a los alumnos de 2° Año Medio de la muestra de establecimientos PISA. Posteriormente, Chile participará en PISA 2012 cuyo foco será Matemática. Entre tanto, se liberarán los resultados de la Prueba Internacional de Educación Cívica y Formación Ciudadana (ICCS 2009), que se aplicó en Chile a fines del 2008, y ya se ha iniciado por tercera vez en Chile, el Estudio de Tendencias en Ciencias y Matemática (TIMSS 2011). La aplicación experimental y definitiva de este último estudio, se realizará durante 2010.

Parece deseable que la experiencia del esfuerzo conjunto que representa esta publicación, permita realizar análisis semejantes en relación a los estudios internacionales en el futuro. En cualquier caso, el SIMCE siempre pondrá a disposición de los investigadores las bases de datos de los estudios que realice.



Anexo

Ejemplos de preguntas.

Figura A.1.
Pregunta de Ciencias, del tipo “selección múltiple compleja”

EL GRAN CAÑÓN

El Gran Cañón está ubicado en un desierto de los Estados Unidos. Es un cañón muy grande y profundo que contiene muchas capas de roca. En algún momento del pasado, los movimientos de la corteza terrestre levantaron estas capas. Hoy en día el Gran Cañón tiene 1.6 km de profundidad en algunas partes. Al fondo del cañón corre el Río Colorado.

Observa la siguiente fotografía del Gran Cañón tomada desde su orilla Sur. En las paredes del cañón pueden verse varias capas de roca distintas.



Piedra caliza A

Lutita A

Piedra caliza B

Lutita B

Esquistos y Granito

Estímulo

EL GRAN CAÑÓN

5429007

Cada año, cerca de cinco millones de personas visitan el parque nacional del Gran Cañón. El daño que tantos visitantes causan al parque es motivo de preocupación.

¿Es posible contestar las siguientes preguntas mediante investigaciones científicas? Encierra en un círculo "Sí" o "No" para cada pregunta.

¿Es posible responder estas preguntas mediante investigaciones científicas?	¿Sí o No?
¿Cuánta erosión es causada por caminar por los senderos?	Sí / No
¿Es hoy el parque tan hermoso como hace 100 años?	Sí / No

Pregunta selección múltiple compleja

Puntaje Completo
Código 1: Ambos correctos: Sí, No, en ese orden.


Sin Puntaje:
Código 0: Otras Respuestas

Figura A.2. Pregunta de Ciencias, del tipo “respuesta abierta”

LLUVIA ÁCIDA

A continuación se muestra una foto de las estatuas llamadas Canitides, que fueron construidas en la Acrópolis de Atenas hace más de 2.500 años. Las estatuas están hechas de un tipo de roca llamada mármol. El mármol está compuesto de carbonato de calcio.

En 1980, las estatuas originales que estaban siendo reconstruidas por la lluvia ácida, fueron trasladadas al interior del Museo de la Acrópolis y fueron reemplazadas por réplicas.



PUNTAJE PARA LLUVIA ÁCIDA 2

Puntaje completo

Código 2: Cualquiera de las siguientes: emisiones de los autos, de las industrias, combustión de combustibles fósiles como carbón y petróleo, gases de los volcanes u otras cosas similares.

- De quemar carbón y gas.
- Los gases del aire vienen de la contaminación producida por fábricas e industrias.
- Volcanes.
- Humo de plantas de producción de energía. (Plantas de producción de energía se llama como que incluye plantas de producción de energía que queman combustibles fósiles.)
- Vienen de la quema de materiales que contienen azufre y nitrógeno.

Puntaje parcial

Código 1: Respuestas que incluyen una incorrecta como también una correcta fuente de contaminación.

- Combustibles fósiles y plantas nucleares. (Las plantas nucleares no son una fuente de lluvia ácida.)
- Los óxidos vienen del ozono, la atmósfera y meteoritas que vienen hacia la Tierra. También de la quema de combustibles fósiles.

Respuestas que se refieren a la “contaminación” pero que no dan una fuente de contaminación que sea una causa significativa de lluvia ácida.

- Contaminación.
- El medio ambiente en general, la atmósfera en la que vivimos, por ejemplo, polución.
- La gasificación, polución, incendios, cigarrillos. (No es claro lo que se quiere decir con “gasificación”, “incendios” no es suficientemente específico, el humo del cigarrillo no es una fuente significativa de lluvia ácida.)
- Polución como la de plantas de energía nuclear.

Comentario: Solo mencionar “contaminación” es suficiente para código 1. Cualquier otro acompañante como ejemplo es evaluado para determinar si la respuesta, en cambio, merece un código 2.

Sin puntaje

Código 0: Otras respuestas, incluyendo respuestas que no mencionan “contaminación” y que no dan una causa significativa de lluvia ácida.

- Son emitidos por los plásticos.
- Son componentes naturales del aire.
- Cigarrillos.
- Carbón y petróleo. (No es suficientemente específico-no hay referencia a la “quema”.)
- Plantas de energía nuclear.
- Desechos industriales. (No es suficientemente específico.)

Código 9: Omittida.

Tabla A.1.
Puntajes en la Escala General de Ciencias

País	Puntaje	Error estándar	Diferencia con Chile
Finlandia	563	(2,0)	▲
Hong Kong, China	542	(2,5)	▲
Canadá	534	(2,0)	▲
Taipei, China	532	(3,6)	▲
Estonia	531	(2,5)	▲
Japón	531	(3,4)	▲
Nueva Zelanda	530	(2,7)	▲
Australia	527	(2,3)	▲
Holanda	525	(2,7)	▲
Liechtenstein	522	(4,1)	▲
Corea	522	(3,4)	▲
Eslovenia	519	(1,1)	▲
Alemania	516	(3,8)	▲
Reino Unido	515	(2,3)	▲
República Checa	513	(3,5)	▲
Suiza	512	(3,2)	▲
Macao, China	511	(1,1)	▲
Austria	511	(3,9)	▲
Bélgica	510	(2,5)	▲
Irlanda	508	(3,2)	▲
Hungría	504	(2,7)	▲
Suecia	503	(2,4)	▲
Polonia	498	(2,3)	▲
Dinamarca	496	(3,1)	▲
Francia	495	(3,4)	▲
Croacia	493	(2,4)	▲
Islandia	491	(1,6)	▲
Letonia	490	(3,0)	▲
Estados Unidos	489	(4,2)	▲

(Continuación en página siguiente)

(Continuación Tabla A.1)

Eslovaquia	488	(2,6)	▲
España	488	(2,6)	▲
Lituania	488	(2,8)	▲
Noruega	487	(3,1)	▲
Luxemburgo	486	(1,1)	▲
Federación Rusia	479	(3,7)	▲
Italia	475	(2,0)	▲
Portugal	474	(3,0)	▲
Grecia	473	(3,2)	▲
Israel	454	(3,7)	▲
Chile	438	(4,3)	●
Serbia	436	(3,0)	●
Bulgaria	434	(6,1)	●
Uruguay	428	(2,7)	▼
Turquía	424	(3,8)	▼
Jordania	422	(2,8)	▼
Tailandia	421	(2,1)	▼
Rumania	418	(4,2)	▼
Montenegro	412	(1,1)	▼
México	410	(2,7)	▼
Indonesia	393	(5,7)	▼
Argentina	391	(6,1)	▼
Brasil	390	(2,8)	▼
Colombia	388	(3,4)	▼
Túnez	386	(3,0)	▼
Azerbaiyán	382	(2,8)	▼
Qatar	349	(0,9)	▼
Kirguistán	322	(2,9)	▼

▲: Puntaje significativamente superior al de Chile

▼: Puntaje significativamente inferior al de Chile

●: Puntaje que no es significativamente distinto al de Chile.

Fuente: OCDE, 2008, *op. cit.*

Tabla A.2.
Puntajes en la Escala General de Lectura

País	Puntaje	Error estándar	Diferencia con Chile
Corea	556	(3,8)	▲
Finlandia	547	(2,1)	▲
Hong Kong, China	536	(2,4)	▲
Canadá	527	(2,4)	▲
Nueva Zelanda	521	(3,0)	▲
Irlanda	517	(3,5)	▲
Australia	513	(2,1)	▲
Liechtenstein	510	(3,9)	▲
Polonia	508	(2,8)	▲
Suecia	507	(3,4)	▲
Holanda	507	(2,9)	▲
Bélgica	501	(3,0)	▲
Estonia	501	(2,9)	▲
Suiza	499	(3,1)	▲
Japón	498	(3,6)	▲
Taipei, China	496	(3,4)	▲
Reino Unido	495	(2,3)	▲
Alemania	495	(4,4)	▲
Dinamarca	494	(3,2)	▲
Eslovenia	494	(1,0)	▲
Macao, China	492	(1,1)	▲
Austria	490	(4,1)	▲
Francia	488	(4,1)	▲
Islandia	484	(1,9)	▲
Noruega	484	(3,2)	▲
República Checa	483	(4,2)	▲
Hungría	482	(3,3)	▲
Letonia	479	(3,7)	▲

(Continuación en página siguiente)

(Continuación Tabla A.2)

Luxemburgo	479	(1,3)	▲
Croacia	477	(2,8)	▲
Portugal	472	(3,6)	▲
Lituania	470	(3,0)	▲
Italia	469	(2,4)	▲
Eslovaquia	466	(3,1)	▲
España	461	(2,2)	▲
Grecia	460	(4,0)	▲
Turquía	447	(4,2)	●
Chile	442	(5,0)	●
Federación Rusa	440	(4,3)	●
Israel	439	(4,6)	●
Tailandia	417	(2,6)	▼
Uruguay	413	(3,4)	▼
México	410	(3,1)	▼
Bulgaria	402	(6,9)	▼
Serbia	401	(3,5)	▼
Jordania	401	(3,3)	▼
Rumania	396	(4,7)	▼
Indonesia	393	(5,9)	▼
Brasil	393	(3,7)	▼
Montenegro	392	(1,2)	▼
Colombia	385	(5,1)	▼
Túnez	380	(4,0)	▼
Argentina	374	(7,2)	▼
Azerbaiyán	353	(3,1)	▼
Qatar	312	(1,2)	▼
Kirguistán	285	(3,5)	▼

▲: Puntaje significativamente superior al de Chile.

▼: Puntaje significativamente inferior al de Chile.

●: Puntaje que no es significativamente distinto al de Chile.

Fuente: OCDE, 2008, *op. cit.*

Tabla A.3.
Puntajes en la Escala General de Matemática

País	Puntaje	Error estándar	Diferencia con Chile
Taipei, China	549	(4,1)	▲
Finlandia	548	(2,3)	▲
Hong Kong, China	547	(2,7)	▲
Corea	547	(3,8)	▲
Holanda	531	(2,6)	▲
Suiza	530	(3,2)	▲
Canadá	527	(2,0)	▲
Macao, China	525	(1,3)	▲
Liechtenstein	525	(4,2)	▲
Japón	523	(3,3)	▲
Nueva Zelanda	522	(2,4)	▲
Bélgica	520	(3,0)	▲
Australia	520	(2,2)	▲
Estonia	515	(2,7)	▲
Dinamarca	513	(2,6)	▲
República Checa	510	(3,6)	▲
Islandia	506	(1,8)	▲
Austria	505	(3,7)	▲
Eslovenia	504	(1,0)	▲
Alemania	504	(3,9)	▲
Suecia	502	(2,4)	▲
Irlanda	501	(2,8)	▲
Francia	496	(3,2)	▲
Reino Unido	495	(2,1)	▲
Polonia	495	(2,4)	▲
Eslovaquia	492	(2,8)	▲
Hungría	491	(2,9)	▲
Luxemburgo	490	(1,1)	▲
Noruega	490	(2,6)	▲

(Continuación en página siguiente)

(Continuación Tabla A.3)

Lituania	486	(2,9)	▲
Letonia	486	(3,0)	▲
España	480	(2,3)	▲
Azerbaijón	476	(2,3)	▲
Federación Rusa	476	(3,9)	▲
Estados Unidos	474	(4,0)	▲
Croacia	467	(2,4)	▲
Portugal	466	(3,1)	▲
Italia	462	(2,3)	▲
Grecia	459	(3,0)	▲
Israel	442	(4,3)	▲
Serbia	435	(3,5)	▲
Uruguay	427	(2,6)	▲
Turquía	424	(4,9)	●
Tailandia	417	(2,3)	●
Rumania	415	(4,2)	●
Bulgaria	413	(6,1)	●
Chile	411	(4,6)	●
México	406	(2,9)	●
Montenegro	399	(1,4)	▼
Indonesia	391	(5,6)	▼
Jordania	384	(3,3)	▼
Argentina	381	(6,2)	▼
Colombia	370	(3,8)	▼
Brasil	370	(2,9)	▼
Túnez	365	(4,0)	▼
Qatar	318	(1,0)	▼
Kirguistán	311	(3,4)	▼

▲: Puntaje significativamente superior al de Chile.

▼: Puntaje significativamente inferior al de Chile.

●: Puntaje que no es significativamente distinto al de Chile.

Fuente: OCDE, 2008, *op. cit.*

Factores Asociados 2



- Evidencia sobre el uso de tecnologías y su correlación con el desempeño en Pisa-Ciencias 2006
- ¿Cómo las escuelas chilenas pueden mejorar el aprendizaje en Ciencias?
- ¿Qué explica las diferencias de resultados PISA Matemática entre Chile y algunos países de la OCDE y América Latina?
- ¿Qué explica las diferencias de resultados PISA Lectura entre Chile y algunos países de la OCDE y América Latina?

Evidencia sobre el uso de tecnologías y su correlación con el desempeño en Pisa-Ciencias 2006

Por
Martha Kluttig
Claudia Peirano
Constanza Vergara¹

Resumen

Este estudio permite avanzar en el conocimiento que existe sobre acceso y uso de las Tecnologías de Información y Comunicación (TIC) y su relación con el aprendizaje de las ciencias. La hipótesis de este estudio fue que el mero acceso a la tecnología no tiene impacto en los resultados de los estudiantes en Ciencias y que la frecuencia de uso, no siempre se relaciona positivamente con el rendimiento escolar: la relación está condicionada al tipo de actividad desarrollada en el computador.

Para ello, primero se describió el acceso y uso dado por los estudiantes chilenos al computador; se hicieron comparaciones con la OCDE y entre diferentes grupos del país, y se realizaron estudios econométricos para analizar la correlación condicionada entre el uso de las TIC y los resultados académicos. Todo lo anterior, sobre la base de la evidencia recopilada en las Pruebas PISA 2000 y 2006, sobre el desempeño de los estudiantes en Ciencias.

El principal hallazgo del análisis descriptivo fue que existe un limitado acceso a la tecnología y la conectividad, por parte de los estudiantes de Educación Media en Chile, en comparación con los países de la OCDE. Estas diferencias se acentúan en los grupos de alumnos de niveles socioeconómicos más bajos. En el análisis econométrico, por su parte, se encontró una correlación positiva entre frecuencia de uso de computadores y desempeño, la cual, sin embargo, oculta que distintos usos, se relacionan de manera diferente con el aprendizaje: ciertos usos están relacionados de forma positiva, otros de manera nula e incluso algunos se relacionan negativamente.

¹ Martha Kluttig, Claudia Peirano y Constanza Vergara son profesionales del Centro de Microdatos del Departamento de Economía de la Universidad de Chile.

Las autoras agradecen el financiamiento proporcionado por la Iniciativa Científica Milenio al Centro de Microdatos, Proyecto P07S-023-F; el apoyo proporcionado por Telefónica Chile a la agenda de investigación sobre tecnología y educación, y la participación de Edmundo Hermosilla, como asistente de esta investigación.

Introducción

En los últimos años, el mundo ha vivido procesos de transformación en la manera en la cual se organiza la economía. Muchos países han pasado de tener economías industriales, agrícolas y extractoras de materias primas, para avanzar hacia una situación donde el conocimiento y la información se perfilan como un nuevo y valioso tipo de capital (Law, Pelgrum y Plomo, Eds., 2008). Las tecnologías de información y conocimientos (TIC) están ejerciendo un rol central en el desarrollo de estas nuevas economías y en la participación social de los ciudadanos. Por ello, el mundo occidental ha realizado importantes inversiones destinadas a que los estudiantes puedan acceder a las TIC en los establecimientos educativos, confiando en el propósito de que la brecha tecnológica no sea un elemento más de desigualdad social. Las familias, por su parte, también realizan esfuerzos importantes para que sus hijos tengan acceso a equipos y conectividad. En el caso de Chile, el Ministerio de Educación ha impulsado proyectos de informática educativa hace más de 15 años y entre sus resultados, se registra el 2007, un promedio de 26 alumnos por computador, en los establecimientos subvencionados² (MINEDUC, 2008).

Pese a estos esfuerzos, la evidencia aún entrega pocos elementos para poder entender la relación que existe entre acceso y uso de las TIC, y el proceso de enseñanza-aprendizaje; ello, básicamente porque los resultados de las investigaciones, no son concluyentes. En efecto, existen múltiples investigaciones orientadas a conocer cuál ha sido la evolución de los países en materia de acceso y uso de las TIC para fines educativos y la manera en que los jóvenes y los profesores se involucran con la tecnología. Algunos de estos estudios establecen efectos nulos. Por ejemplo, Cuban (2003) señaló que la inversión en tecnología realizada por Estados Unidos en los establecimientos educacionales

no había generado ningún impacto en la educación. Otros estudios (como los de Beltran, Das y Fairlie, 2006; Wößmann y Fuchs, 2004, y Angrist y Lavy, 2002) han evidenciado un impacto negativo de las TIC en el desempeño. En ellos se plantea, por ejemplo, que el computador constituye una fuente de distracción, que sustituye métodos efectivos de enseñanza, facilita el acceso a fuentes no confiables y limita la creatividad de los alumnos, ya que los computadores vendrían programados para funcionar de una manera determinada. Finalmente, diversos estudios (como los de Machin, McNally y Silva, 2006; Papanastasiou, Zembylas y Vrasidas, 2003, y Sivin-Kachala y Bialo, 2000) han encontrado un efecto positivo y significativo del uso del computador sobre el desempeño; y hay autores que plantean incluso, que es posible que las nuevas generaciones de alumnos requieran un mayor uso de tecnologías en el proceso educativo, para alcanzar mejores resultados. Tal es el caso de Pedró (2006), quien define a los *New Millenium Learners* (NML) como una generación altamente integrada en lo comunicacional, que además suele hacerlo de manera múltiple: son jóvenes que ven televisión, mientras hacen sus tareas y chatean en internet. Según este autor, los alumnos que nacieron en la década de los ochenta, han vivido en un contexto donde la tecnología está constantemente presente. Este cambio en los sujetos de la educación plantea la necesidad de adaptar las prácticas pedagógicas a las nuevas estructuras de pensamiento. En este contexto, promover el uso de multimedia como recurso educacional y la innovación de temas relacionados con las TIC, puede ser un factor importante a la hora de educar a esta nueva generación. Junto con lo anterior, el uso de computadores puede tener otros efectos, tales como: disminuir la criminalidad juvenil y otras acciones antisociales, al “sacar” a los estudiantes de la calle (Beltran, Das y Fairlie, 2006); aumentar la motivación en los alumnos con dificultades de aprendizaje (Karpati, Ed., 2004), y desarrollar otro tipo de habilidades, que aunque no se midan en las pruebas de aprendizaje, son valoradas por el mercado laboral (Borghans y Ter Weel, 2004).

2 Establecimientos subvencionados son aquellos financiados total o parcialmente por el Estado.

Ahora bien, quienes defienden el uso de las TIC, tienden a coincidir en la complejidad del proceso. La obtención de retornos a la inversión en tecnología educativa requiere mucho más que la mera introducción de tecnología con software y recursos web alineados curricularmente. Requiere de la triangulación de contenidos, principios de calidad en la enseñanza y enseñanza de alta calidad. Todo lo anterior, alineado con la evaluación y la rendición de cuentas (Papanastasiou, Zembylas y Vrasidas, 2003; Wößmann y Fuchs, 2004, y Barrera-Osorio y Linden, 2009).

En el contexto de aportar con información a este respecto, que resulte relevante a la realidad chilena, en el presente estudio se plantea como objetivo, mostrar evidencia sobre el acceso y uso de computadores de los estudiantes chilenos y determinar las correlaciones que existen entre ese acceso y uso, y el desempeño de los mismos, en la Prueba de Ciencias PISA 2006 (OCDE, 2006). La hipótesis de este estudio es que el mero acceso a la tecnología no tiene impacto en los resultados y que la frecuencia de uso no siempre se relaciona positivamente con el rendimiento escolar: la relación está condicionada al tipo de actividad desarrollada en el computador.

A continuación, se presenta la metodología usada para ambas fases del análisis (descriptivo y econométrico) y luego se exponen los principales resultados de dichos análisis. Finalmente, se plantean las conclusiones del estudio, incorporando recomendaciones para el ámbito de las políticas públicas.

1. Metodología

En este estudio, con el objeto de presentar evidencia sobre el acceso y uso de los computadores, se realizó un análisis descriptivo del acceso y uso de tecnología. En una segunda parte y con el propósito de encontrar correlaciones condicionadas entre el uso del computador y el desempeño de los alumnos, se estimaron funciones de producción, entre desempeño

y tecnología, controlando por características de los alumnos y los establecimientos.

Para el análisis descriptivo se utilizaron los siguientes cuestionarios, aplicados para la prueba PISA, en 2000 y 2006 (OCDE, 2000 y 2006):

- Cuestionario de estudiantes, el cual incluye preguntas de contexto, de interés en Ciencias y que, en PISA 2006 incluyó además un cuestionario sobre el uso de tecnologías.
- Cuestionario de establecimientos.

Mediante el análisis econométrico, se estimó la correlación condicionada del acceso, frecuencia y tipo de uso de computadores, con el desempeño logrado en la Prueba PISA de Ciencias, por los estudiantes chilenos. Se modeló una ecuación multinivel para cada estimación.

La función de producción que se estimó para obtener correlaciones condicionadas entre desempeño en Ciencias y acceso y uso de tecnología, tiene la siguiente estructura:

$$P_{ij} = \beta_{0j} + \beta_1 F_{ij} + \beta_2 C_j + \beta_3 T_{ij} + \mu_{ij}$$

donde P es el puntaje obtenido por el alumno i , en el establecimiento educacional j , en la prueba PISA de Ciencias. F es el conjunto de variables relacionadas con el contexto familiar y personal del estudiante, C es el conjunto de variables relacionadas con el establecimiento donde asiste y T son las variables relacionadas con el uso de computadores. Los parámetros β representan el valor estimado de la correlación condicionada entre la variable indicada y el desempeño en Ciencias.

La especificación de la función de producción permite que exista un intercepto diferente para cada establecimiento (β_0) pero restringe a que la correlación entre las variables explicativas con el desempeño (β) sea constante para todos los establecimientos.

Se estimaron cinco regresiones:

- Ecuación básica: se incorporan sólo las variables de control, es decir las variables relacionadas con el contexto familiar y personal del estudiante (F) y las variables relacionadas con el establecimiento donde asiste (C). No se incluyen variables de tecnología, ya que se plantean como base de comparación.
- Acceso: se incluyen todas las variables de la ecuación básica más las variables de acceso en el hogar y el acceso en el establecimiento.
- Frecuencia casa: se consideran sólo los alumnos que tienen acceso a un computador en el hogar. Se incluyen todas las variables de la ecuación básica más la frecuencia de uso en el hogar y el acceso en el establecimiento.
- Frecuencia establecimiento: se incluyen todas las variables de la ecuación básica más la frecuencia de uso en el establecimiento educacional y el acceso en el hogar.
- Tipo de uso: se incluyen todas las variables de la ecuación básica más la frecuencia de los siguientes usos: internet, tecno-adictos y Office.

Dentro de las variables de control, se incorporaron:

- Características del alumno: género, curso y si es extranjero.
- Nivel socioeconómico y cultural: cantidad de libros en el hogar, índice de nivel socioeconómico construido por PISA y educación de la madre. La educación de la madre puede considerarse como una aproximación de la habilidad no observada del alumno (Hoxby, 2000; Gallego, no publicado, y Card, 1995). Sobre el nivel socioeconómico, cabe mencionar que, el índice construido por PISA incluye el acceso a un computador. En este contexto, utilizar en la estimación tanto el índice PISA como la variable de acceso, habría generado alta correlación entre ambas variables lo cual habría provocado que la variable de acceso perdiera significancia por construcción, distorsionando las posibles conclusiones del estudio. Entonces, cuando se intentó

correlacionar acceso con desempeño en Ciencias, no se incorporó el índice de nivel socioeconómico construido por PISA, ya que no permitía distinguir la relación individual entre acceso y desempeño, que este estudio intentaba obtener.

- Interés en Ciencias: para la medición se utilizó el índice de gusto en Ciencias, el cual se construyó en base a cinco preguntas del cuestionario, relacionadas con cuánto le agradan al alumno los temas científicos. A través de este índice se intentó aislar el efecto motivacional de ciertos estudiantes, que podría estar relacionado con el desempeño en Ciencias.
- Características del establecimiento: considera calidad de recursos, ruralidad, dependencia, autonomía y grupo socioeconómico. La autonomía fue medida a través del “índice de autonomía para distribuir recursos” de PISA, el cual se deriva de la cantidad de decisiones relacionadas con los recursos del establecimiento, que son de responsabilidad del establecimiento educacional (OCDE, 2007). El “grupo socioeconómico” fue construido por el Ministerio de Educación, y parte del supuesto de que las escuelas atienden a alumnos de similares características socioeconómicas. Por ello, los establecimientos se clasifican de acuerdo con las características socioeconómicas predominantes de sus alumnos. Las variables consideradas para clasificar los establecimientos son: el nivel educacional de la madre, el nivel educacional del padre, el ingreso económico mensual familiar y el índice de vulnerabilidad escolar del establecimiento. Las primeras 3 variables se obtuvieron del cuestionario aplicado a padres y apoderados de los alumnos de 2° Año Medio, que rindieron la prueba SIMCE en 2006, y la cuarta variable es el Índice de Vulnerabilidad Escolar (IVE), calculado anualmente por la Junta Nacional de Auxilio Escolar y Becas (JUNAEB). Ahora bien, mediante el análisis de conglomerados se obtuvieron cinco grupos que representan los diferentes niveles socioeconómicos: Bajo, Medio-Bajo, Medio, Medio-Alto y Alto.

Este índice permitió medir el efecto “pares”, es decir, controlar por el contexto del establecimiento al que asiste el alumno.

En el Anexo se presentan las variables utilizadas en la educación básica para explicar el desempeño en Ciencias. Para el análisis econométrico se utilizaron los cuestionarios mencionados, pero sólo los aplicados en Chile, en el contexto de la medición 2006 (OCDE, 2006). Adicionalmente se utilizaron las siguientes variables del establecimiento educacional³:

- Dependencia administrativa.
- Matrícula.
- Grupo socioeconómico.

Dado que existen variables no observadas, el análisis econométrico puede llevar sesgo por variables omitidas, cuando éstas están correlacionadas con variables de interés. Como se ha demostrado en diversos estudios (Schulz, 2005), los estudiantes que provienen de un contexto familiar más aventajado, como el con padres mejor educados o con mejores trabajos, en general tienden a tener un desempeño mejor en términos educacionales. Padres más educados proveen de más herramientas para mejorar el aprendizaje de sus hijos, tales como, mejor apoyo académico, mayor motivación o mejores recursos. Además, los estudiantes de mayor nivel socioeconómico tienden a tener más computadores en la casa. Por lo tanto, tener un computador en el hogar puede ser considerado una *proxy* del nivel socioeconómico y, al mismo tiempo, puede tener su propio impacto en el aprendizaje de los alumnos. Este mismo problema de identificación puede darse a nivel de establecimiento educacional. En este estudio se buscó minimizar la probabilidad de omitir las variables relevantes, incluyendo el conjunto de características de los alumnos y de los establecimientos proporcionado por PISA 2006 (OCDE, 2006) e información entregada por el Ministerio de Educación⁴ para cada establecimiento educacional.

3 Información solicitada al SIMCE en septiembre de 2008.

4 Información solicitada al SIMCE en septiembre de 2008.

Sin embargo, permanecieron características que no pudieron ser observadas y que por lo tanto, no fueron controladas.

Por otro lado, existe un problema de endogeneidad que impide determinar la dirección de causalidad. Es decir, no es posible afirmar si el uso de computadores potencia el aprendizaje o si los alumnos con mejor rendimiento tienden a utilizar las TIC con mayor frecuencia. La relación inversa podría presentarse, si los alumnos con buenos rendimientos son “premiados” con computadores o con permisos para usar los computadores.

Por todo lo anterior, las estimaciones de este estudio deben ser interpretadas con precaución, como correlaciones condicionales, en el sentido de que reportan una relación entre computadores y aprendizaje, manteniendo constante otros aspectos familiares y de los establecimientos. Aún así, esta correlación condicional entrega mayor información que una correlación bivariada simple.

2. Resultados

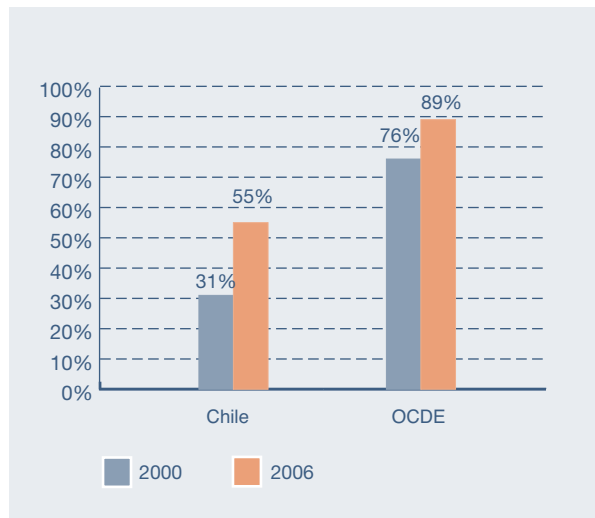
2.1 Análisis descriptivo

A continuación, se presenta evidencia sobre el acceso y uso de los computadores, por parte de los estudiantes chilenos. El objetivo es presentar el estado de la penetración de la tecnología en el país, además de realizar comparaciones que permitan identificar el nivel de avance de los indicadores de interés y determinar si éste ha sido suficiente como para alcanzar los estándares de otros países.

2.1.1 Acceso al computador

Respecto del acceso a un computador, por parte de los estudiantes chilenos que rindieron PISA 2006. En primer lugar se puede señalar que, como se observa en el Gráfico 1, la mitad de los alumnos tenía acceso a un computador en su hogar.

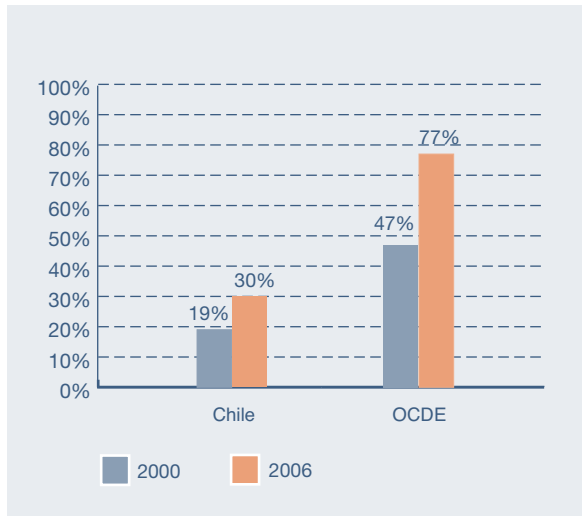
Gráfico 1: Porcentaje de alumnos con computador en el hogar



Fuente: OCDE, 2000 y 2006

Como se observa en el Gráfico 1, en el año 2000, 31% de los estudiantes chilenos tenía acceso a un computador en su hogar y en 2006, el acceso aumentó en 77% (ese año, 55% de los estudiantes contaba con un computador en su casa). Sin embargo, si bien el crecimiento en el acceso de los estudiantes chilenos es superior al mostrado por los equivalentes de la OCDE, la brecha de acceso en el hogar que existe entre los países integrantes de la OCDE y Chile es de 34 puntos porcentuales. A este respecto, cabe mencionar que, según las estadísticas nacionales de posesión de computador en los hogares chilenos, reportadas por la Encuesta CASEN (MIDEPLAN, 2007), 34% de los hogares de Chile, en 2006, poseía un computador, por lo que se puede afirmar que, en promedio, los hogares formados por al menos un alumno de 15 años tienen mayor probabilidad de tener un computador disponible.

Gráfico 2: Porcentaje de alumnos que tiene acceso a internet en el hogar



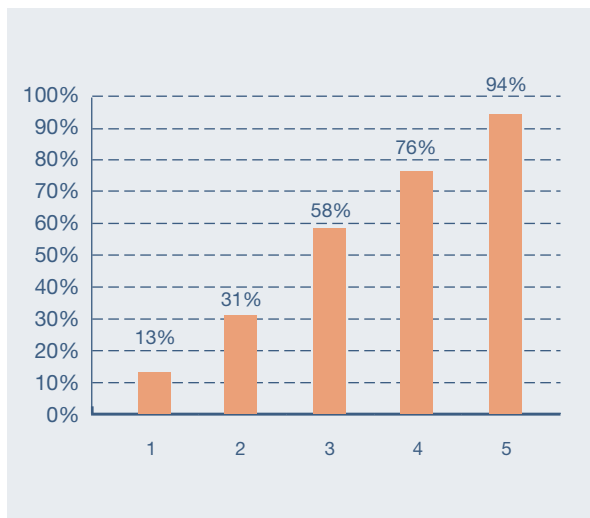
Fuente: OCDE, 2000 y 2006

En cuanto a la conexión a internet, como se observa en el Gráfico 2, un tercio de los alumnos chilenos tenía, en 2006, acceso a un computador conectado a internet en su hogar, lo que equivale a 55% (poco más de la mitad) de los que tenían un computador.

Sin embargo, pese a que el aumento de acceso a un computador en el hogar fue de 24 puntos porcentuales entre el 2000 y 2006, el aumento del acceso a un computador conectado a internet, sólo fue de 11 puntos porcentuales, y la brecha en este plano, con los alumnos de los países OCDE, es de 47 puntos porcentuales.

Ahora bien, analizados el acceso, por grupo socioeconómico del alumno, se tiene que, en los grupos socioeconómicos más altos, el acceso es sustantivamente superior.

Gráfico 3: Porcentaje de alumnos con un computador en el hogar, por nivel socioeconómico



Fuente: OCDE, 2000 y 2006

Así se muestra en el Gráfico 3, donde se puede observar el porcentaje de estudiantes chilenos que tenía acceso a un computador en el hogar el año 2006, según nivel socioeconómico. El grupo 1 representa al nivel más desaventajado y el 5, al grupo más aventajado.

Junto con lo anterior, cabe mencionar que, aunque entre los años 2000 y 2006, el acceso a un computador en el hogar, creció en todos los grupos socioeconómicos, los alumnos del grupo más bajo, experimentaron un alza menor. Asimismo, si se compara la conectividad por grupo socioeconómico, se observa que el 3% del quintil socioeconómico más bajo, contaba con un computador en el hogar conectado a internet, contra un 76% del quintil más alto (OCDE, 2000 y 2006), sin embargo, esta mayor proporción de conexión, por parte de los alumnos con mejor nivel socioeconómico, igualmente no alcanza el promedio de conectividad de los países de la OCDE para el mismo año.

Respecto del acceso a un computador en el establecimiento educacional, se observa que, en el año 2006, los establecimientos educacionales chilenos poseían en promedio 0,06 computadores por alumno y 0,04 disponibles para clases (OCDE, 2006), lo que implica que 67% de los computadores disponibles en los establecimientos educacionales estaban disponibles para las clases, teniendo así que, en promedio, un establecimiento debía administrar los recursos tecnológicos para realizar clases, sujeto a la restricción de tener un computador disponible por cada 25 alumnos. Lo anterior, aun cuando en Chile, la disponibilidad de computadores por alumno en los establecimientos, se duplicó entre 2000 y 2006. En los establecimientos de los países de la OCDE, en cambio, hay 6 alumnos por computador disponible para hacer clases (OCDE, 2000 y 2006).

Ahora bien, entre los establecimientos chilenos, se advierte una disponibilidad de recursos diferenciada: los públicos son los que cuentan con menos computadores disponibles por alumnos para hacer clases (OCDE, 2006).

2.1.2 Frecuencia de uso del computador

Conforme el acceso al computador, tanto en el hogar como en el establecimiento educacional, aumenta, es esperable que cambien las frecuencias de uso.

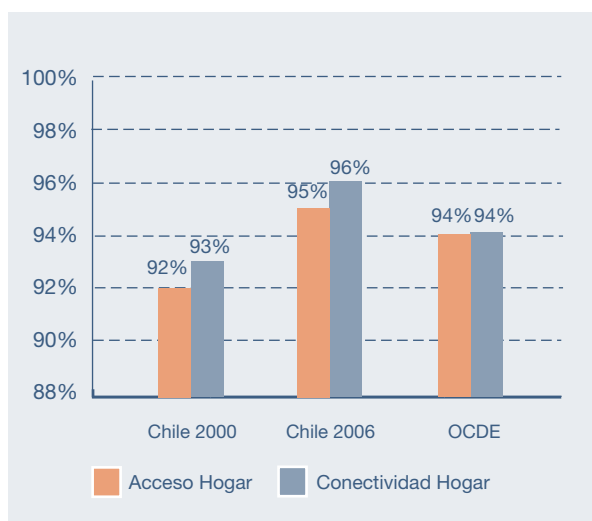
A continuación, se revisan las frecuencias de uso del computador en el hogar, en el establecimiento y en otros lugares, de manera de determinar si la frecuencia se ha visto afectada por el aumento en el acceso. Para facilitar la revisión, se define como “uso

5 Esta información se obtuvo de la encuesta respondida por los directores de los establecimientos educacionales, a quienes se les preguntó cuántos de los computadores que en total había disponibles en el establecimiento, estaban disponibles para las clases de los alumnos y, de esos, cuántos tenían conexión a internet. Esta última pregunta, sin embargo, no fue considerada para este estudio, porque entregaba información contradictoria. En efecto, por construcción, la cantidad de computadores disponibles con conexión a internet debió ser menor o igual a la cantidad de computadores disponibles, sin embargo, ello no ocurrió en todos los casos.

poco frecuente” aquel cuyo tramo va desde “alguna vez al mes” hasta “alguna vez al año”; como “uso frecuente”, el que se realiza una vez a la semana o más, y como “nunca”, cuando se responde que no se usa el computador.

En términos de la frecuencia, se puede señalar que si el alumno tiene un computador en el hogar, en general hace de él un “uso frecuente”.

Gráfico 4: Porcentaje de alumnos que usa el computador frecuentemente en el hogar según acceso y conectividad

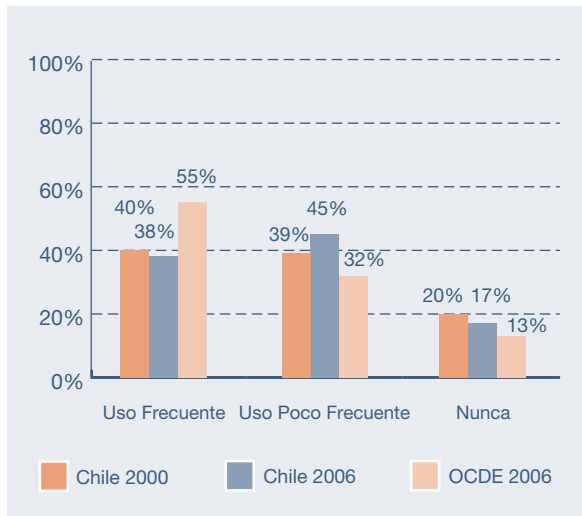


Fuente: OCDE, 2000 y 2006

Así, como se observa en el Gráfico 4, en el año 2006, 95% de los alumnos chilenos que tenía acceso a un computador en el hogar, lo usaba de manera frecuente. Es interesante notar que la frecuencia de uso de los alumnos chilenos el año 2006 es levemente superior a la de los estudiantes de los países de la OCDE, en particular a la de aquellos alumnos con un computador conectado a internet.

A diferencia de la situación anterior, se puede señalar que, como se observa en el Gráfico 5, los alumnos usan con moderada frecuencia el computador disponible en el establecimiento.

Gráfico 5: Frecuencia de uso del computador, en el establecimiento educacional



Fuente: OCDE, 2000 y 2006

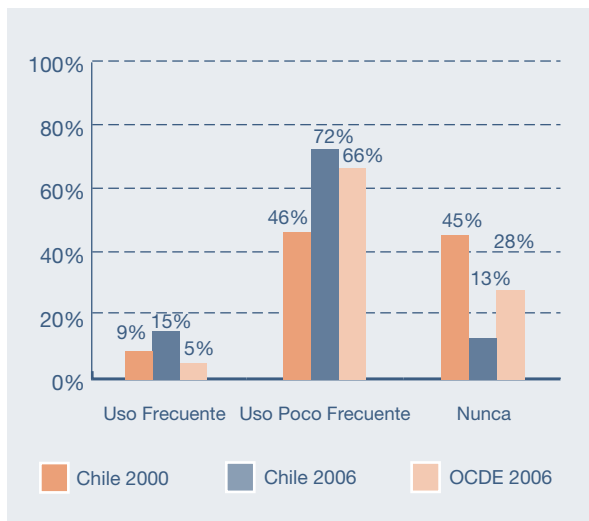
En efecto, en el Gráfico 5 se compara la frecuencia de uso del computador en el establecimiento, mostrando, para cada grupo, tres columnas: la primera representa a Chile en el año 2000, la segunda a Chile, en el año 2006 y la tercera, a la OCDE, en el año 2006.

Considerando el primer grupo de columnas (“uso frecuente”), se puede observar que el año 2006, menos de la mitad (38%) de los estudiantes chilenos usaba el computador frecuentemente en el establecimiento educacional (incluso solo 5% declaró usarlo “casi todos los días”). En contraste, poco más de la mitad (55%) de los estudiantes de los países de la OCDE declaró usarlo frecuentemente.

A este respecto, cabe mencionar que, si bien entre 2000-2006 la razón de computadores por alumno en los establecimientos de Chile se duplicó (OCDE, 2000 y 2006), la frecuencia de uso experimentó sólo un alza moderada. La proporción de alumnos que nunca usa el computador en el establecimiento disminuyó en 3 puntos porcentuales.

De forma complementaria a lo anteriormente expuesto, resulta interesante notar que, según los datos, existe una creciente frecuencia de uso de los computadores, en lugares diferentes al hogar y al establecimiento educacional, como se muestra en el Gráfico 6.

Gráfico 6: Frecuencia de uso del computador en otros lugares



Fuente: OCDE, 2000 y 2006

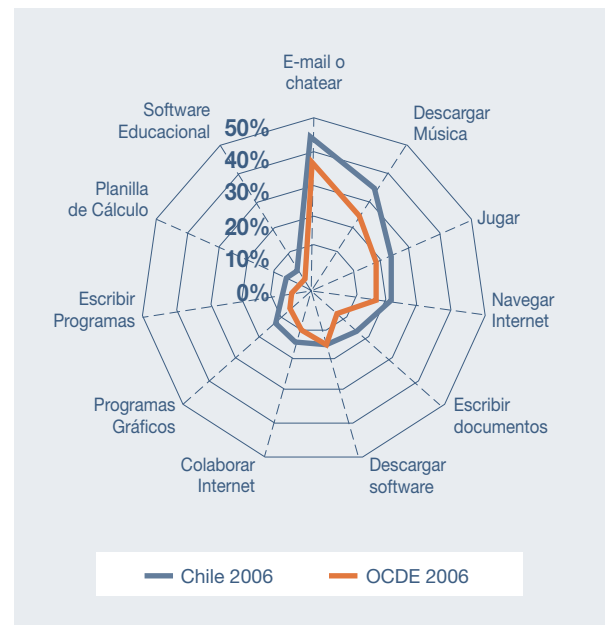
Probablemente, producto de la proliferación de lugares que proveen el servicio de computadores y/o internet, la proporción de alumnos que usa con alguna frecuencia computadores fuera de su casa y del establecimiento educacional, aumentó en 29 puntos porcentuales entre 2000 y 2006. Específicamente, en el año 2006, 83% de los estudiantes chilenos usaba el computador en otros lugares.

Comparativamente, en el 2006 en Chile, el uso del computador en un lugar diferente al hogar y al establecimiento era más frecuente que en los países de la OCDE, donde 28% de los alumnos de 15 años nunca usa el computador en otros lugares.

2.1.3 Patrón de uso del computador

Utilizando la información recolectada por PISA, sobre las actividades que se ejecutan con el computador y sobre qué tan bien pueden hacer esas tareas de manera autónoma, se puede afirmar de los estudiantes que hacen un “uso frecuente” del computador, que, aunque existen ciertos patrones comunes entre Chile y los países miembros de la OCDE, se observan comportamientos particulares de los estudiantes chilenos.

Gráfico 7: Porcentaje de alumnos que respondió usar frecuentemente el computador, para las diferentes actividades



Fuente: OCDE, 2000 y 2006

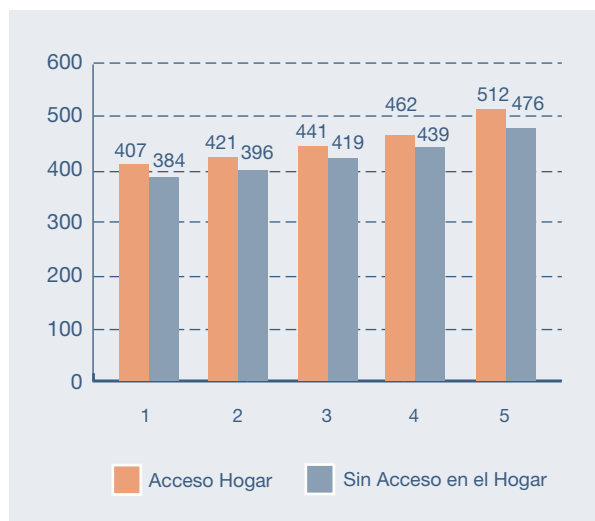
En efecto, como se muestra en el Gráfico 7, los estudiantes chilenos demuestran una mayor intensidad de uso en todo tipo de actividades, sobre todo en aquellas de comunicación e internet. Específicamente, 46% de los alumnos respondió que enviaba mails o chateaba todos los días, en comparación al 39% de los alumnos de la OCDE. Asimismo, cuando se compara a los alumnos chilenos con y sin acceso

a un computador en el hogar, también se verifican similares patrones de uso, siendo las actividades de entretención, como chateo, envío de correos, descarga de música y jugar, las más frecuentes. La diferencia se presenta en la intensidad de uso, donde, como se señaló anteriormente, quienes tienen acceso a un computador presentan una mayor frecuencia de uso (OCDE, 2006).

2.1.4 Relación entre los resultados en la prueba PISA de Ciencias y el acceso a un computador

En el Gráfico 8, se compara el promedio en Ciencias de los estudiantes chilenos, con o sin computador en el hogar, según grupo socioeconómico.

Gráfico 8: Rendimiento en la prueba PISA 2006 de Ciencias, por nivel socioeconómico, y según el acceso de computador en el hogar



Fuente: OCDE, 2000 y 2006

Al comparar a los alumnos que cuentan con un computador en el hogar con los que no lo tienen, se encuentra que el puntaje obtenido en la prueba PISA 2006 de Ciencias de los primeros es consistentemente mayor, para todos los grupos socioeconómicos. Asimismo, la brecha de resultado entre el grupo con

acceso y sin acceso es casi idéntica, aproximadamente 24 puntos, en todos los grupos socioeconómicos, excepto en el más alto, donde, habida cuenta de esta información, pareciera que la falta de acceso de computador en el hogar, genera un mayor impacto marginal negativo⁶.

2.1.5 Relación entre los resultados de la prueba PISA de Ciencias y la frecuencia por tipo de uso del computador

Dado que es posible que el impacto en el aprendizaje, dependa del tipo de uso que se le dé al computador, a continuación, se relaciona la frecuencia de uso para diferentes actividades, con el desempeño en la prueba PISA 2006 de Ciencias. Ahora bien, dado que los registros PISA refieren a una gran cantidad de usos, para efectos de este estudio, éstos se agruparon mediante un análisis factorial (para detalles, véase el Anexo):

1. Usuario internet: estudiantes que usan el computador para tareas relacionadas con internet, como e-mail o chat, descargar música, navegar, descargar software y colaborar en internet.
2. Tecno-adictos: estudiantes que podrían considerarse como que poseen una “gran fascinación” por la tecnología y la informática, haciendo o utilizando la tecnología por diversión y/o por el reconocimiento que conlleva. Este tipo de usuarios, realiza frecuentemente, tareas como, descargar música, jugar, programar, e-mail o chat y descargar software.

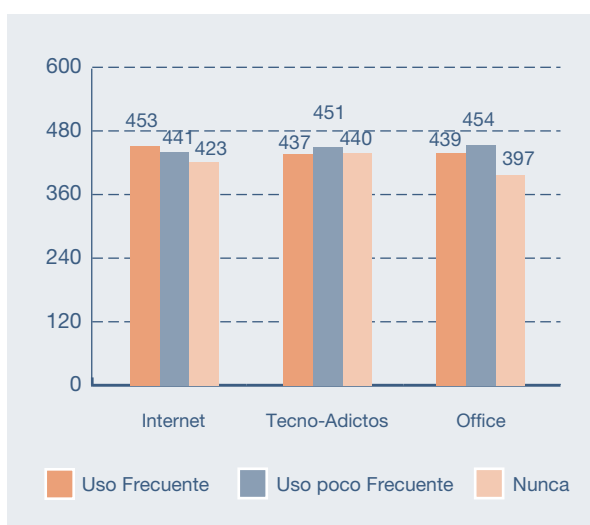
A los anteriores, se agrega un tercer usuario, que no se desprende del análisis factorial, y que se ha definido como “usuario Office”. Este es aquel que utiliza programas tipo Office, relacionados principalmente con tareas académicas, y se incorpora al estudio, pues se considera interesante verificar la relación entre estas actividades y el desempeño.

⁶ Estos resultados son estadísticamente significativos.

Para el análisis de los datos que se presentan a continuación, es necesario tener presente que un estudiante puede encontrarse en más de un grupo, pues el interés está en relacionar distintos tipos de usos con el desempeño.

En el Gráfico 9, se compara el promedio de los estudiantes chilenos en Ciencias, considerando distintas frecuencias de uso, según el tipo de uso dado.

Gráfico 9: Rendimiento en la prueba PISA 2006 de Ciencias, según tipo de uso del computador



Fuente: OCDE, 2000 y 2006

Para cada tipo de usuario, la primera columna muestra el puntaje de aquellos alumnos que “frecuentemente” usan el computador, la segunda para aquellos que lo usan de forma “poco frecuente” y la última, para aquellos “nunca” lo usan. El primer grupo de columnas compara los promedios para los usuarios de internet; la segunda, para los “tecno-adictos” y la última, para los usuarios de Office. Se observa que, a medida que aumenta la frecuencia de uso del computador en actividades relacionadas con internet, el promedio de rendimiento de los alumnos también aumenta. Por su parte, la relación entre rendimiento y frecuencia, en

los tecno-adictos y usuarios Office tiene forma de U invertida, lo cual da cuenta de la posible existencia de un óptimo en la frecuencia de uso del computador, cuando se usa para cierto tipo de tareas⁷.

2.2 Análisis econométrico

En esta sección del estudio, se presentan las correlaciones condicionadas entre tecnología y desempeño en Ciencias. En este ámbito, se puede señalar que la relación entre acceso en el establecimiento y desempeño es inexistente, y la relación entre acceso en el hogar y desempeño es de baja magnitud. En cambio, existen relaciones positivas y significativas entre: frecuencia y desempeño, y entre uso del tipo Office. Por su parte, la relación entre desempeño y usos relacionados con internet es no significativa, y para usos del tipo tecno-adictos, es negativa y significativa. En el Anexo, se muestran los resultados encontrados para cada ecuación.

A este respecto, cabe señalar, que las regresiones estimadas explican alrededor de 12% de la varianza dentro del establecimiento educacional y 86% de la varianza entre establecimientos educacionales. La baja proporción de la varianza explicada a nivel de alumno, se debe a que las características de los alumnos usadas para explicar el desempeño, no se distribuyen homogéneamente entre la población; sino que distintos tipos de alumnos, se encuentran en diferentes grupos de establecimientos educacionales. Por lo tanto, no es posible separar las características de los alumnos de las características de su establecimiento (véase una explicación más detallada en el Anexo). Dado que los puntajes se han estandarizado de tal forma que la desviación estándar es de 100 puntos, los estimadores pueden interpretarse como puntos porcentuales de la desviación estándar internacional.

⁷ Estos resultados son estadísticamente significativos.

2.2.1 Variables de control

En relación con las variables de control, en general se encuentra que, para todos los modelos, las variables tienen una relación estadísticamente significativa con el desempeño en Ciencias y están alineados con estudios previos.

Así, el desempeño de las mujeres en Ciencias es menor que el de los hombres, presentando una diferencia promedio significativa de 19 puntos. Asimismo, considerando que a los estudiantes de 15 años les corresponde, por edad, estar cursando 2° Año Medio, se observa que aquellos alumnos que cursan niveles inferiores, tienen un peor desempeño, lo que podría atribuirse al hecho de que es probable que los alumnos que se encuentran en niveles más bajos con 15 años, sean aquellos que han tenido que repetir algún curso. Por otro lado, los estudiantes que se encontraban en 3° Año Medio, tienen un desempeño superior que el resto. Todas estas estimaciones están hechas, utilizando como referencia a los estudiantes que se encontraban en 8° ó 7° Año de Educación Básica. Los alumnos extranjeros que rindieron la Prueba PISA en Chile, muestran un peor desempeño que los nacionales, presentando una diferencia de alrededor de 22 puntos.

En otro ámbito, la cantidad de libros que tiene el alumno en su hogar, así como el índice de nivel socioeconómico general, son factores relacionados positivamente con el desempeño en Ciencias: a mayor cantidad de libros y mayor índice socioeconómico, mejor desempeño. Consistentemente, aquellos alumnos cuya madre no terminó la Educación Básica obtienen puntajes menores que sus pares.

En cuanto a otras variables, cabe señalar que el gusto por las ciencias tiene un efecto positivo sobre el puntaje, como así también la autonomía y el nivel socioeconómico de los establecimientos a los que asisten los alumnos.

2.2.2 Variables de acceso a computador

Como se observa en la Tabla 1, el acceso a computador en el hogar tiene una relación positiva y significativa con el desempeño, sin embargo, su magnitud es baja: 5% de una desviación estándar de la distribución de puntajes PISA 2006⁸. Al contrario, el acceso en el establecimiento educacional, medido como la razón de computadores a matrícula, no muestra una relación significativa con el desempeño.

8 Al incluir en la regresión el Índice de Nivel Socioeconómico, la relación entre acceso y desempeño se vuelve no significativa.

Tabla 1.
Resultados regresión acceso variables de frecuencia

Parámetro	Estimación	Error típico	t	Sig.	Intervalo de confianza 95%	
Tiene computador en la casa	4,912816	2,347896	2,092	0,036	0,309792	9,515839
Razón computador-matricula	122,182361	84,669211	1,443	0,151	-45,224859	289,589581

Tabla 2.
Resultados regresiones frecuencias

Parámetro	Estimación	Error típico	t	Sig.	Intervalo de confianza 95%	
Uso poco frecuente casa	25,911736	13,494633	1,920	0,055	-0,549607	52,373079
Uso frecuente casa	30,239741	11,992827	2,521	0,012	6,723264	53,756219
Uso poco frecuente establecimiento	8,541579	3,221341	2,652	0,008	2,225978	14,857179
Uso frecuente establecimiento	0,745226	3,567892	0,209	0,835	-6,249960	7,740411

2.2.3 Variables de frecuencia de uso del computador

Del análisis de las variables de frecuencia de uso (Tabla 2), se tiene que, aquellos alumnos que nunca usan el computador en el hogar, obtienen 24 puntos menos en la prueba, que aquellos que lo usan con baja frecuencia, y 30 puntos menos, que aquellos que lo usan frecuentemente en el hogar. Sin embargo, un uso poco frecuente, es sólo significativo al 10% de confianza estadística. Respecto de la frecuencia de uso en el establecimiento educacional, el uso poco frecuente está relacionado positivamente con desempeño y el uso frecuente no tiene una relación estadísticamente significativa⁹.

Al comparar el uso de computador en el hogar con la cantidad de libros del hogar, considerado como referente de capital cultural, se verifica que el uso frecuente del computador, sería equivalente a que el alumno tuviera entre 200 y 500 libros en el hogar.

2.2.4. Variables de tipo de uso del computador

Respecto de los tipos de uso, como se observa en la Tabla 3, el uso de internet no está relacionado con un mayor desempeño en la Prueba PISA de Ciencias; sin embargo, los usuarios tecno-adictos parecieran

tener un desempeño inferior. Esto último, puede estar asociado al hecho de que estos alumnos destinan una gran proporción de su tiempo a actividades relacionadas con la tecnología, despreocupándose de sus labores escolares.

En cuanto al uso tipo Office, se observa una relación positiva entre el uso poco frecuente y frecuente, con el desempeño; siendo mayor el efecto en aquellos que hacen un uso poco frecuente del computador. En este sentido, las tareas de Office pueden ser consideradas herramientas básicas del uso de computadores, y que permiten acceder a usos más sofisticados.



⁹ Este resultado se mantiene, aun si se incluye el Índice de Nivel Socioeconómico del alumno. Nuevamente el acceso en el hogar pierde significancia.

Tabla 3.
Resultados regresión por tipo de uso

Parámetro	Estimación	Error típico	t	Sig.	Intervalo de confianza 95%	
Uso poco frecuente internet	-2,202237	3,586116	-0,614	0,539	-9,233211	4,828737
Uso frecuente internet	-4,239798	4,346597	-0,975	0,329	-12,761782	4,282186
Uso poco frecuente tareas de “tecnoadictos”	-8,714779	3,121824	-2,792	0,005	-14,835473	-2,594086
Uso frecuente tareas de “tecnoadictos”	-19,069756	4,233771	-4,504	0,000	-27,370533	-10,768979
Uso poco frecuente Office	16,353293	4,544145	3,599	0,000	7,443996	25,262590
Uso frecuente Office	10,549175	4,857730	2,172	0,030	1,025079	20,073271

3. Conclusiones

Como se señaló, este estudio tiene dos objetivos principales. Por un lado, mostrar evidencia sobre acceso y uso de computadores de los estudiantes chilenos y por otro, determinar las correlaciones que existen entre dicho acceso y uso, y el desempeño en la prueba PISA 2006 de Ciencias. La hipótesis, es que el mero acceso a la tecnología, no tiene impacto en los resultados, y que la frecuencia del uso, no siempre se relaciona positivamente con el rendimiento: la relación está condicionada al tipo de actividad desarrollada en el computador.

Respecto de la evidencia que se buscaba, cabe señalar que, en términos generales, se observa que los estudiantes chilenos tienen un acceso limitado a computadores e internet, y que ambos accesos están determinados por su nivel socioeconómico. Sin embargo, tanto la proporción de alumnos con acceso a un computador como la con acceso a internet, ha experimentado grandes tasas de crecimiento entre 2000 y 2006, aun cuando no alcancen la penetración de tecnología a la que acceden los estudiantes de los países de la OCDE. Este hecho, en el largo plazo, puede significar una desventaja económica relativa, producto de la diferencia de habilidades tecnológicas desarrolladas hoy por la futura fuerza laboral.

Si se consideran las estadísticas nacionales de posesión de computador en los hogares chilenos, se observa una proporción menor a la encontrada en la muestra de estudiantes PISA; en otras palabras, los hogares formados por al menos un alumno de 15 años, al parecer tienen mayor probabilidad de tener un computador disponible. Esto podría indicar que la decisión de compra de un computador es afectada por la composición del hogar, especialmente si el hogar posee estudiantes que pueden potenciar su uso, lo que coincidiría con lo señalado por Pedró (2006), en términos de que es posible que los *New Millenium Learners* estén presionando por un ambiente con mayor tecnología.

A nivel socioeconómico, sin embargo, se encuentra una brecha de acceso de gran magnitud: 81 puntos porcentuales entre el grupo socioeconómico más alto y el más bajo. Esta situación provoca que la posesión de un computador sea un buen indicador de grupo socioeconómico, lo cual presenta dificultades para el análisis econométrico.

El acceso a computadores en el establecimiento también es limitado si se compara con la situación enfrentada por los estudiantes de la OCDE. En Chile, hay un computador disponible por cada 25 alumnos mientras que en promedio, en la OCDE hay un computador disponible por cada 6 alumnos.

En términos de frecuencia y tipo de uso de los computadores, los resultados descriptivos indican que todos los alumnos tienen preferencias de uso similares, tanto dentro del país, como en los países de la OCDE.

Los estudiantes que poseen computador en su hogar, reportan altos niveles de uso, los que se refuerzan cuando el computador tiene acceso a internet. En el establecimiento educacional, en cambio, la frecuencia de uso se concentra en niveles medios, lo que coincide con la factibilidad real de acceder a un computador, dada la cantidad de alumnos por equipo o la falta de experiencia de los docentes en el uso efectivo de la tecnología en el proceso educativo (Barrera-Osorio y Linden, 2009).

El patrón de uso, por su parte, se concentra en actividades comunicativas y de entretenimiento. Los alumnos que tienen un computador en el hogar intensifican el uso bajo el mismo patrón.

Al comparar el rendimiento entre alumnos con y sin acceso a computador en el hogar, se encuentra una brecha de alrededor de 24 puntos. En consecuencia, se espera que el acceso a la tecnología se relacione positivamente con el rendimiento escolar, condicionado a las características del alumno y del establecimiento al cual asiste. Sin embargo, hay que tener en consideración que esta observación no implica causalidad: el uso de computadores puede potenciar el aprendizaje o puede que los alumnos con mejores puntajes sean “premiados” con acceso a un computador.

La relación de la frecuencia de uso del computador con el rendimiento escolar, es controversial. Tal como se mencionó anteriormente, existen estudios que permiten afirmar que es positiva y otros que indican que es negativa. De este estudio, se obtiene que existe una relación condicionada al tipo de uso dado por el alumno. Así, la frecuencia de uso de internet se relaciona positivamente con el rendimiento, y

la frecuencia de uso de tecno-adictos y Office, se relaciona positivamente en un principio, pero luego, para uso frecuente, la relación se vuelve negativa. En consecuencia, el uso del computador, para ciertas tareas relacionadas con un usuario tecno-adicto o tareas relacionadas con Office, podría aportar al rendimiento escolar, hasta que se transformara en una distracción o hasta que restringiera o limitara los tiempos dedicados al estudio.

Ahora bien, en cuanto a los resultados del análisis econométrico, éstos no permiten afirmar con certeza que existe un efecto del acceso sobre el rendimiento. La baja relación encontrada entre acceso en el hogar y desempeño, podría ser explicada por variables no observables, relacionadas con el nivel socioeconómico o la motivación de los padres. Por otro lado, la ausencia de relación entre acceso en el establecimiento y resultados, puede deberse al bajo uso que se le da al computador, dado los niveles actuales de inversión, o porque los docentes no los usan en forma apropiada (Barrera-Osorio y Linden, 2009). Por lo tanto, estos resultados estarían alineados con la hipótesis de este estudio, de que el mero acceso a los computadores no mejora el desempeño.

Por otro lado, se encuentra una relación positiva entre desempeño y frecuencia de uso, tanto en el establecimiento, como en el hogar. Sin embargo, esta relación oculta que la herramienta tecnológica tiene diferentes usos, los que se relacionan de distintas formas con el desempeño. El uso de computadores podría tener un efecto sobre el rendimiento cuando este es usado como herramienta de apoyo, pero también podría utilizarse como distracción, lo cual tendría un efecto nulo o incluso negativo sobre el aprendizaje. Específicamente, se encuentra que no existe relación entre el uso de internet y el conocimiento de Ciencias; una relación negativa entre uso del tipo “tecno-adicto”, y una relación positiva con programas de productividad tipo Office.



Finalmente, del análisis econométrico se obtiene que, en el establecimiento educacional, se encuentra una relación positiva entre desempeño y frecuencia, sólo para un uso poco frecuente. La relación entre uso frecuente y desempeño es no significativa. Probablemente, a pesar de los avances en el acceso a computadores en el establecimiento, éste sigue siendo tan deficiente que no permite transformarse en un herramienta activa y válida para realizar o apoyar las clases e impactar la frecuencia del uso, al interior del establecimiento educacional. Por otro lado, no hay que olvidar la función que cumplen los docentes en la introducción del computador en el aula. Si los docentes no saben cómo utilizar el computador como herramienta educativa, probablemente la frecuencia de uso no se vea alterada (Barrera-Osorio y Linden, 2009).

A modo de conclusiones generales, cabe señalar que la evidencia recogida en este estudio, indica que la alta desigualdad en el acceso a las TIC en los hogares, podría ser uno de los elementos condicionantes de la desigualdad en los resultados académicos de los alumnos. Además, de todos los elementos relacionados con la falta de capital social de las familias, actualmente, los grupos socioeconómicos más pobres concentran la falta de acceso a las TIC en entornos que favorezcan mejores aprendizajes de manera directa o indirecta.

Asimismo, las políticas de telecomunicaciones deberían revisar las opciones para que las familias de menores ingresos puedan tener acceso a internet de forma gratuita o a muy bajo costo. La actual brecha tecnológica en conectividad es mucho más profunda que la tradicional brecha en libros o material didáctico en el hogar. En Chile, el costo mensual de una conexión es equivalente a más del 10% del salario mínimo de un trabajador. Resulta imprescindible buscar opciones que permitan conexiones de bajo costo.

Por otra parte, la evidencia sobre la relación entre uso de programas de productividad tipo Office y el rendimiento de los alumnos, también permite avanzar en el análisis de la formación de competencias en los alumnos y los docentes. Las orientaciones de informática educativa tienden a especificar que no importa que los docentes sean buenos usuarios de herramientas de productividad, sino que sean capaces de hacer uso pedagógico de ellas. Dado que son los docentes los que guían el trabajo de los alumnos, esta evidencia también permite profundizar en la discusión sobre cuán importante puede ser que los docentes sean buenos usuarios de estos programas, como condición de hacer un buen uso pedagógico de los mismos y de poder desarrollar estas competencias en sus alumnos.

En virtud de lo anterior, también resulta interesante analizar los beneficios de establecer certificaciones para alumnos y docentes en el uso de herramientas de productividad, como el primer escalón en el desarrollo de competencias TIC.

Los resultados de este estudio entregan antecedentes que permiten ampliar la discusión sobre equidad tecnológica, y colaborar en la toma de decisiones, respecto de dónde debe priorizarse la futura inversión tecnológica y cuáles son las competencias mínimas a desarrollar en los alumnos y los docentes. Sin embargo,

hay que tener presente que los resultados presentados no pueden considerarse relaciones causales. En primer lugar, hay un problema de identificación, ya que no es posible controlar por variables no observadas correlacionadas, tanto con acceso a las TIC, como con desempeño. El uso de computadores puede ser una *proxy* del nivel socioeconómico o ser el resultado de padres más educados y/o motivados que proveen de más herramientas para mejorar el desempeño de sus hijos. Por otro lado, existe un problema de endogeneidad que impide determinar la dirección de causalidad, es decir, no se puede distinguir si la relación encontrada corresponde a un impacto del uso de tecnología sobre el aprendizaje o si el uso de tecnologías es el resultado de un buen desempeño académico, lo que ocurriría, por ejemplo, si el alumno es premiado con acceso a computadores cuando obtiene un buen rendimiento.

Para avanzar en propuestas de política pública, sería interesante contar con mayor evidencia cuantitativa y con estudios pilotos experimentales que permitan evaluar el impacto de distintas alternativas, antes de definir las grandes inversiones de los próximos años. Desde una perspectiva metodológica, sería preferible usar un estudio longitudinal para determinar el efecto de la innovación tecnológica que permita eliminar sesgos relacionados con el contexto del alumno (Haertel y Means, 2000).

Otras limitaciones de este estudio, se relacionan con la imposibilidad de indagar en profundidad sobre el contexto donde se desarrollan las innovaciones tecnológicas. El acceso a computadores es sólo un componente en la implementación de tecnología, tanto en la escuela como en los hogares, y no suele ser el componente más influyente (Haertel y Means, 2000 y Barrera-Osorio y Linden, 2009). Es decir, existe una serie de variables que determinan la “calidad de acceso” de los alumnos. A modo de ejemplo, no es

posible afirmar que dos estudiantes que asisten a establecimientos con la misma razón de matrícula-computadores, tienen el mismo acceso; ya que uno de ellos puede tener un profesor que domina perfectamente la herramienta tecnológica y el otro no; o bien, en un establecimiento puede que la sala de computadores permanezca cerrada la mayor parte del tiempo versus otro establecimiento, donde los alumnos tienen libre disposición de los computadores. Es por esta razón, que no podemos afirmar con certeza que el acceso a computadores en el establecimiento no está relacionado con el aprendizaje, ya que no podemos distinguir otros aspectos relacionados con la calidad del acceso, que podrían estar afectando la relación.

Finalmente, es interesante destacar que, en el cuestionario PISA, se incluye una pregunta sobre la experiencia en el uso de computadores. Si el uso del computador efectivamente favorece el aprendizaje, entonces probablemente quienes han tenido acceso a un computador desde hace más tiempo, se verán más beneficiados que quienes sólo recientemente han adquirido un computador. La investigación de este tema podría ampliarse en esta línea.

Referencias

- Angrist, J. y Pischke, J. V. (2002). "New Evidence on Classroom Computers and Pupil Learning". *Economic Journal*, 112. Oxford, Reino Unido: Royal Economic Society.
- Barrera-Osorio, F. y Linden, L. (2009). "The Use and Misuse of Computers in Education. Evidence from a Randomized Experiment in Colombia". *Impact Evaluation Series*, 29. Washington: Banco Mundial.
- Beltran, D., Das, K. y Fairlie, R. (2006). "Do Home Computers Improve Educational Outcomes? Evidence from Matched Current Population Surveys and the National Longitudinal Survey of Youth 1997", *IZA Discussion Papers*, 1912. Bonn, Alemania: Institute for the Study of Labor (IZA).
- Borghans, L. y Ter Weel, B. (2004). "Are Computer Skills the New Basic Skills? The Returns to Computer, Writing and Math Skills in Britain", *Labour Economics*, 11(1). Amsterdam, Países Bajos: Elsevier B.V.
- Card, D. (1995). "Using Geographic Variation in College Proximity to Estimate the Return to Schooling", en Christofides, L., Kenneth Grant E. y P. Swidinsky, R. (Eds.), *Aspects of Labour Market Behavior: Essays in Honor of John Vanderkamp*. Toronto, Canadá: University of Toronto Press.
- Cuban, L. (2003). *Oversold and Underused: Computers in the Classroom*. Cambridge, Massachusetts: Harvard University Press.
- Gallego, F. (no publicado). "School Choice, Incentives, and Academic Outcomes: Evidence for Chile". Extraído desde <http://repec.org/esLTM04/up.17368.1080314323.pdf>
- Haertel, G. y Means, B. (2000). *Stronger Designs for Research on Educational Uses of Technology: Conclusions and Implications*. Extraído desde https://www.msu.edu/course/cep/807/zOld807.1998Gentry/snapshot.afs/*cep240studyrefs/beckersynthe1b.pdf
- Hoxby, C. (2000). "Does Competition Among Public Schools Benefit Students and Taxpayers?". *American Economic Review*, 90(5). Nashville, Tennessee: American Economic Association.
- Karpati, A. (Ed.) (2004). *Promoting Equity Through ICT in Education: Projects, Problems, Prospects*. Paris, Francia: Organización para la Cooperación y el Desarrollo Económico (OCDE)/ Hungarian Ministry of Education.
- Law, N., Pelgrum, W. y Plomp, T. (Eds.) (2008). *Pedagogy and ICT Use in Schools Around the World: Findings from the IEA SITES 2006 Study*. Hong Kong: CERC-Springer.
- Machin, S., McNally, S. y Silva, O. (2006). "New Technology in Schools: Is There a Payoff?", *Economic Journal*, 117(522). St. Andrews, Reino Unido: Royal Economic Society.
- MINEDUC (Ministerio de Educación de Chile) (2008). *15 años integrando tecnologías a la educación chilena*. Santiago, Chile.
- MIDEPLAN (Ministerio de Planificación y Cooperación) (2007). *Encuesta CASEN 2006*. Santiago, Chile.
- OCDE (Organización para la Cooperación y el Desarrollo Económico) (2000). *Database - PISA 2000*. Extraído desde <http://pisa2000.acer.edu.au/downloads.php>
- _____ (2003). *PISA 2003. Data Analysis Manual*. Extraído desde <http://www.OCDE.org>
- _____ (2006). *Database-PISA 2006*. Extraído desde <http://pisa2006.acer.edu.au/downloads.php>
- _____ (2007). *PISA 2006 Science Competencies for Tomorrow's World*. Extraído desde <http://www.OCDE.org>
- Papanastasiou, E., Zembylas, M. y Vrasidas, C. (2003). "Can Computer Use Hurt Science Achievement? The USA Results from PISA", *Journal of Science Education and Technology*, 12(3). Dordrecht, Holanda: Springer.
- Pedró, F. (2006). *The New Millennium Learners: Challenging our Views on ICT and Learning*. Paris, Francia: OCDE/Centre for Educational Research and Innovation (CERI).
- Sivin-Kachala, J. y Bialo, E. (2000). *2000 Research Report on the Effectiveness of Technology in Schools*. Washington: Software Information Industry Association (SIIA).
- Schulz, W. (2005). *Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003*. San Francisco, California: American Educational Research Association.
- Wößmann, L. y Fuchs, T. (2004). "What Accounts for International Differences in Student Performance? A Re-examination using PISA Data", *IZA Discussion Papers*, 1287. Bonn, Alemania: IZA.

Anexo

1. Variables utilizadas en el estudio econométrico

Tabla A.1
Variables a nivel de estudiantes y establecimiento

Variables a nivel de estudiante	Media	Mínimo	Máximo	Perdidos
Mujer	46%	–	–	0%
Curso				
Séptimo Educación Básica	1%	–	–	0%
Octavo Educación Básica	3%	–	–	0%
Primero Educación Media	19%	–	–	0%
Segundo Educación Media	71%	–	–	0%
Tercero Educación Media	6%	–	–	0%
Extranjero	1%	–	–	3%
Cantidad de libros en casa				
Menos de 10 libros	23%	–	–	3%
11-25 libros	30%	–	–	3%
26-100 libros	31%	–	–	3%
101-200 libros	10%	–	–	3%
201-500 libros	4%	–	–	3%
Más de 500 libros	2%	–	–	3%
Madre no terminó Educación Básica (8 años de educación)	15%	–	–	7%
Índice gusto por las ciencias	0,25	-2,15	2,06	2%

(Continuación en página siguiente)

(Continuación Tabla A.1)

Variables a nivel de establecimiento educacional	Media	Mínimo	Máximo	Perdidos
Índice Calidad recursos del establecimiento educacional	-0,57	-3,43	2,14	2%
Rural	4%	–	–	0%
Dependencia establecimiento educacional				
Municipal (público)	44%	–	–	0%
Particular subvencionado (privado con financiamiento público)	49%	–	–	0%
Particular pagado (privado)	7%	–	–	0%
Índice autonomía para distribuir recursos	-0,21	-1,08	2,02	2%
Grupo socioeconómico (indicador SIMCE)				
Bajo	22%	–	–	0%
Medio Bajo	39%	–	–	0%
Medio	22%	–	–	0%
Medio Alto	9%	–	–	0%
Alto	8%	–	–	0%

2. Análisis factorial

El análisis factorial es una técnica de reducción de datos que sirve para encontrar grupos homogéneos de variables, a partir de un conjunto numeroso de éstas. Estos grupos homogéneos se forman con las variables con alta correlación entre sí. Aplicando un análisis factorial, se pueden encontrar grupos de variables con significado común.

En primer lugar se realizó un análisis factorial con toda la muestra, y se encontraron dos componentes, como se muestra en la Tabla A.2. Los porcentajes corresponden a la correlación entre la variable y cada uno de los factores. El primer componente no permitía segmentar la muestra, y representaba un uso generalizado del computador. El segundo componente, permitió generar un grupo que usa el computador para tareas relacionadas con internet, tales como navegar, usar correo electrónico y realizar descargas.

Tabla A.2
Matriz de Componentes
Toda la muestra

	Componentes	
	Uso de computador	Internet
E-mail o chat	60%	62%
Descargar música	66%	53%
Navegar por internet	65%	43%
Descargar software	75%	8%
Colaborar en internet	69%	5%
Jugar	52%	-9%
Procesador de texto	65%	-13%
Programar	68%	-26%
Programas gráficos	61%	-34%
Software educacional	66%	-44%
Planilla de cálculo	66%	-44%

Con objeto de encontrar algún otro tipo de alumnos, se hizo un análisis factorial, considerando sólo a aquellos que usan el computador casi todos los días en su casa. Este análisis arrojó tres componentes, como se muestra en la Tabla A.3, los dos mencionados anteriormente y un tercero, que podría relacionarse con un tipo de persona con una “gran fascinación” por la tecnología y la informática, al que se ha denominado “tecnoadicto”, y que se caracteriza por generar tecnología

Tabla A.3
Matriz de Componentes
Alumnos que usan el computador
en su casa casi todos los días

	Componentes		
	Uso de computador	Internet	Tecno-adicto
Jugar	33%	-27%	74%
Descargar software	71%	18%	38%
Descargar música	53%	64%	19%
Programas gráficos	56%	-34%	13%
Programar	68%	-21%	6%
E-mail o chat	43%	73%	-3%
Software educacional	64%	-41%	-16%
Planilla de cálculo	64%	-44%	-19%
Colaborar en internet	68%	10%	-24%
Navegar por internet	54%	50%	-26%
Procesador de texto	54%	-29%	-34%

o usarla, por diversión y/o por el reconocimiento que conlleva.

Finalmente, se realizó un análisis factorial, considerando a los alumnos que usan el computador “casi todos los días”, en el establecimiento educacional. Como se muestra en la Tabla A.4, se encontraron los mismos tres componentes que en el análisis anterior.

Tabla A.4
Matriz de Componentes
Alumnos que usan frecuentemente el computador
en el establecimiento casi todos los días

	Componentes		
	Uso de computador	Internet	Tecno-adicto
Descargar música	55%	21%	63%
Jugar	45%	-28%	38%
Programar	65%	-21%	26%
E-mail o chat	50%	62%	20%
Descargar software	77%	-12%	18%
Programas gráficos	58%	-52%	-2%
Navegar por internet	56%	58%	-7%
Software educacional	71%	-37%	-13%
Planilla de cálculo	70%	-15%	-34%
Colaborar en internet	69%	17%	-36%
Procesador de texto	73%	21%	-39%

3. Resultados

Tabla A.5

	Ecuación básica	Acceso	Frecuencia casa	Frecuencia establecimiento	Tipo de uso
Intersección	339,59 (**)	326,4 (**)	314,29 (**)	328,36 (**)	354,85 (**)
Mujer	-19,08 (**)	-19,32 (**)	-17,42 (**)	-18,71 (**)	-20,73 (**)
Curso					
Primero Medio	45,27 (**)	43,39 (**)	39 (**)	47,12 (**)	40,52 (**)
Segundo Medio	88,82 (**)	86,38 (**)	87,67 (**)	89,94 (**)	79,82 (**)
Tercero Medio	103,07 (**)	101,31 (**)	94,16 (**)	102,94 (**)	92,72 (**)
Extranjero	-23,14 (**)	-22,37 (**)	-20,8 (**)	-20,48 (**)	-24,47 (**)
Cantidad de libros en el hogar					
11-25	1,6	2,81	-0,53	2,54	1,43
26-100	10,47 (**)	13,43 (**)	10,25 (**)	15 (**)	13,68 (**)
101-200	17,7 (**)	21,62 (**)	18,59 (**)	22,45 (**)	19,59 (**)
201-500	27,77 (**)	32,78 (**)	25,7 (**)	33,06 (**)	28,68 (**)
Más de 500	22,28 (**)	26,49 (**)	26,67 (**)	25,02 (**)	21,46 (**)
Índice nivel socioeconómico	5,45 (**)	-	5,25 (**)	-	6,65 (**)
Madre no terminó 8° Básico	-5,47(*)	-9,18 (**)	-9,43	-8,64 (**)	-5,15
Índice gusto ciencias	10,28 (**)	10,17 (**)	13,2 (**)	11,72 (**)	12,66 (**)
Índice calidad recursos establecimiento	2,87	1,35	2,37	3,4	3,76
Área rural	8,41	4,95	2,64	13,22	4,33
Dependencia establecimiento					
Particular subvencionado	1,34	0,14	-7,93	-0,78	-2,01
Particular pagado	13,72	10,48	3,22	13,99	10,72
Índice autonomía para distribuir recursos	5,67(*)	5,86(*)	4,31	4,76	4,64
Grupo socioeconómico (GSE)					
Medio Bajo	3,63	4,16	8,5	5,1	4,54
Medio	49,09 (**)	53,82 (**)	47,96 (**)	47,29 (**)	45,43 (**)
Medio Alto	69,11 (**)	76,81 (**)	71,97 (**)	68,82 (**)	63,74 (**)
Alto	91,04 (**)	100,74 (**)	94,18 (**)	92,39 (**)	86,82 (**)

(Continuación en página siguiente)

(Continuación Tabla A.5)

	Ecuación básica	Acceso	Frecuencia casa	Frecuencia establecimiento educacional	Tipo de uso
Tiene computador en la casa	-	4,91 (**)	-	6,12 (**)	-
Razón computador establecimiento educacional	-	122,18	83,33	-	-
Uso poco frecuente casa	-	-	25,91(*)	-	-
Uso frecuente casa	-	-	30,24 (**)	-	-
Uso poco frecuente establecimiento educacional	-	-	-	8,54 (**)	-
Uso frecuente establecimiento educacional	-	-	-	0,75	-
Uso poco frecuente internet	-	-	-	-	-2,2
Uso frecuente internet	-	-	-	-	-4,24
Uso poco frecuente tecno-adicto	-	-	-	-	-8,71 (**)
Uso frecuente tecno-adicto	-	-	-	-	-19,07 (**)
Uso poco frecuente Office	-	-	-	-	16,35 (**)
Uso frecuente Office	-	-	-	-	10,55 (**)
Varianza explicada					
Varianza dentro del establecimiento educacional	12%	12%	10%	12%	13%
Varianza entre establecimientos educacionales	85%	85%	87%	86%	87%
Observaciones	4.664	4.606	2.595	4.090	3.739

** Significativo al 5%

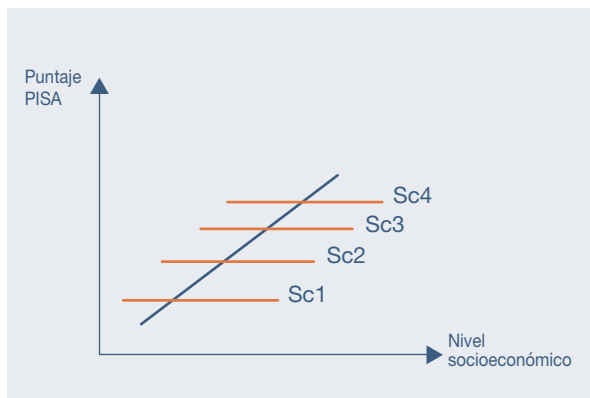
* Significativo al 10%

4. Varianza explicada

¿Cómo es posible que las variables a nivel de estudiante puedan explicar alrededor del 75% de la varianza entre establecimientos educacionales y sólo 12% de la varianza dentro del establecimiento educacional? Esto refleja principalmente la segregación de los establecimientos educacionales (OCDE, 2003).

Si en un determinado país, los establecimientos educacionales no contienen todos los rangos de nivel socioeconómico que existen en la población y, al mismo tiempo, no cubren todos los rangos de desempeño que existen; no habrá relación entre el nivel socioeconómico y el desempeño de cada establecimiento. Lo relevante para determinar el desempeño de un alumno es el nivel socioeconómico del establecimiento educacional al que asiste, el cual depende del nivel socioeconómico de los alumnos que atiende. Esta situación se refleja en el Gráfico A.1.

Gráfico A.1:



El ajuste de la regresión multinivel, representado por las líneas horizontales, difiere considerablemente de la regresión simple, representada por la línea diagonal.

En Chile, la varianza entre establecimientos educacionales es de 4.799,89 y la varianza dentro del establecimiento educacional es de 4.630,17. La proporción de la varianza total que es explicada por el establecimiento, es 51%, lo que demuestra el grado de segregación, en términos de desempeño, de los establecimientos educacionales. Esta proporción tiene un promedio de 36% para los países que rindieron PISA 2006 y varía entre 6% en Finlandia y 61% en Hungría.

Las características de los alumnos usadas para explicar el desempeño, no se distribuyen homogéneamente entre la población, sino que distintos tipos de alumnos se encuentran en diferentes grupos de establecimientos educacionales. Por lo tanto, no es posible separar las características de los alumnos de las características de su establecimiento, logrando explicar una alta proporción de la varianza entre establecimientos educacionales, pero una baja proporción de la varianza entre alumnos.

¿Cómo las escuelas chilenas pueden mejorar el aprendizaje en Ciencias?

Por
Ernesto Treviño
Francisca Donoso
Macarena Bonhomme¹

Resumen

En la presente investigación se indaga acerca de los factores escolares que se relacionan con el aprendizaje de las ciencias en Chile, usando los datos de PISA 2006. Para ello, primero se hace un análisis de la relación entre el nivel socioeconómico y cultural de los estudiantes y sus logros de aprendizaje, tanto en Chile, como en los otros países de la región (Argentina, Brasil, Colombia, México y Uruguay). En este análisis, se utilizan gradientes socioeconómicos y perfiles escolares. En segundo lugar, se realiza un análisis de los factores escolares que explican el rendimiento, a través de modelos lineales jerárquicos.

Los resultados de estos análisis indican, por una parte, que el nivel socioeconómico de los estudiantes está íntimamente relacionado con el logro en Ciencias de la prueba PISA 2006, especialmente en el caso de Chile, y por otra, que el promedio del nivel socioeconómico de los estudiantes en la escuela es la variable que se vincula más fuertemente con el rendimiento. Esto quiere decir que la segregación económica entre escuelas —entendida como el agrupamiento de estudiantes de un nivel socioeconómico similar en la misma escuela— es la variable que más explica las diferencias de aprendizaje.

Si bien el nivel socioeconómico y cultural incide en el aprendizaje, los hallazgos del análisis de modelos lineales jerárquicos muestran que la calificación de los profesores y sus prácticas pedagógicas, así como las políticas de gestión escolar que muestran responsabilización por los resultados, son elementos claves que las escuelas pueden manejar para mejorar el nivel de desempeño de los estudiantes.

1 Ernesto Treviño y Francisca Donoso son investigadores de la Facultad de Educación de la Universidad Diego Portales, y Macarena Bonhomme es socióloga de la Pontificia Universidad Católica de Chile.

Los autores agradecen los comentarios del Comité Técnico por las exhaustivas retroalimentaciones que hicieron a versiones anteriores de este capítulo. También queremos expresar nuestro agradecimiento a Claudia Matus, por sus contribuciones para mejorar los análisis estadísticos que se presentan en el texto.

Introducción

Chile es el segundo país, entre todos los participantes en PISA 2006, donde mayor influencia tiene el promedio del nivel socioeconómico de los estudiantes que asisten a una misma escuela sobre el rendimiento académico en el área de Ciencias (OCDE, 2008). En este contexto de amplias disparidades, se hace indispensable conocer las características de las escuelas que podrían ayudar a mitigar el efecto de las desigualdades sociales, sobre el logro académico de los estudiantes en Chile. Tal es el objetivo de la presente investigación.

Para lograr este objetivo, en primer lugar se presenta una breve revisión de la literatura, sobre la relación entre aprendizaje y nivel socioeconómico, y de los factores escolares que podrían mejorar el logro académico, aun en situaciones de desmedro socioeconómico. En segundo lugar, se describe el diseño metodológico del estudio, el que incluye la muestra, los instrumentos y las estrategias analíticas utilizadas. En tercer lugar, se presentan los resultados obtenidos del análisis comparativo de la relación entre aprendizaje y nivel socioeconómico, entre los países latinoamericanos participantes en PISA 2006. Luego se muestran los principales hallazgos relativos a los factores escolares que tienen potencial de mejoramiento de los logros de aprendizaje, y finalmente se presentan conclusiones y recomendaciones de política educativa.

1. Antecedentes

La investigación es contundente en señalar que el origen social de los estudiantes ejerce gran influencia sobre sus resultados académicos. De acuerdo a las teorías y estudios sobre reproducción social, los sistemas educativos están organizados de manera tal que reproducen la estructura de desigualdad social (Bourdieu y Passeron, 2003; Bowles y Gintis, 2001, y Torche, 2005). Pese a esta evidencia, se sabe que existen factores escolares que pueden tener incidencia en el

aprendizaje de los estudiantes, independientemente del nivel socioeconómico y cultural de los mismos. Por ello, en los párrafos siguientes, se destaca la situación del sistema educacional chileno, así como la desigualdad social y su relación con los resultados educativos. También se aborda la importancia de las competencias científicas en la obtención de oportunidades educativas y en el desarrollo social y, finalmente, se analizan los principales factores escolares que se asocian al aprendizaje de los estudiantes.

1.1 La relación entre el nivel socioeconómico y cultural y los resultados de aprendizaje

Chile tiene un sistema educativo donde el nivel socioeconómico y cultural muestra una estrecha relación con los resultados de aprendizaje (OCDE, 2008). Sin embargo, la asociación no se evidencia tan claramente a nivel de individuo, como de escuela. En efecto, en Chile se puede observar que la segregación escolar —entendida como el agrupamiento de estudiantes de nivel socioeconómico similar en el mismo establecimiento— afecta en mayor medida al rendimiento académico que el nivel socioeconómico individual de los estudiantes. Tal segregación se da en el contexto de un sistema de cuasi-vouchers —o subvención escolar— instaurado en 1982, que, entre otros factores, acentúa la dinámica de un sistema educativo que agrupa en las escuelas a estudiantes de estratos socioeconómicos similares. En efecto, de acuerdo a cálculos propios, y usando los datos de PISA 2006, se pudo establecer que un 53% de las diferencias en las características socioeconómicas y culturales de los estudiantes se dan entre escuelas.

Esta segregación ayuda a construir la falacia de que aquellos establecimientos que obtienen altos puntajes en las evaluaciones nacionales, son los que ofrecen la mejor educación, cuando en realidad las altas puntuaciones en los exámenes parecen ser más bien un producto de la selección de estudiantes que realizan

estos colegios, reteniendo a aquellos cuyas familias tienen el más elevado capital cultural y las mejores condiciones económicas. Por el contrario, los colegios que obtienen bajos puntajes, son aquellos que concentran un alto porcentaje de estudiantes de bajo estrato socioeconómico (Bellei, 2007). En síntesis, el espejismo de las diferencias en las puntuaciones, derivado de la segregación socioeconómica, se ha visto reforzado por el sistema de vouchers y el esquema de financiamiento compartido (Torche, 2005), así como por los mecanismos de selección de estudiantes de las escuelas y los procesos de autoselección en los que incurren las familias, al buscar un colegio para sus hijos (Bellei, 2007).

Como medida reparatoria de lo anterior, el Estado de Chile creó el financiamiento compartido: una modalidad instaurada en 1993, que permite que los padres realicen un aporte complementario a la subvención estatal, con la promesa de mejorar el nivel de aprendizaje de los estudiantes. La evidencia muestra que desde 1999 a 2006, se duplicó el aporte realizado por los padres, de \$7.000 a \$14.000 mensuales, pero ello no condujo a la obtención de mejores resultados académicos, que reflejaran un incremento de la calidad (Reyes, 2008). En consecuencia, dentro del marco organizacional de las escuelas chilenas, los más beneficiados siguen siendo los colegios con estudiantes de mayor nivel socioeconómico. Por el contrario, los estudiantes más vulnerables, mantienen niveles precarios de aprendizaje. De esta manera, el sistema escolar chileno va generando mayor inequidad y una triple segregación—residencial, por capacidad de pago y por calidad de recursos necesarios para el aprendizaje— que dificulta la integración de distintas clases sociales, al interior de un mismo colegio (Valenzuela, 2007).

En atención a los mayores costos que implica educar a estudiantes de menor nivel socioeconómico (problema que se agudiza con la segregación socioeconómica de las escuelas), desde el año 2008 en Chile se ha implementado la Ley de Subvención Escolar

Preferencial. Esta medida ofrece un mayor monto de subvención por alumno a las escuelas que reciben estudiantes de estrato socioeconómico bajo, a los cuales se denomina “estudiantes prioritarios”². Además, esta política destina un monto adicional de recursos financieros a las escuelas, por concentración de estudiantes prioritarios. Esto implica que, en general, las escuelas pueden recibir hasta 70% más de recursos por estudiante, de lo que reciben con la subvención escolar regular.

Ahora bien, junto con el problema de la segregación antes mencionado, se puede observar que, si bien Chile ha tenido una sólida expansión del sistema educativo, sus beneficios se han repartido inequitativamente. Los resultados de la encuesta CASEN 2006³, señalan que ha habido un aumento de los niveles de escolaridad de la población mayor de 18 años, entre 1990 y 2006, pasando de 9 a 10,2 años de estudios, en promedio (MIDEPLAN, 2006). Sin embargo, tal como se advirtió anteriormente, esta expansión ha beneficiado de forma desigual a los distintos estratos socioeconómicos del país. Así, durante ese periodo, mientras que 99,1% de los estudiantes de más recursos (pertenecientes al último quintil de ingresos), completaron su Educación Básica; sólo llegó a este punto 71,9% de los del quintil más pobre. En la Educación Secundaria, la situación es más grave, puesto que 30% de los estudiantes del quintil más pobre completó la Educación Media, comparado con 95% de los del quintil más rico. Finalmente, en la Educación Superior, sólo completó la

2 La clasificación de los estudiantes como “prioritarios”, se realiza, en función de distintos instrumentos sociales que existen en Chile, a saber: a) participar en el programa Chile Solidario; b) estar dentro del tercil más bajo de ingresos, de acuerdo a la Ficha de Protección Social y/o c) pertenecer al tramo A de FONASA (sistema de salud estatal). En caso de no contar con los datos anteriores, se consideran para la clasificación, los ingresos familiares del hogar, la escolaridad de la madre, el nivel de ruralidad y la pobreza comunal. Si bien la subvención escolar preferencial está en la primera etapa de implementación, orientada a los estudiantes desde Pre-kinder a 4° Básico, se espera que exista una extensión gradual de la norma, hasta abarcar toda la Educación Básica (MINEDUC, 2008).

3 CASEN es una encuesta de caracterización socioeconómica de los hogares que ofrece indicadores a nivel nacional, regional y comunal.

Educación Universitaria 3% de los jóvenes del quintil más pobre, frente al 48% de los del quintil más rico (Torche, 2005).

Como se señaló al comienzo, estas desigualdades del sistema educativo redundan en amplias brechas de aprendizaje entre los estudiantes. En efecto, las evaluaciones de logros de aprendizaje nacionales (realizadas por el Sistema de Medición de la Calidad de la Educación, SIMCE) evidencian diferencias entre los que pertenecen a escuelas con distintos tipos de dependencia administrativa: municipal (escuelas financiadas y administradas por entidades públicas); particular subvencionada (escuelas financiadas con recursos públicos, pero administradas por privados); y particular pagada (escuelas de financiamiento y administración privados).

La tendencia se mantiene invariable en el tiempo: los estudiantes que asisten a escuelas particulares pagadas, superan sistemáticamente a aquellos que pertenecen a escuelas municipales o particulares subvencionadas, que son las de menores recursos. Específicamente, desde los años 90, la diferencia promedio entre ambos tipos de escuelas ha sido de 0,3 a 0,4 desviaciones estándar en los puntajes de la prueba SIMCE.

Estas diferencias de rendimiento por dependencia, son análogas a las originadas por el nivel socioeconómico. Al igual que en estas últimas, no necesariamente indican que las escuelas privadas ofrecen mayor calidad de educación, pues al efecto de la selección de estudiantes, se suma el efecto de pares o “compañero” (Bellei, 2007), por la alta homogeneidad entre los estudiantes al interior de las escuelas. A este respecto, baste señalar que las escuelas particulares pagadas reciben a los estudiantes provenientes de familias de mayores ingresos, las particulares subvencionadas abarcan un rango amplio de población, pero concentran a la clase media, y la población más pobre del país, se concentra en los colegios municipales (Auguste y Valenzuela, 2004 y Valenzuela, 2007).

1.2 La relación entre el aprendizaje de las ciencias y el género de los estudiantes

Respecto del aprendizaje de las ciencias en particular (cuyos resultados PISA constituyen el corpus de este estudio), cabe señalar que, junto con la dificultad que los estudiantes pueden encontrar derivada de su nivel socioeconómico, existe otro factor determinante: el género. En América Latina, los resultados de evaluaciones de aprendizaje de Ciencias, indican que las mujeres suelen obtener resultados menores que los hombres (OCDE, 2008 y UNESCO, 2008 y en prensa). La evaluación PISA 2006, muestra que Chile es el país de la región con la mayor diferencia de rendimiento por género, seguido por Brasil y México. Sin embargo, estas diferencias no son una norma: en Argentina, las mujeres obtienen mejores resultados que los hombres, y en Uruguay y Colombia, no hay diferencias significativas en el aprendizaje de las ciencias por género (OCDE, 2008). Según la literatura especializada, son múltiples los factores que pueden explicar las diferencias de rendimiento por género en las ciencias, aunque dos ideas opuestas han marcado el debate teórico: la de que las disparidades de rendimiento se producen por una socialización diferenciada para hombres y mujeres (Scaife, 1998 y UNESCO, 2009) y la que postula que se debe a diferencias en las características del funcionamiento cerebral, innatas entre hombres y mujeres (Goldstein y otros, 2001). El estudio de las causas de estas disparidades es uno de los desafíos más importantes de la investigación educativa en Chile.

1.3 La importancia del desarrollo de las competencias científicas

Desde la perspectiva de la investigación educativa, es claro que el desarrollo de competencias científicas entre los estudiantes es indispensable para mejorar las oportunidades de aprendizaje. Los estudiantes que son capaces de analizar fenómenos naturales o sociales, que reconocen las principales relaciones que

están a la base de los fenómenos, que pueden realizar inferencias y generalizaciones, y que comunican sus ideas efectivamente, tienen mayores probabilidades de cursar con éxito sus estudios e insertarse en la sociedad (OCDE, 2008 y UNESCO, 2009). El desarrollo del pensamiento científico ayuda a los estudiantes a involucrarse en prácticas cognitivas que les permiten monitorear sus progresos y promover cambios conceptuales (Bransford, Brown y Cocking, 2000). Por otro lado, la investigación disponible indica que las escuelas hacen un mayor aporte al aprendizaje de las ciencias que el que hacen al de lenguaje, porque es casi exclusivamente la escuela la que dota a los estudiantes de los conocimientos científicos, en cambio, el desarrollo del lenguaje recibe más influencias del capital cultural de las familias (Heyneman, 2004 y UNESCO, 2008).

La importancia del conocimiento científico también puede verse en relación con el desarrollo social. Desde mediados del siglo XX fue claro que el desarrollo tecnológico estaba ligado al crecimiento de la economía (Solow, 1957). Dicha tesis fue recogida y extendida en la década de los noventa en América Latina y el Caribe, al vincular el desarrollo económico, con el conocimiento, el desarrollo tecnológico y la educación, como ejes para el crecimiento económico con equidad (CEPAL, 1992). El ámbito de la educación, también ha hecho eco de estos postulados, reconociendo la centralidad del desarrollo científico para el crecimiento de la economía y del desarrollo social (OCDE, 2008 y UNESCO, 1996). En suma, la obtención de altos logros de aprendizaje en el área de las ciencias por parte de todos los estudiantes, podría contribuir a la consecución de objetivos nacionales de largo plazo, como el crecimiento económico y la integración social.

1.4 Los factores escolares que inciden en el aprendizaje de las ciencias

A pesar de la incidencia que las disparidades socioeconómicas y culturales tienen sobre los resultados

de aprendizaje, la educación sigue representando una oportunidad para la movilidad social. En este sentido, Bourdieu y Passeron (2003) señalan que la educación puede convertirse en la mejor opción para los estudiantes de sectores más desfavorecidos. En efecto, en el año 2008, aproximadamente 70% de los estudiantes universitarios en Chile representaban la primera generación de sus familias que lograba acceder a ese nivel educativo (Donoso y otros, 2004). Sin embargo, para que la educación sea un medio eficaz de movilidad social, es imprescindible apoyar a los estudiantes cuyos padres tienen baja escolaridad, para que puedan sortear las diferencias culturales que se presenten entre el hogar y la escuela. Por esta razón, para los estudiantes más pobres, el éxito en la escuela pasa por un proceso de aculturación que permita superar tal discontinuidad (Bourdieu y Passeron, 2003). Cuando las escuelas se hacen cargo de este proceso de manera consciente y adecuada, la educación se puede transformar en una vía directa a la superación social.

Tal como se vio anteriormente, el nivel socioeconómico y cultural tiene menos influencia en los resultados de Ciencias que en los de Lenguaje. Por este motivo, las ciencias pueden ser una importante vía para acortar las brechas culturales entre los niños de familias más pobres y las escuelas, mejorando las oportunidades de aprendizaje y de movilidad social de los estudiantes. En este contexto, cabe preguntarse, ¿qué factores de la escuela tienen incidencia sobre el aprendizaje, después de considerar el origen social?

El primero de ellos, ampliamente estudiado, es el de la calidad de los docentes. Investigaciones recientes proponen que ningún sistema educativo puede lograr resultados de aprendizaje por encima de la calidad de sus docentes (Barber y Mourshed, 2007), lo que implica que el motor del rendimiento académico, después de las disparidades sociales, son los docentes (Hanushek, 2005). En el caso de las ciencias, se ha comprobado que el desarrollo de competencias científicas está

íntimamente ligado a las discusiones en la sala de clase, pues en ellas los estudiantes adquieren un lenguaje que les permite expresar ideas, explicitar sus procesos de pensamiento y aprender a plantear argumentos para resolver problemas y explicar fenómenos (Bransford, Brown y Cocking, 2000). Esto sólo es posible si el profesor juega un rol gravitante en la estructuración de las oportunidades de aprendizaje.

Ahora bien, la calidad de los docentes se cristaliza en las prácticas en el aula y su eficacia para promover el aprendizaje entre los estudiantes. En efecto, se ha demostrado que las prácticas tienen una fuerte incidencia en el desempeño de los estudiantes (Kane y Steiger, 2008; Treviño y Treviño, 2004; Wenglinsky, 2002 y 2003, y Wright, Horn y Sanders, 1997). Las buenas prácticas pedagógicas suponen la movilización cognitiva de las ideas previas de los estudiantes respecto de los fenómenos científicos (trascendiendo la comprensión basada en el sentido común, para sustituirla por explicaciones científicas), y motivan y acompañan la curiosidad por aprender y encontrar sentido a la ciencia. Los docentes más efectivos en la enseñanza de las ciencias son aquellos que reconocen este ámbito de conocimiento como uno en el cual los seres humanos construyen significados sobre fenómenos de la realidad. Estos profesores organizan sus clases haciendo a los estudiantes partícipes de dicha construcción. Para ello, se apoyan en evidencia científica, diseñando e implementando sus intervenciones pedagógicas (clases, tareas, trabajos, etc.) de forma cuidadosa, y siempre con el objetivo de promover el aprendizaje de todos los estudiantes. Además, se apoyan en el supuesto de que las conclusiones de la ciencia tienen carácter provisorio y que pueden sustituirse por mejores explicaciones a los fenómenos, dando así oportunidad a los estudiantes de valorar explicaciones alternativas (UNESCO, 2009). En suma, se trata de desarrollar en los estudiantes habilidades para comprender, aplicar, analizar, sintetizar y evaluar evidencias y conclusiones (Bain, 2007), manteniendo siempre el gusto por aprender.

Considerando lo anteriormente expuesto, la disponibilidad de docentes de calidad en las escuelas resulta fundamental para el aprendizaje de las ciencias. Sin embargo, contar con docentes calificados y capaces de desarrollar buenas prácticas, supone que se cumpla una serie de condiciones. En primer lugar, es necesario mejorar la calidad de la docencia en el aula, aumentar su cobertura y generar mecanismos de monitoreo que guíen y controlen la instrucción en el sistema educacional; lo que, a su vez, requiere generar cambios en la formación de los docentes, de tal manera que se centren en promover el aprendizaje (Carnoy, 2007). En segundo lugar, es necesario que las escuelas administren sus recursos humanos en función del aprendizaje de los estudiantes. Esto se logra a través de estrategias de selección y asignación racional de los docentes a las distintas tareas; de un proceso continuo y colectivo de capacitación interna, y de un mínimo de rotación (Arancibia, 1992 y Bellei y otros, 2004). Finalmente, se requiere capacitar a los administradores educacionales para que puedan liderar el cambio en la formación docente, procurar una distribución más equitativa de la calidad docente y promover mayor tiempo efectivo de enseñanza (Carnoy, 2007).

Un segundo factor escolar determinante en el aprendizaje en general y, por ende, también en el de las ciencias, lo constituye la rendición de cuentas de las escuelas, que permite que éstas se responsabilicen por el aprendizaje de sus estudiantes. De acuerdo a la literatura, la rendición de cuentas se basa en que los padres puedan exigir un buen servicio educativo y en que, de no obtenerlo, puedan retirar a sus hijos de la escuela (Hirschman, 1970). Para que esto sea posible se requieren padres con altos niveles de capital cultural; disponibilidad irrestricta de cupos en distintas escuelas —lo que facilitaría la opción de salir—, y capacidad de las familias o de los estudiantes para movilizarse hacia las mejores escuelas (Gershberg y Meade, 2006). En Chile no se da ninguna de estas condiciones, al menos no para la población en general, y menos en

sectores rurales. Adicionalmente, la publicación de rankings de escuelas no ha garantizado por sí sola una mejora generalizada en la educación (Bellei, 2007 y Hsieh y Urquiola, 2003). La rendición de cuentas que mejor promueve el aprendizaje es aquella que motiva la participación de la comunidad y la responsabilización por los resultados, por parte de las escuelas y de los docentes (Gershberg y Meade, 2006). En este contexto, cabe mencionar que éste es precisamente el tipo de rendición de cuentas que pretende promover la subvención escolar preferencial chilena (MINEDUC, 2008), aunque será necesario esperar para ver si su implementación efectivamente conduce a una mejora en los resultados.

Un tercer factor escolar determinante en el aprendizaje es la autonomía escolar. La transferencia de responsabilidades y recursos permiten que la escuela sea más eficaz en la prestación de servicios, ya que la transforman en el foco principal de gerenciamiento de la educación. Esto conduce a mejoras de calidad y equidad, pues existe la posibilidad de generar espacios para innovar e incentivar gastos eficientes, y porque se enfatiza la responsabilidad por el servicio entregado y el control de los procesos educativos, mediante la participación de los padres y de la comunidad (Blanco y Treviño, 2000 y Espínola, 2000). Pese a los beneficios de la autonomía escolar, esta no debe ser aplicada de manera uniforme y sin considerar las diferencias en la situación de los establecimientos, pues ello puede generar desigualdades entre las escuelas. En efecto, dado que la capacidad institucional de los establecimientos es distinta, algunos estarán mejor preparados para asumir una autonomía escolar, en la dirección que apunta la teoría (Beeby, 1986 y Slavin, 2005). Las escuelas mejor preparadas podrán alinear su proyecto educativo a las necesidades de la comunidad y administrar los recursos de forma más eficiente y acorde a sus objetivos. En cambio, los establecimientos con débiles capacidades institucionales, correrán el riesgo de encarar la autonomía como una especie de abandono que podría incluso debilitarlos aún más.



El cuarto factor escolar que se releva es el de la organización del sistema escolar. Como se dijo anteriormente, resulta muy complejo separar este factor de los otros relacionados con la dependencia, como el nivel socioeconómico del alumnado; sin embargo, igualmente se puede advertir que la regulación para los tres tipos de escuela es distinta. Las escuelas particulares pagadas enfrentan requisitos mínimos de operación y se financian a través de los pagos de los padres de familia. Las escuelas particulares subvencionadas reciben recursos públicos y de los padres, por lo que tienen mayor flexibilidad para la contratación y despido de docentes, y no están sujetos a la evaluación del desempeño docente. Por su parte, las escuelas municipales, al estar financiadas completamente por recursos públicos, tienen menos libertad en su regulación, entre otras cosas, porque están sujetos al estatuto docente, que fija diferentes reglas para salvaguardar los derechos laborales de los docentes y, en algunos casos, limita las posibilidades de cambiar a profesores que muestran un bajo desempeño. En este sentido, cabe mencionar que, al menos, los docentes de establecimientos municipales están obligados a evaluar su desempeño con un instrumento nacional, cuyos resultados tienen consecuencias concretas sobre su situación laboral (Espínola, 2000).

El último factor escolar que se considera determinante de los aprendizajes, es la selección de estudiantes y las regulaciones de admisión y expulsión. Éstas son diferentes para las escuelas de distinta dependencia administrativa (Bellei, 2007), aunque se está en proceso de cambio. En efecto, la reciente aprobación de la Ley General de Educación, y la tramitación en el congreso de leyes para instaurar una Agencia de Calidad de la Educación y una Superintendencia de Educación, buscan disminuir la posibilidad de selección de estudiantes y establecer más mecanismos de control y fiscalización en el uso de recursos públicos en el sistema escolar.

A partir de esta revisión de literatura queda en evidencia que, si bien el origen social incide en el aprendizaje, también lo hacen en gran medida la calidad de los profesores, la implementación de medidas de autonomía escolar y un esquema de rendición de cuentas orientado al aprendizaje. Dado que estos factores tienen el potencial de impactar en los resultados académicos de los estudiantes, es que se escogieron para el análisis que se realizó en este estudio.

2. Metodología

En este apartado se explica el diseño metodológico del estudio y contempla la descripción de la muestra, los instrumentos y estrategias de análisis, así como los alcances y limitaciones del estudio.

2.1 Muestra

El presente trabajo utiliza las bases de datos de la evaluación de PISA 2006, enfocándose en los resultados de aprendizaje en Ciencias. Los resultados de esta evaluación, dadas sus particulares características, reflejan las competencias de los estudiantes para la aplicación práctica de conocimientos científicos, en situaciones nuevas y desafiantes.

Además de aplicar pruebas de rendimiento a los estudiantes, PISA recoge información sobre las escuelas, a través de un cuestionario para el director y otro para los estudiantes. En la evaluación 2006 además, se incluyó, por primera vez, una serie de preguntas a los estudiantes, relativas a las prácticas docentes en el aula, que resultan claves para los propósitos de este estudio.

Para el análisis se utilizan dos muestras: una compuesta por todos los estudiantes de América Latina y otra que considera solo a los estudiantes chilenos. Con la primera, se realiza un análisis comparativo de la influencia del nivel socioeconómico en el rendimiento; con la segunda, se estudian los factores que se asocian al aprendizaje de las ciencias. En la Tabla 1 se describe la composición de ambas muestras.

Tabla 1.
Muestra de establecimientos y estudiantes que rindieron la prueba PISA en 2006, en los seis países participantes de América Latina⁴

	Am. Latina	Chile	Argentina	Brasil	Colombia	México	Uruguay
Nº de estudiantes	59.155	5.233	4.339	9.295	4.478	30.971	4.839
Nº de establecimientos	2.557	173	176	625	165	1.140	278

⁴ Las muestras de estudiantes responden a los criterios establecidos por PISA, y son representativas de la población de estudiantes de 15 años que asiste a la escuela, en cada uno de los países, siempre y cuando se usen los ponderadores muestrales. En todos los análisis las muestras se ponderan, utilizando el factor de expansión provisto por PISA para los estudiantes. En el análisis de perfiles escolares, se usan los ponderadores por escuela.

2.2 Instrumentos y estrategias de análisis

Para esta investigación se utilizan el análisis de regresión y los modelos jerárquicos lineales⁵. El análisis de regresión se usa para estudiar los perfiles escolares, que muestran la relación entre el rendimiento académico promedio de las escuelas y el nivel socioeconómico y cultural promedio del establecimiento. Los modelos jerárquicos lineales tienen dos aplicaciones en esta investigación: se usan para el análisis de gradientes socioeconómicos (donde se asocia el rendimiento de los estudiantes con el índice socioeconómico y cultural a nivel individual) y para encontrar los factores escolares que podrían ser determinantes en la mejora del rendimiento de los estudiantes en Ciencias.

Respecto del análisis de perfiles escolares, se ajustan modelos de regresión para cada país de América Latina y para la Región en general, donde el aprendizaje promedio por escuela en Ciencias, se asocia al índice socioeconómico y cultural promedio de la escuela. De forma análoga, los gradientes socioeconómicos relacionan el rendimiento de los estudiantes con su nivel socioeconómico y cultural. En este caso, por tratarse de estudiantes agrupados en escuelas, se ajustan modelos jerárquicos lineales que se corrigen por la autocorrelación de las observaciones (representadas por estudiantes agrupados en escuelas). El modelo general, aplicado a cada uno de los países de América Latina y a la Región como un todo, asocia la puntuación del estudiante en la escuela en la prueba de Ciencias con el índice socioeconómico y cultural del estudiante, centrado en la *gran media*.

El análisis de perfiles escolares y gradientes socioeconómicos permite conocer con mayor profundidad la naturaleza de la relación que existe entre

las desigualdades sociales y el aprendizaje; entendiendo por “desigualdades sociales”, las diferencias entre estudiantes y escuelas en el índice socioeconómico y cultural calculado por PISA. Ambos análisis se realizan para todos los países latinoamericanos participantes en PISA, lo que otorga la oportunidad de comparar el caso de Chile y contextualizar los hallazgos sobre las desigualdades socioeconómicas y el aprendizaje.

El análisis de gradientes y perfiles utiliza el esquema provisto por PISA de ponderadores muestrales para cada país, usando los factores de expansión de estudiantes y escuelas, respectivamente. Bajo este enfoque, la muestra de estudiantes se pondera para que represente a la población de 15 años que asiste a la escuela en cada país. Así, los estudiantes de países más poblados —como Brasil y México— tendrán un peso relativo mayor que los estudiantes de países con menor cantidad absoluta de población, cuando se realice el análisis para los países latinoamericanos en conjunto. Al utilizar este esquema se representa adecuadamente a la población real de estudiantes de 15 años del grupo de países de América Latina que participó en PISA 2006, y se obtienen estimadores insesgados de los promedios de aprendizaje de los estudiantes y de la relación de los mismos con el índice socioeconómico y cultural del estudiante y de la escuela.

Después de contextualizar las desigualdades (con los análisis de perfiles escolares y gradientes socioeconómicos en América Latina), este estudio se concentra en desentrañar los factores escolares que influyen en el rendimiento académico. Para ello se construye una taxonomía de modelos jerárquicos lineales de dos niveles: estudiante y escuela. En esta taxonomía se incorpora primero el nivel socioeconómico y cultural del estudiante y de la escuela, para después agregar los factores escolares provenientes de información del cuestionario del director y de los estudiantes. En la Tabla 2 se describen las variables que se utilizan en los modelos jerárquicos, con su definición, métrica y centrado.

5 Se usa también el análisis de componentes principales con rotación Varimax, para la construcción de índices compuestos que den cuenta de los factores escolares. Para economizar el espacio, no se expone esta metodología, cuya aplicación puede encontrarse en Affifi, Clark y May (2004). En el Anexo puede encontrarse la especificación general de los modelos de regresión y jerárquicos lineales, así como los resultados del análisis de componentes principales.

Tabla 2.
Variables utilizadas en los análisis

Variables del nivel escuela	Definición	Métrica y centrado
Escuela municipal	Identificación de escuelas municipales.	Variable dicotómica que toma el valor 1 para las escuelas municipales y 0 para otro tipo de establecimientos.
Escuela subvencionada	Identificación de escuelas subvencionadas.	Variable dicotómica que toma el valor 1 para las escuelas subvencionadas y 0 para otro tipo de establecimientos.
Escuela particular pagada	Identificación de escuelas particulares pagadas.	Variable dicotómica que toma el valor 1 para las escuelas particulares pagadas y 0 para otro tipo de establecimientos.
Rendición de cuentas	Índice compuesto de las medidas que toma la escuela para rendir cuentas a los padres respecto del aprendizaje de los estudiantes, y de las expectativas de logro que se han trazado los mismos apoderados.	Índice con media cero y desviación estándar uno, para la muestra de Chile. Centrado en la gran media.
Profesores calificados	Índice compuesto de la dotación de profesores calificados en la escuela.	Índice con media cero y desviación estándar uno, para la muestra de Chile. Centrado en la gran media.
Autonomía pedagógica con respecto a estudiantes	Índice compuesto que mide la autonomía pedagógica del establecimiento para tomar decisiones relativas a medidas disciplinarias, evaluación, admisión, libros a utilizar y cursos a impartir.	Índice con media cero y desviación estándar uno, para la muestra de Chile. Centrado en la gran media.
ISEC promedio escuela	Promedio por escuela del índice de estatus socioeconómico y cultural de PISA.	Índice con media cero y desviación estándar uno, para la muestra de América Latina, en el caso de gradientes y perfiles. En el caso de regresiones lineales jerárquicas, el índice tiene media cero y desviación estándar uno, para la muestra de Chile. Centrado en la gran media.

(Continuación en página siguiente)

(Continuación Tabla 2)

Variables del nivel estudiante	Definición	Métrica y centrado
Mujer (género)	Identificación del género del estudiante, específicamente, las mujeres.	Variable dicotómica que toma el valor 1 para las mujeres y 0 para los hombres. Se utiliza en los modelos con su métrica original.
Afinidad con las ciencias	Índice compuesto de la inclinación de los estudiantes por aprender Ciencias.	Índice con media cero y desviación estándar uno, para la muestra de Chile. Centrado en la media del grupo.
Profesores incentivan aplicación práctica de las ciencias	Índice compuesto de la frecuencia con que los estudiantes realizan experimentos prácticos en el laboratorio, diseñan sus propios experimentos y piensan en la forma de contestar preguntas científicas, a través de la experimentación.	Índice con media cero y desviación estándar uno, para la muestra de Chile. Centrado en la media del grupo.
Proyección de los estudiantes en temas científicos	Índice compuesto de las expectativas futuras de los estudiantes para estudiar Ciencias.	Índice con media cero y desviación estándar uno, para la muestra de Chile. Centrado en la media del grupo.
Profesores hacen partícipes a los estudiantes	Índice compuesto de las oportunidades que ofrecen los docentes para que los estudiantes participen y comprendan la ciencia.	Índice con media cero y desviación estándar uno, para la muestra de Chile. Centrado en la media del grupo.
Oportunidades de investigación científica en la escuela	Índice compuesto de la percepción de los estudiantes sobre las instancias que les otorga el establecimiento para incursionar en investigación científica.	Índice con media cero y desviación estándar uno, para la muestra de Chile. Centrado en la media del grupo.
ISEC	Índice de estatus socioeconómico y cultural de PISA.	Índice con media cero y desviación estándar uno, para la muestra de América Latina, en el caso de gradientes y perfiles. En el caso de regresiones lineales jerárquicas, el índice tiene media cero y desviación estándar uno, para la muestra de Chile. Centrado en la gran media para los gradientes socioeconómicos y centrado en la media del grupo en los modelos multinivel para estimar efectos composicionales.
Rendimiento académico en Ciencias	Escala de puntuaciones de la prueba de Ciencias PISA.	Escala con una media de 500 y desviación estándar de 100, para todos los países participantes en PISA 2006.

Como se señaló, los modelos de regresión ajustados para los perfiles escolares utilizan los ponderadores por escuela calculados por PISA. En el caso de los modelos jerárquicos lineales, en este estudio se usan los ponderadores por estudiantes normalizados, por ofrecer una estimación más adecuada del rendimiento promedio de Chile. Para todas las estimaciones se

consideran los cinco valores plausibles del rendimiento de los estudiantes, con los cuales se ajustan los modelos de regresión y se promedian sus coeficientes. Por su parte, en los modelos jerárquicos lineales, se consideran los valores plausibles, utilizando el procedimiento que ofrece el software HLM para estos casos.

2.3 Alcances y limitaciones del estudio

Como cualquier investigación, el presente estudio tiene limitaciones que deben considerarse al momento de ponderar sus hallazgos. En primer lugar, se trata de un estudio de carácter observacional, por lo que las relaciones aquí encontradas representan correlaciones entre variables y no necesariamente efectos causales. En segundo lugar, los indicadores construidos y utilizados para el análisis están subordinados a la información que entregan los cuestionarios PISA, lo que restringe la cantidad y calidad de éstos; por lo que no se desconoce la existencia de otros indicadores que podrían explicar el rendimiento de los estudiantes en Ciencias, independientemente de su nivel socioeconómico. Finalmente, se asume que los factores escolares son exógenos al modelo y no se considera el hecho de que variaciones en el rendimiento escolar podrían implicar cambios en los factores escolares. De no cumplirse este supuesto, las variables predictivas podrían tener algún grado de endogeneidad y, por consiguiente, los coeficientes de regresión podrían estar sobrestimados.

3. Resultados

En la presente sección se analizan los gradientes socioeconómicos y perfiles escolares, que estudian la relación entre el aprendizaje de los estudiantes y su nivel socioeconómico⁶. Como se señaló en la sección anterior, los perfiles relacionan el promedio de rendimiento con el promedio del nivel socioeconómico y cultural de los estudiantes en la escuela; mientras que los gradientes hacen lo propio para el rendimiento y el nivel socioeconómico por estudiante. Luego del análisis de gradientes y perfiles, se presenta el de los factores escolares asociados al rendimiento en Ciencias.

6 A lo largo del documento se usan los conceptos de nivel "socioeconómico y cultural", "sociocultural" y "socioeconómico". Todos estos términos se refieren al indicador de estatus socioeconómico y cultural calculado por PISA, y se usan como sinónimos para hacer más fluida la lectura.

3.1 Gradientes socioeconómicos

El análisis de gradientes para el conjunto de países latinoamericanos que rindió PISA 2006, indica que cerca de una quinta parte de la variación en el aprendizaje en Ciencias entre escuelas se explica por el índice socioeconómico y cultural de los estudiantes. En los casos de Chile, Colombia y Uruguay, se observa que el ISEC da cuenta de una cuarta parte de la variación en el logro académico entre escuelas (véase la Tabla 3)⁷.

La comparación del gradiente socioeconómico de Chile con el de los países de América Latina, muestra que el promedio de aprendizaje en Ciencias en Chile, después de descontar el efecto del ISEC del estudiante, es mayor a la media del resto de los países de América Latina. La excepción es Uruguay, cuyo promedio de logro, independientemente del ISEC del alumno, es similar al de Chile.

Las pendientes de los gradientes reflejan el cambio que puede ocurrir en el rendimiento de los estudiantes al cambiar el ISEC. Al comparar la pendiente del gradiente de Chile con las pendientes del resto de los países, se observa que la pendiente en Chile es distinta⁸.

Los resultados de la Tabla 3 muestran que en Uruguay, el ISEC del alumno tiene una incidencia más pronunciada sobre el aprendizaje que en el resto de los países; ello, por cuanto un aumento de una unidad en el ISEC llevaría a incrementar el logro en

7 Cabe resaltar que los modelos individuales por país muestran diferencias significativas con el modelo para América Latina en la magnitud de la constante y las pendientes. La única excepción es la constante del modelo de Colombia, la cual no es significativamente distinta a la del modelo regional. Para mayor información respecto de la comparación entre modelos, véase el Anexo.

8 La contrastación entre países se realiza a través de modelos lineales jerárquicos en los que se analizan los datos de dos países simultáneamente y se incluye el ISEC por estudiante, una variable dicotómica que toma el valor 1 para los estudiantes chilenos y cero para los del país de comparación. También se introduce una interacción entre el ISEC y la variable identificatoria del país. De esta forma, se pretende conocer si hay diferencias significativas en la relación del ISEC con el aprendizaje en Ciencias, en los dos países. En el Anexo se incluyen los resultados de los contrastes.

Tabla 3.
Gradientes socioeconómicos para los seis países de América Latina participantes en PISA 2006

Coefficientes y varianzas	América Latina	Argentina	Brasil	Chile	Colombia	México	Uruguay
Constante	388,9*	367,7*	382,7*	422,0*	385,9*	401,7*	428,1*
ISEC Alumno	13,1*	16,9*	13,7*	15,5*	16,3*	9,8*	20,7*
% Varianza explicada nivel 1 (estudiantes)	1,4%	2,5%	1,1%	1,4%	2,3%	0,8%	3,1%
% Varianza explicada nivel 2 (escuela)	19,7%	20,1%	21,1%	26,9%	26,2%	19,6%	27,7%

(*) Los coeficientes son significativos al 0,1%

20,7 puntos. De acuerdo con las pruebas estadísticas realizadas, existen diferencias significativas entre las pendientes y constantes de los modelos de todos los países, y las de Chile.

Otro dato que se observa, es que los gradientes socioeconómicos por país son significativamente distintos al gradiente regional. Incluso los gradientes de Brasil, México y Argentina —que tienen mayor influencia en los resultados globales, dado el tamaño de su población— son distintos al regional. A partir de este resultado se puede deducir que el modelo regional ofrece una visión general que no necesariamente refleja características específicas de la relación entre el aprendizaje en Ciencias y el ISEC de cada uno de los países.

Las desigualdades socioeconómicas y culturales se relacionan fuertemente con el aprendizaje logrado, y las disparidades sociales a nivel individual explican solo una pequeña porción de la variación de los aprendizajes. De hecho, llama la atención que entre una quinta y una cuarta parte de la variación en los resultados de aprendizaje por escuela, se explique por el nivel socioeconómico y cultural individual de los estudiantes⁹.

9 Aunque este alto porcentaje podría estar sobrestimado por la especificación del modelo que omite algunas variables con el fin de estudiar la asociación entre el ISEC y el logro.

Para profundizar este análisis de la relación entre ISEC y aprendizaje en Chile, en la Tabla 4, se presenta un modelo que separa el efecto que tienen sobre el aprendizaje el ISEC individual y el ISEC promedio de la escuela. Este modelo se denomina “efectos composicionales”¹⁰. Los resultados muestran que el aprendizaje en Ciencias se relaciona con el ISEC, tanto a nivel del estudiante, como del promedio del establecimiento en Chile. La incidencia del ISEC promedio de la escuela sobre el aprendizaje es apreciable, pues un estudiante de un establecimiento cuyo ISEC está una desviación estándar por debajo del promedio de Chile, obtiene 58 puntos menos que los estudiantes de centros educativos con ISEC igual al promedio nacional. Asimismo, la inclusión del ISEC promedio por escuela lleva a que se explique 75% de la variación en el rendimiento académico entre escuelas, mientras que el ISEC del nivel estudiante explica una baja proporción del logro.

10 Los efectos composicionales se refieren a descomponer el efecto individual y grupal de un predictor sobre la variable dependiente. Para esto, se incluye la variable predictiva en los distintos niveles de anidación de los datos. Para el caso de este trabajo, se analizan los efectos composicionales del ISEC en el aprendizaje, distinguiendo entre el ISEC por alumno y el promedio por escuela.

Tabla 4.
Efectos composicionales del ISEC sobre el logro en Chile

Efectos Composicionales Chile	Coefficientes
Constante	426,8*
ISEC Establecimiento	57,9*
ISEC Alumno	12,4*
Varianza explicada nivel 1 (estudiantes) ¹¹	1,1%
Varianza explicada nivel 2 (escuela) ¹²	74,9%

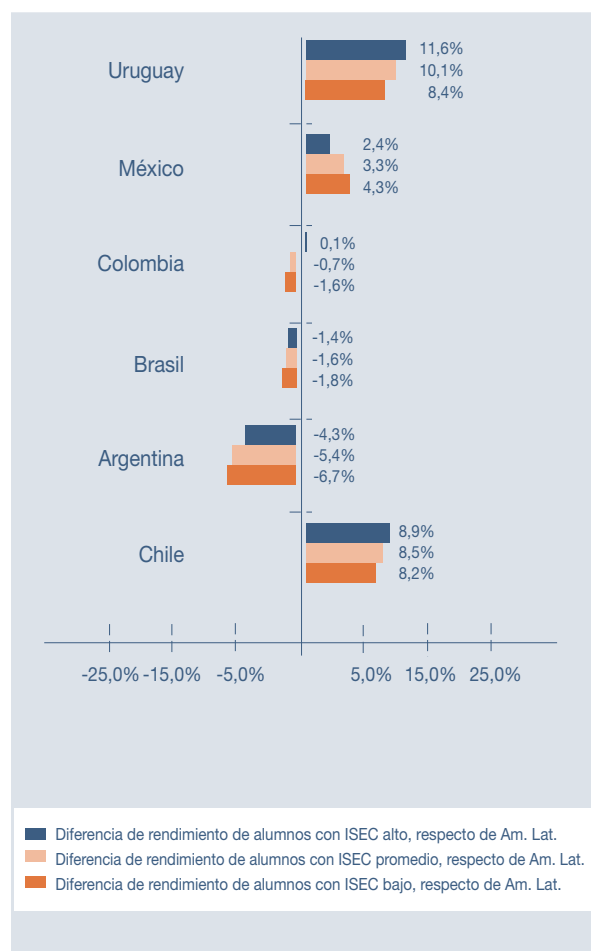
(*) Los coeficientes son significativos al 0,1%

El comportamiento de los efectos composicionales que se observa en la Tabla 4, sugiere que en Chile, la segregación escolar — vista a través del ISEC por establecimiento— se relaciona con el rendimiento. Una forma de ver la segregación escolar es a través del análisis de la varianza del ISEC entre establecimientos y al interior de ellos. Al llevar a cabo este cálculo se observa que, del total de la variación en el ISEC, 53% se da entre escuelas, y 47%, entre estudiantes al interior de la escuela. Estas cifras son similares al análisis de la varianza en el aprendizaje de Ciencias: de acuerdo al *modelo nulo* que se presentará más adelante, cerca de 51% de la variación total en el logro ocurre entre escuelas, y 49% entre estudiantes, al interior de la escuela. Por ello, se puede afirmar que la distribución del aprendizaje y el nivel socioeconómico siguen patrones similares, en los que las desigualdades sociales se traducen en distintos logros de aprendizaje.

A modo de síntesis, en el Gráfico 1, se comparan las diferencias porcentuales en el rendimiento de los

estudiantes de cada país, con respecto al promedio de América Latina. Para facilitar la visualización de la información, se establecen tres estratos del nivel socioeconómico y cultural: ISEC promedio, ISEC alto (entendido como una desviación estándar por sobre el promedio) e ISEC bajo (entendido como una desviación estándar bajo el promedio).

Gráfico 1: Diferenciales de rendimiento de los estudiantes de cada país respecto de los estudiantes de América Latina, controlado por ISEC



Del gráfico, se puede concluir que en Chile, en todos los estratos, los puntajes obtenidos por los estudiantes, superan al contexto regional. En efecto, un alumno chileno con ISEC promedio, obtiene un puntaje mayor en 8,5%, al de un alumno de similares condiciones perteneciente al contexto regional. Una

11 La varianza explicada para el nivel 1, se obtiene mediante la siguiente fórmula: Varianza explicada nivel 1 = 1- (varianza del nivel 1 en el modelo de gradiente (variance component en HLM) / varianza del nivel 1 en el modelo nulo).

12 La varianza explicada para el nivel 2 se obtiene mediante la siguiente fórmula: Varianza explicada nivel 2 = 1- (varianza del nivel 2 en el modelo de gradiente/ varianza del nivel 2 en el modelo nulo).

interpretación similar se puede hacer con los otros dos estratos. En este análisis, llama la atención el caso de Argentina, ya que es el país más rezagado con respecto al rendimiento de América Latina en los tres estratos. En ese país, un alumno con ISEC promedio, obtiene un puntaje inferior en 5,4%, al de un alumno de similares condiciones, perteneciente al contexto regional. Uruguay por su parte, es el país que mayores diferencias positivas presenta en relación al contexto regional para los tres estratos: un alumno con ISEC promedio en Uruguay, obtiene un puntaje mayor en 11,6% al de uno de similares condiciones en el contexto regional. Colombia se comporta de manera muy similar al contexto regional en sus tres estratos, al igual que Brasil. México, por otro lado, tiene pequeñas diferencias respecto de América Latina, siendo estas positivas para los tres estratos.

En resumen, el análisis de los gradientes socioeconómicos, indica que la incidencia del ISEC del alumno en el aprendizaje de los estudiantes en Chile está en un rango alto en comparación con el resto de los países de América Latina, con la excepción de Uruguay. Específicamente, los estudiantes chilenos obtienen resultados por encima del promedio regional y levemente menores a los que alcanzan los estudiantes en Uruguay. Esto indicaría que las desigualdades en Chile tienen incidencia en el logro. Aún así, los alumnos de ISEC alto, medio y bajo en Chile, obtienen resultados mucho mayores a los que obtendría un alumno representativo de América Latina. Esto se



debe a que el promedio de rendimiento de los estudiantes chilenos (representado por el intercepto de la ecuación de gradientes socioeconómicos, y después de controlar con el nivel socioeconómico y cultural de los estudiantes), es el segundo mayor entre los países estudiados, solo bajo el de Uruguay.

3.2 Perfiles escolares

Como ya se señaló, los perfiles escolares son otra forma de estudiar la relación entre rendimiento y nivel socioeconómico y cultural. En efecto, a partir de sus resultados se puede describir la distribución de las escuelas, de acuerdo a la relación entre rendimiento académico e ISEC promedio por escuela. Los hallazgos de este análisis sirven para retroalimentar políticas educativas, pues permiten establecer un mapa de la eficacia escolar¹³.

Del análisis, se observa que el promedio de ISEC por escuela es altamente explicativo del aprendizaje, lo que confirma que la segregación socioeconómica entre las escuelas se traduce en desigualdades de aprendizaje.

En particular, se puede observar que Chile tiene el perfil escolar con mayor fuerza y pendiente de América Latina. Esto supone que es el país en que existe mayor incidencia del nivel socioeconómico y cultural sobre el aprendizaje promedio de los estudiantes, en una misma escuela. En concreto, el ISEC por establecimiento en Chile explica 60% de la variación en el aprendizaje promedio de la escuela (como se puede ver en el indicador R2 de la Tabla 5).

¹³ Los perfiles escolares se calcularon utilizando el método de regresión lineal del software SPSS para cada uno de los valores plausibles promedio por escuela. Finalmente se promediaron los valores de los coeficientes de las cinco regresiones, para obtener un coeficiente promedio, siguiendo las indicaciones de la literatura de imputación múltiple y de análisis con valores plausibles.

Tabla 5.
Perfiles escolares para América Latina y los países de la Región, participantes en PISA 2006

Perfil escolar	América Latina	Argentina	Brasil	Chile	Colombia	México	Uruguay
Constante	386,9*	352,0*	385,8*	400,9*	389,5*	401,1*	398,5*
ISEC establecimiento	35,2*	43,3*	41,4*	57,5*	32,9*	28,7*	47,7*
R2	0,4	0,4	0,5	0,6	0,5	0,4	0,5

(*) Los coeficientes son significativos al 0,1%

Tal como muestra el perfil escolar de América Latina (Figura 1), existe un conjunto reducido de escuelas con bajo nivel socioeconómico promedio y exiguo rendimiento en Ciencias¹⁴. Al revisar los perfiles de los países, se constata que estas escuelas se encuentran en Argentina, Brasil, Colombia y México. Los establecimientos de bajo ISEC y bajo rendimiento, contrastan con otro conjunto de centros educativos que, presentando ISEC menores a -2, obtienen resultados en Ciencias que fluctúan entre los 300 y los 400 puntos. De ello se desprende que existen escuelas que son más eficaces para promover el aprendizaje entre estudiantes de nivel socioeconómico bajo, lo que motiva a estudiar las características que les permiten superar las expectativas de aprendizaje determinadas por la condición social de sus estudiantes.

En el perfil escolar de Chile, se advierte una pendiente más pronunciada que en el resto de los países y que en la Región; lo que sugiere una alta correlación entre el ISEC promedio del establecimiento y los resultados de aprendizaje en Ciencias. Por otro lado, algunas escuelas chilenas con ISEC igual o menor a cero, registran puntuaciones promedio ostensiblemente más bajas que otros establecimientos con el mismo nivel socioeconómico.

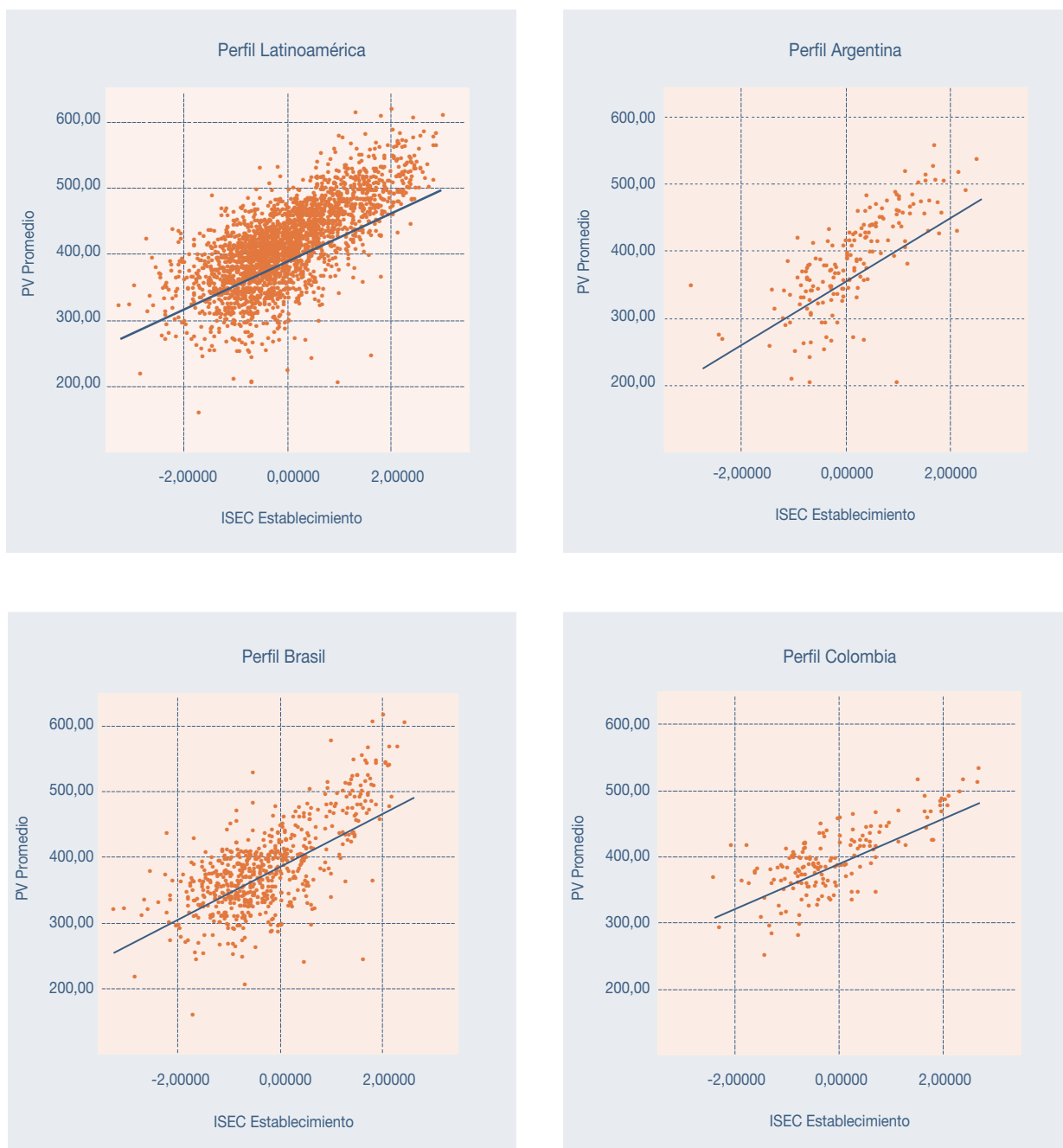
En general, los perfiles muestran una mayor variación de resultados entre las escuelas, cuando éstas tienen un ISEC promedio menor a cero. A medida que aumenta el ISEC medio de la escuela, la distribución

de resultados se va haciendo más homogénea, en torno a la línea de regresión. Esto sugiere que existe una mayor posibilidad de caer en círculos viciosos de bajo logro, al ubicarse por debajo del ISEC promedio. En contraste, a medida que aumenta el ISEC promedio de la escuela, disminuyen las posibilidades de que éstas tengan logros muy lejanos a los esperados. El patrón descrito en los perfiles se manifiesta en Argentina, Chile, Colombia y Uruguay.

Si se compara el rendimiento promedio de las escuelas de cada país, según el rango de ISEC en el que se encuentran, se observa que las escuelas de Chile con ISEC más bajo, obtienen rendimientos ostensiblemente menores al promedio de la Región. Esto contrasta con lo observado en los resultados de gradientes socioeconómicos, donde los estudiantes chilenos alcanzan aprendizajes mayores a los de la Región. Sin embargo, al agregar el ISEC y las puntuaciones en Ciencias por escuela, se confirma lo visto en los efectos composicionales para Chile. Es decir, que existe una alta correlación entre el ISEC promedio de la escuela y los resultados de aprendizaje (Gráfico 2). De lo anterior, se puede concluir que la segregación socioeconómica de la escuela (o el “efecto par”), tiene una incidencia relativamente mayor sobre el aprendizaje, que la que origina el nivel socioeconómico individual de los estudiantes. Las escuelas de Chile con ISEC medio y alto, por el contrario, alcanzan niveles de rendimiento superiores a la media de la región. Entre los países, llama la atención el caso de Argentina, donde en los tres rangos del ISEC, las escuelas ostentan un rendimiento promedio menor al de la Región.

¹⁴ Se trata de las escuelas representadas por los puntos cuyo ISEC promedio es menor a -2 y cuya puntuación es menor a 300 puntos.

Figura 1.
Perfiles escolares para América Latina y los países de la Región, participantes en PISA 2006



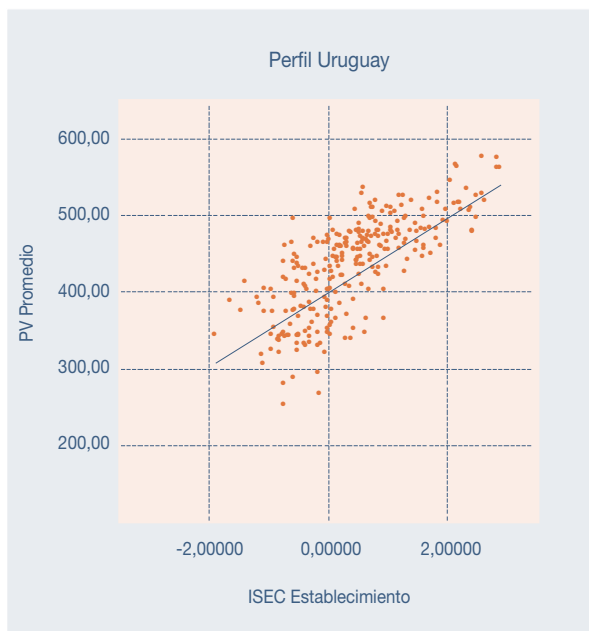
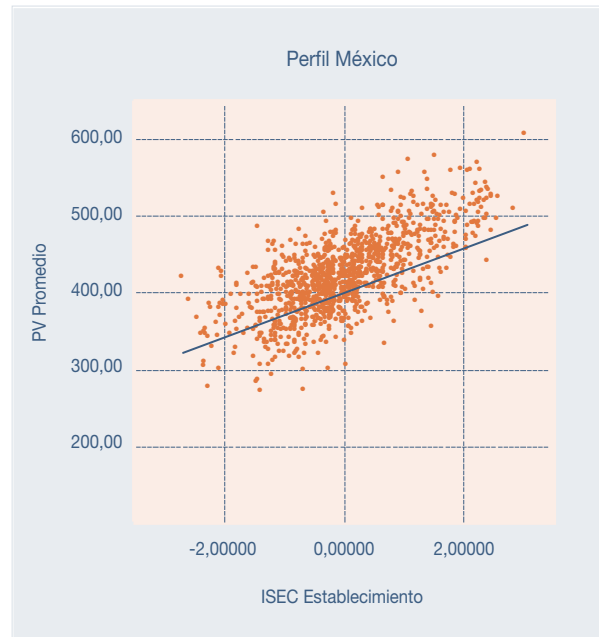
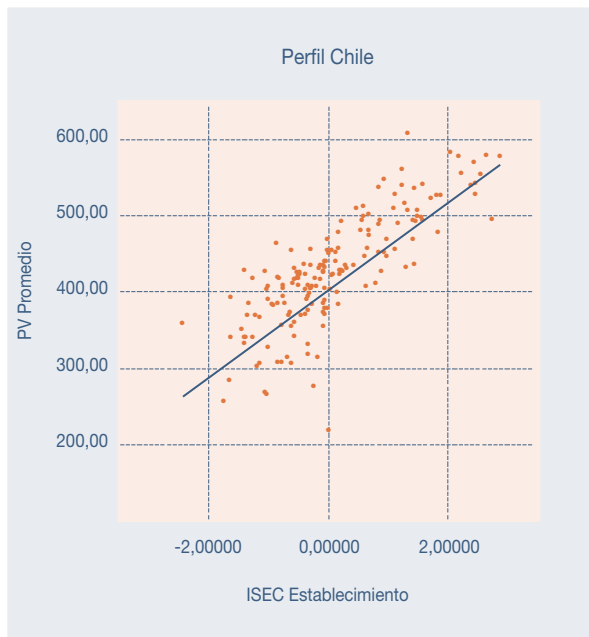
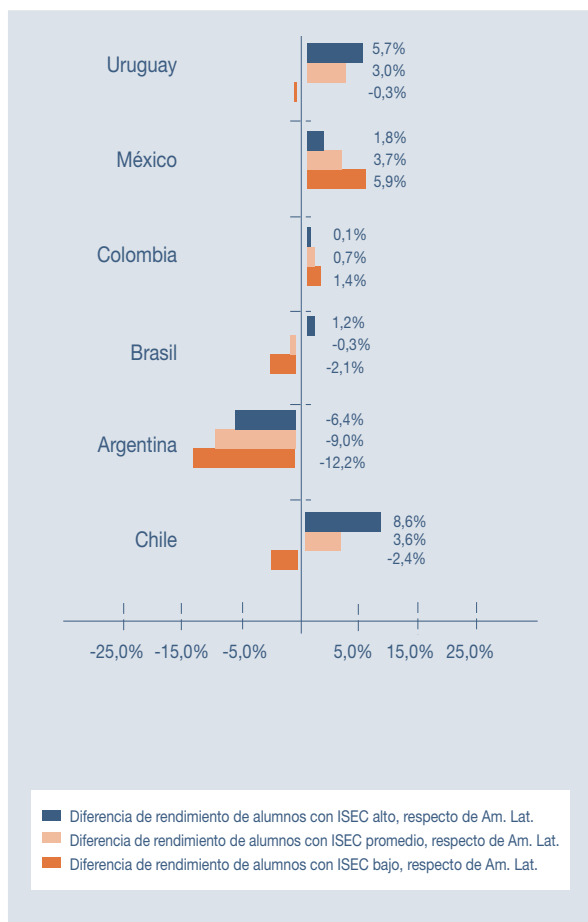


Gráfico 2: Diferenciales de rendimiento promedio de las escuelas de cada país con respecto al promedio de América Latina, controlado por ISEC del establecimiento



En síntesis, de la evidencia antes expuesta, se puede afirmar que los aprendizajes están fuertemente mediados por la segregación socioeconómica y cultural de los estudiantes, en los distintos tipos de escuela. Asimismo, si bien la relación entre desigualdades sociales y aprendizaje sigue patrones similares en América Latina, destacan las escuelas chilenas de ISEC bajo, en las que se obtienen puntuaciones menores al promedio de la región. Este hallazgo parecería contradictorio con lo que se encontró en los gradientes (donde los estudiantes chilenos de todos los estratos de ISEC, obtenían resultados por encima del promedio de la Región); sin embargo, no existe tal contradicción: el

problema radica en que los perfiles, al considerar el ISEC promedio por escuela, dan cuenta de la influencia que tiene sobre el rendimiento la segregación socioeconómica de las escuelas.

Ahora bien, aun cuando las diferencias socioeconómicas marcan los logros de aprendizaje de los estudiantes y reproducen las disparidades sociales, el hecho de que existan escuelas donde los estudiantes obtienen logros que superan las expectativas impuestas por las condiciones sociales, releva la importancia de indagar en los factores que determinan estas diferencias, pues la adopción de medidas exitosas, por parte de escuelas con bajos rendimientos, podrían contribuir a aminorar la estrecha relación existente entre las desigualdades sociales y el aprendizaje.

3.3 Análisis de factores asociados al aprendizaje

En este estudio, como se señaló, el análisis de modelos lineales jerárquicos está orientado a determinar los factores escolares que potencialmente podrían mejorar el aprendizaje de los estudiantes chilenos en Ciencias, independientemente de su nivel socioeconómico. Para ello, como se muestra en la Tabla 6, se construyó una taxonomía de trece modelos de regresión multinivel que prueban: variables sociales y económicas, factores escolares (de acuerdo a las respuestas del director) y factores relativos a la calidad de los docentes y a la afinidad de los estudiantes con las ciencias.

A través del *modelo nulo*, se puede observar que 51% de la variación de los aprendizajes se debe a diferencias entre escuelas, lo que se interpreta como el “efecto escolar bruto”, esto es, el potencial de la escuela para elevar los resultados de aprendizaje, sin considerar la influencia del nivel socioeconómico de los estudiantes.

Tabla 6.

Taxonomía de modelos lineales jerárquicos para estimar, a partir de los datos PISA 2006 obtenidos por Chile, la relación entre el aprendizaje de los estudiantes en Ciencias, el nivel socioeconómico y cultural de los mismos y los factores escolares

	Modelo Nulo	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6	Modelo 7	Modelo 8	Modelo 9	Modelo 10	Modelo 11	Modelo 12	Modelo Final (13)
Nivel escuela													
Escuela subvencionada				16,48	16,93	15,53	32,23	28,89	29,02	29,05	29,07	29,16	29,48
Escuela particular pagada				17,35	17,42	22,17	33,69	27,30	27,39	27,43	27,44	27,54	27,87
Rendición de cuentas						11,40	9,41	9,60	9,65	9,65	9,65	9,68	9,82
Profesores calificados							17,89	17,09	17,15	17,16	17,16	17,20	17,31
Autonomía de decisión respecto a estudiantes								6,35	6,40	6,41	6,42	6,45	6,55
ISEC promedio escuela			57,95	53,76	53,78	47,60	49,32	49,78	49,75	49,75	49,74	49,71	49,60
Nivel estudiantes													
Niña (género)					-16,47	-16,53	-16,59	-16,59	-17,20	-17,22	-17,35	-18,04	-22,38
Afinidad con las ciencias									11,02	11,13	11,24	11,33	11,63
Profesores incentivan aplicación práctica de las ciencias										4,06	4,08	4,07	4,60
Proyección de los estudiantes en temas científicos											3,53	3,68	3,89
Profesores hacen partícipes a los estudiantes												8,74	9,12
Oportunidades de investigación científica en la escuela													15,34
ISEC		15,49	12,42	12,42	11,52	11,51	11,51	11,51	10,91	10,76	10,48	10,60	11,07
Intercepto	418,07	422,03	426,82	418,13	425,41	426,07	417,43	419,22	419,40	419,39	419,43	419,69	421,43
Varianza Explicada Modelo													
Nivel 2	0,51	0,27	0,75	0,75	0,76	0,77	0,80	0,80	0,80	0,80	0,80	0,80	0,79
Nivel 1	0,49	0,01	0,00	0,00	0,01	0,01	0,01	0,01	0,03	0,04	0,04	0,06	0,10

Significativa al 5%

Significativa al 10%

El *modelo 2*, por su parte, logra explicar 27% de la variación en el aprendizaje entre escuelas y 1,4% de la variación entre estudiantes. Este hallazgo sugiere que el ISEC puede explicar mejor la diferencia de logro entre escuelas, que entre estudiantes.

El *modelo 3*, muestra que el ISEC por establecimiento tiene un efecto potente sobre la puntuación en Ciencias, con un coeficiente de 58 puntos, y que explica tres cuartas partes de la variación del aprendizaje entre escuelas.

El *modelo 4* sugiere que, controlando por el ISEC del alumno y establecimiento, no existen diferencias significativas entre el aprendizaje de los estudiantes de las escuelas municipales y el de las escuelas particulares pagadas. Esto podría deberse a la segregación socioeconómica de los estudiantes en las escuelas, dado que, como se sabe, la dependencia de la escuela está fuertemente asociada con las características socioculturales de los estudiantes¹⁵. Por otro lado, los estudiantes de escuelas particulares subvencionadas aventajan en 16 puntos a los de las municipales, y a partir del *modelo 7*, que controla por factores escolares, se observa que existen diferencias significativas en el rendimiento, que favorecen a los establecimientos particulares subvencionados y pagados, por sobre los municipales.

El *modelo 5*, que incorpora la variable de género, revela que las mujeres, aun después de controlar la variable nivel socioeconómico individual y de la escuela, obtienen puntuaciones promedio 16,5 puntos por debajo de los varones. Estas desigualdades de

género, se mantienen significativas en el resto de la taxonomía, y se acrecientan a medida que se incorporan nuevos predictores. De allí que se pueda admitir que, sin duda, este es uno de los temas más importantes a tratar por la investigación educacional en Chile, dada la complejidad para acumular evidencia empírica que ayude a explicar con certeza este fenómeno.

A partir del *modelo 6*, la taxonomía progresivamente incorpora las variables escolares de rendición de cuentas a los padres, dotación de profesores calificados y autonomía para tomar decisiones. Estas incorporaciones, llevan a que se reduzca la magnitud del coeficiente del ISEC por establecimiento, de 58 a 49 puntos. En otras palabras, los estudiantes que logran mayores niveles de aprendizaje, asisten a establecimientos que:

- a. se responsabilizan por los resultados,
- b. rinden cuentas de los resultados a los apoderados y
- c. cuentan con procedimientos para contrastar los niveles de aprendizaje obtenidos con las expectativas de los padres.

Entre estas variables incorporadas, la autonomía del establecimiento en el ámbito pedagógico, es la que presenta la asociación más débil con el rendimiento (nivel de significancia de 10%).



¹⁵ Por ello, es probable que exista algún grado de multicolinealidad en el modelo, es decir, correlación entre dos variables predictoras. Esto provoca que las varianzas y covarianzas de las estimaciones sean más grandes de lo que serían, en ausencia de multicolinealidad, y lleva a que algunos coeficientes resulten no significativos para predecir el logro, aunque su magnitud sea elevada.

También a partir del *modelo 6*, se puede observar que un incremento de una unidad en el ISEC por establecimiento trae un ascenso de 50 puntos en el rendimiento académico de los estudiantes. Esto sucede aun después de considerar los factores escolares. Además, los resultados de los modelos 8 al 13 indican que las diferencias de logro entre las escuelas municipales y las particulares pagadas son significativas solamente al 10%, lo que implica que una de cada diez veces, esta afirmación no será verdadera.

En otro aspecto y, de manera similar a lo que se señaló en la revisión de la literatura, se puede establecer que la calidad de las escuelas está determinada por la calidad de sus docentes. Los establecimientos que cuentan con profesores calificados en todas las áreas del currículum, promueven niveles de aprendizaje más altos entre los estudiantes. Esta constatación se condice con las características de las prácticas docentes que se presentan a continuación, donde se observa que los profesores más efectivos son aquellos que estimulan un aprendizaje aplicado y de responsabilidad compartida con los estudiantes.

En efecto, en cuanto a las prácticas pedagógicas, los resultados muestran que, cuando un profesor promueve la participación en el aprendizaje (con un índice superior en una desviación estándar a la media), los estudiantes obtienen hasta 9 puntos más, en promedio, que los que tienen profesores que se ubican en el promedio de este indicador. Por otra parte, los profesores que incentivan la aplicación práctica de las ciencias (una desviación estándar por encima del promedio en este indicador), consiguen que los estudiantes alcancen en promedio, 4 puntos más en la prueba. Estos resultados son consistentes y estables, tal como se observa en los modelos 9 al 13.

Las oportunidades de aprendizaje de Ciencias son otro factor que incide en el logro: una desviación estándar por encima del promedio en el indicador

de oportunidades de investigación científica en la escuela, implica que los estudiantes alcancen 15 puntos adicionales en la prueba de Ciencias. El índice de oportunidades de aprendizaje da cuenta de las actividades que realizan los establecimientos para promover el compromiso con las ciencias, y de la intención explícita de los docentes por desarrollar en sus estudiantes habilidades y conocimientos que les permitan acceder a carreras científicas en la Educación Superior.

Por su parte, el interés o motivación intrínseca, también resulta ser un factor explicativo: los estudiantes con proclividad hacia las ciencias obtienen resultados más altos que los estudiantes con menos inclinaciones hacia esta área del conocimiento. Como se observa en la Tabla 6, cuanto mayor es la afinidad de los estudiantes hacia las ciencias y su proyección para incursionar en carreras científicas, mayores son sus resultados de aprendizaje. En total, la contribución de estas variables (índices compuestos) implica que los estudiantes cuya puntuación en la afinidad con las ciencias y la proyección en temas científicos está una desviación estándar por encima de la media, obtienen puntuaciones 15 puntos mayores que aquellos con valores promedio en estas dos variables predictivas.

3.4 Síntesis de los hallazgos del estudio

Analizado el comportamiento de las variables independientes en distintas especificaciones de modelos predictores del logro académico en Ciencias, a continuación se presentan los resultados del *modelo final* (13), que sintetizan el efecto de los factores escolares, sobre el aprendizaje de los estudiantes en Ciencias, independientemente del nivel socioeconómico y cultural de los estudiantes y del establecimiento.

Los resultados del modelo final, ilustrados en el Gráfico 3, indican que el índice socioeconómico y cultural promedio de la escuela tiene una alta incidencia

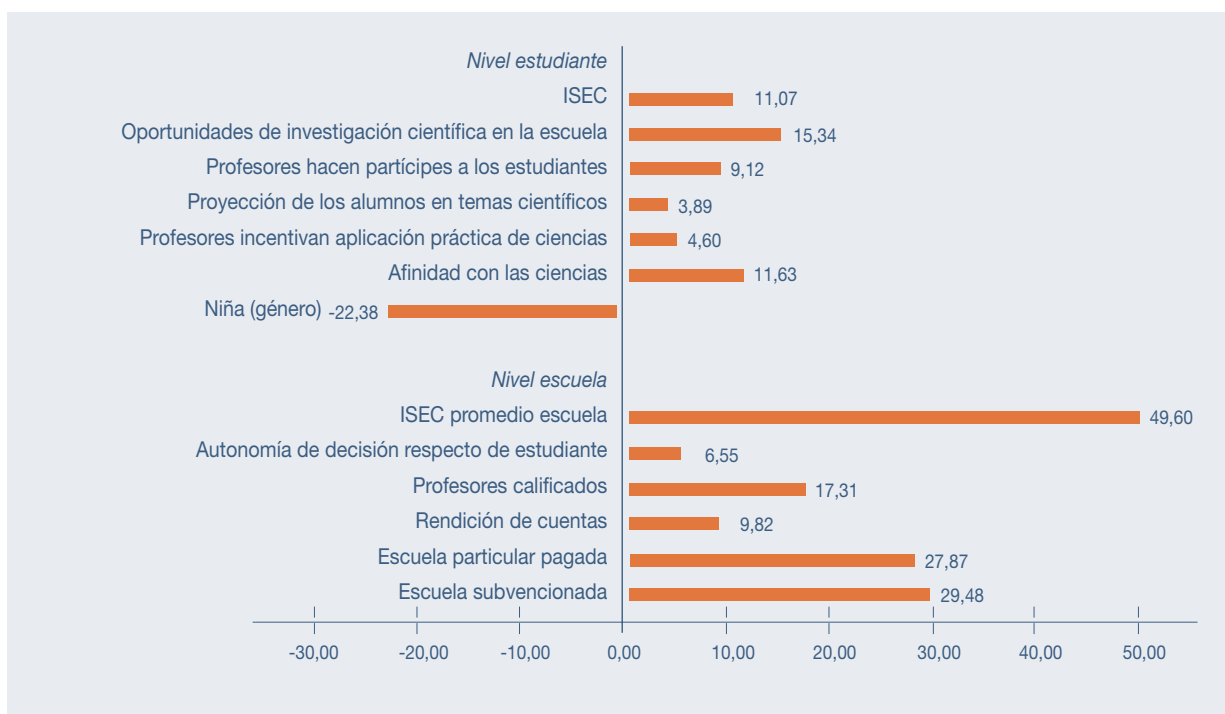
en el aprendizaje. El coeficiente del ISEC promedio de la escuela alcanza cerca de 50 puntos, lo que implica que por una desviación estándar de incremento en el ISEC promedio de la escuela, el rendimiento de los estudiantes aumenta en 50 puntos.

Comparando el coeficiente del ISEC promedio de la escuela en el *modelo final*, con el *modelo 3* de efectos composicionales, se observa que al incluir los factores escolares, se reduce en cerca de 8 puntos el coeficiente asociado al ISEC de la escuela sobre el rendimiento. Por otro lado, el coeficiente del ISEC de los estudiantes mantiene una magnitud de entre 11 y 12 en todos los modelos.

Si bien la inclusión de los factores escolares aumenta el porcentaje de varianza explicada por el establecimiento

(nivel 2) y disminuye el coeficiente asociado a las características socioeconómicas y culturales del estudiante y del establecimiento, este efecto es limitado. Sin embargo, existen características de las escuelas y los profesores que representan promisorias vías para mejorar la educación en Ciencias de los estudiantes menos aventajados del sistema. En efecto, tanto la dotación de profesores calificados en la escuela, como las prácticas docentes que hacen a los estudiantes partícipes del aprendizaje y permiten la aplicación práctica de las ciencias, son variables relativas a los docentes que pueden marcar la diferencia en el aprendizaje de los estudiantes, ya que un aumento en una desviación estándar en los tres índices mencionados, traería consigo un incremento promedio del aprendizaje de los estudiantes, de 31 puntos.

Gráfico 3: Valor de los coeficientes de los factores asociados al aprendizaje de las ciencias en Chile, en PISA 2006, para el nivel de estudiantes y de escuela



Por su parte, la estructuración de oportunidades de aprendizaje, también depende de las medidas que tome la escuela. De esta forma, la rendición de cuentas a los apoderados, el ofrecer oportunidades de investigación científica en la escuela y la autonomía pedagógica, son elementos que favorecen el aprendizaje de los estudiantes.

De acuerdo con los resultados del *modelo final*, los estudiantes de escuelas particulares subvencionadas y pagadas, aventajan a los de las escuelas municipales por cerca de 30 puntos. Esto, después de descontar el efecto de todos los factores incluidos en el modelo. Sin embargo, no existen diferencias significativas en el logro entre las escuelas particulares pagadas y las subvencionadas.

Cabe señalar que algunas características de los estudiantes también predicen el rendimiento académico en Ciencias. Así, las mujeres tienen un desempeño 20 puntos menor que los varones, una vez controlado el resto de las variables. Del mismo modo, rasgos de los estudiantes como la afinidad con las ciencias y la proyección de una carrera futura en el campo científico, son características individuales de los estudiantes que influyen positivamente en el rendimiento académico.

4. Conclusiones

La distribución del conocimiento en Ciencias, entre los jóvenes de 15 años en Chile, replica la estructura de desigualdades sociales. En efecto, los estudiantes asisten a escuelas con compañeros de su misma condición socioeconómica, por razones como la selección por capacidad de pago, las características familiares y el rendimiento pasado o potencial. Cualquiera sea el motivo, es un hecho que el sistema escolar tiende a agrupar en un mismo recinto educativo a estudiantes de clase social similar. Esta segregación escolar se transforma en segregación de aprendizaje, puesto

que las desigualdades en el nivel socioeconómico promedio de los estudiantes en cada escuela, se asemejan a la distribución de los aprendizajes promedio por establecimiento¹⁶. A pesar de esta situación, que podría interpretarse como determinismo social, actualmente en Chile, 70% de los estudiantes de Educación Superior son primera generación de sus familias en acceder a ese nivel educativo. De aquí se deduce que los estudiantes y sus familias hacen un gran esfuerzo por superar las barreras de aprendizaje que establece el sistema escolar y, de alguna forma, un porcentaje no menor de jóvenes lo logra.

La estrecha relación entre el nivel socioeconómico promedio de los estudiantes en la escuela y el aprendizaje, sugiere otorgar un papel fundamental a esta variable en materia de política pública. Por ejemplo, un buen ejercicio implicaría reevaluar la contribución económica que hace la Ley SEP por concepto de concentración de estudiantes prioritarios en el recinto. A reserva de realizar estudios en profundidad en esta materia, aparentemente sería adecuado ofrecer un mayor monto de subvención escolar preferencial asociado al porcentaje de estudiantes prioritarios en la escuela.

Las mujeres obtienen puntuaciones menores a los hombres en la prueba de Ciencias en Chile, de acuerdo con los resultados de PISA y otros estudios internacionales (UNESCO, 2008). Las disparidades de género en Chile son las mayores entre los países de América Latina, situación que llama la atención. Si bien no estaba en el centro de este análisis desentrañar las posibles causas de las desigualdades de género en el aprendizaje de las ciencias, es necesario profundizar

16 Podría ser que este factor recoja, en alguna medida, rasgos como la selectividad de las escuelas en los procesos de admisión, ya que ésta es una de las variables que compone el índice de autonomía escolar. Por ello, se recomienda tomar con cautela este hallazgo, pues la selección en la admisión de estudiantes suele segregar socialmente y sus resultados de aprendizaje están condicionados por las características socioculturales de los estudiantes más que por su eficacia en los procesos escolares y pedagógicos.

en la comprensión de este fenómeno que representa uno de los principales desafíos para la investigación en educación y la política educativa.

El propósito de este estudio ha sido encontrar las variables que la escuela podría modificar para mejorar el aprendizaje de las ciencias, y así facilitar el tránsito de los estudiantes hacia niveles más altos de escolaridad y aprendizajes más profundos y auténticos en el ámbito científico. En este sentido, los resultados de los análisis realizados indican que los factores escolares pueden mejorar el aprendizaje en Ciencias, aunque los cambios que la escuela y los profesores puedan implementar no sean suficientes para contrarrestar en su totalidad el peso que ejercen las desigualdades socioeconómicas sobre el aprendizaje. Por tanto, para mejorar el aprendizaje desde la escuela, se requieren no sólo políticas educativas que disminuyan la segregación escolar y mejoren la formación y las prácticas docentes, sino también, políticas sociales vigorosas que permitan reducir las diferencias de origen de los estudiantes.

Las acciones que las escuelas deben tomar para mejorar los aprendizajes, corresponden a los ámbitos de la gestión y de la pedagogía.

Desde la gestión, los establecimientos serán más exitosos en la medida en que se responsabilicen por los resultados y den cuenta a los apoderados de los logros y dificultades de los estudiantes. Asumir la responsabilidad por los resultados, no debe entenderse como una cacería de culpables, sino más bien como una reflexión seria, basada en la evidencia y orientada a comprender el efecto que han tenido las acciones de los profesores y de la escuela sobre el aprendizaje de los estudiantes. Solo así, se puede valorar la eficacia de las medidas adoptadas por la escuela, planteándose posibles acciones correctivas en aquellos ámbitos donde los resultados sean menores que los esperados. Por otro lado, la autonomía escolar puede ayudar a



mejorar los aprendizajes, dando espacios de decisión a los establecimientos en cuanto a la evaluación, admisión y estructuración del currículum; sin embargo, la promoción de autonomía debe contemplar que no todas las escuelas tienen la misma preparación para asumirla, por lo que no se debe considerar esta recomendación como una “receta” genérica para mejorar el logro.

En el ámbito de las acciones pedagógicas, los profesores pueden aportar a mejorar el aprendizaje en Ciencias de los estudiantes, independientemente del nivel socioeconómico de estos últimos. Para conseguir esto se requiere de docentes calificados, que inviten a los estudiantes a hacerse partícipes del aprendizaje, mediante la motivación y la vinculación de la ciencia con eventos de la vida cotidiana que les sean significativos. La labor de los docentes consiste en promover en los estudiantes el hábito de pensar el mundo desde la perspectiva científica, dando espacio para que prueben sus propias ideas y vean las eventuales equivocaciones como una oportunidad

para aprender. Desde esta perspectiva, el énfasis debería estar más en el desarrollo de habilidades, que en la cobertura de una gran cantidad de contenidos curriculares.

Contar con excelentes profesores que conozcan la disciplina y la forma de enseñarla, resulta especialmente crítico en escuelas de menor nivel socioeconómico, pues podría ser una medida efectiva que logre mejorar los resultados, más allá del nivel socioeconómico de la escuela y del estudiante. Ahora bien, un trabajo docente fructífero en términos de aprendizaje, está regularmente acompañado de un entorno escolar que apoya el desarrollo de habilidades científicas. Esto se logra a través de acciones deliberadas de las escuelas por generar oportunidades sistemáticas para llevar a cabo investigaciones científicas. Nuevamente, esto sugiere que el fortalecimiento de competencias científicas debe ser parte de un proyecto escolar global para diseñar cuidadosamente la enseñanza de los estudiantes, de forma tal que se puedan aprovechar todas las instancias de aprendizaje, transformar la visión de los estudiantes y ayudarlos a pensar sistemáticamente con la lógica de las ciencias.

A modo de síntesis, diremos que los buenos docentes y las escuelas que se responsabilizan por el aprendizaje de los estudiantes son los elementos claves para atenuar el efecto que las desigualdades sociales ejercen sobre el logro académico. El potencial transformador de la educación se vería magnificado, si los esfuerzos de los docentes y de las escuelas se acompañaran de políticas sociales y educativas vigorosas, que propendan a limitar el impacto de la condición social y de la segregación escolar, en el aprendizaje.

Referencias

- Affifi, A., Clark, V. y May, S. (2004). *Computer-Aided Multivariate Analysis*. Washington: CRC Press.
- Arancibia, V. (1992). "Efectividad escolar: un análisis comparado". Santiago, Chile. Documento inédito.
- Auguste, S., y Valenzuela, J. P. (2004). *Do Students Benefit from School Competition? The Chilean Experience*. University of Michigan.
- Bain, K. (2007). *Lo que hacen los mejores profesores universitarios*. Valencia: Universidad de Valencia.
- Barber, M. y Mourshed, M. (2007). *How the World's Best-Performing School Systems Come Out on Top*. Londres, Reino Unido: MacKinsey Company.
- Beeby, C. (1986). "The Stages of Growth in Educational Systems", en S. Heyneman y D. White (Eds.), *The Quality of Education and Economic Development*. Washington: Banco Mundial.
- Bellei, C. (2007). "Expansión de la educación privada y mejoramiento de la educación en Chile. Evaluación a partir de la evidencia", *Revista Pensamiento Educativo*, 40. Santiago, Chile: Pontificia Universidad Católica de Chile.
- Bellei, C. y otros (2004). *¿Quién dijo que no se puede? Escuelas efectivas en sectores de pobreza*. Santiago, Chile: Fondo de las Naciones Unidas para la Infancia (UNICEF)/ Ministerio de Educación de Chile.
- Blanco, R. y Treviño, E. (2000). *Informe subregional de América Latina. Evaluación de Educación para Todos en el año 2000*. Santiago, Chile: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO)/ Programa de las Naciones Unidas para el Desarrollo (PNUD)/ Fondo de Población de las Naciones Unidas (FNUAP).
- Bordieu, P. y Passeron, J. (2003). *Herederos: los estudiantes y la cultura*. Buenos Aires, Argentina: Siglo XXI.
- Bowles, S. y Gintis, H. (2001). *Schooling in Capitalist America Revisited*. Extraído desde <http://www.umass.edu/preferen/gintis/soced.pdf>
- Bransford, J., Brown, A. y Cocking, R. (2000). *How People Learn: Brain, Mind, Experience, and School Committee on Developments in the Science of Learning*. Washington: Commission on Behavioral and Social Sciences and Education of the National Research Council National Academy Press.
- Carnoy, M. (2007). "La mejora de la calidad y equidad educativa: una visión realista". *Revista Pensamiento Educativo*, 40. Santiago, Chile: Pontificia Universidad Católica de Chile.
- CEPAL (Comisión Económica para América Latina y el Caribe) (1992). *Educación y Conocimiento: Eje de la transformación productiva con equidad*. Santiago, Chile.

- Donoso, S. y otros (2004). "Financiamiento de los estudios universitarios: un desafío estratégico para la sustentabilidad", *Panorama Socioeconómico*, 29 (Octubre).
- Espínola, V. (2000). *Autonomía escolar: factores que contribuyen a una escuela más efectiva*. Santiago, Chile: Banco Interamericano de Desarrollo (BID).
- Gershberg, A. y Meade, B. (2006). *Accountability de la reforma educativa en América Latina: ¿Cómo puede o no la descentralización aumentar la accountability?*. Santiago, Chile: Programa de Promoción de la Reforma Educativa de América Latina y el Caribe (PREAL).
- Goldstein, J. y otros (2001). "Normal Sexual Dimorphism of the Adult Human Brain Assessed by In Vivo Magnetic Resonance Imaging". *Cerebral Cortex*, 11(6). Oxford, Reino Unido: Oxford University Press.
- Hanushek, E. A. (2005). "The economics of school quality". *German Economic Review*, 6(3). Weinheim, Alemania: German Economic Association.
- Heyneman, S. (2004). "International education quality". *Economics of Education Review*, 23(4). Elsevier.
- Hirschman, A. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- Hsieh, C. y Urquiola, M. (2003). "When Schools Compete, How Do They Compete? An Assessment of Chile's Nationwide School Voucher Program", *NBER Working paper series*. Extraído desde <http://www.nber.org>
- Kane, T. y Steiger, D. (2008). "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation", *NBER Working paper series*. Extraído desde <http://www.nber.org>
- MIDEPLAN (Ministerio de Planificación) (2006). *Encuesta CASEN*. Santiago, Chile.
- MINEDUC (Ministerio de Educación) (2008). *Ley de subvención escolar preferencial*. Santiago, Chile.
- OCDE (Organización de Cooperación y Desarrollo Económicos) (2008). *Informe PISA 2006. Competencias científicas para el mundo de mañana*. Madrid, España.
- Reyes, L. (2008). "El financiamiento compartido en la educación pública como factor de segregación social: entrevista a Juan Pablo Valenzuela". *Boletín Referencias*, 23. Santiago, Chile: Foro Latinoamericano de Políticas Educativas (FLAPE)/Programa Interdisciplinario de Investigaciones en Educación (PIIE).
- Scaife, J. (1998). "Science Education for All? Towards a More Equitable Science Education", en A. Clark y E. Millard (Eds.), *Gender in the Secondary Curriculum: Balancing the Books*. Londres, Reino Unido: Routledge.
- Slavin, R. (2005). "Sand, Bricks, and Seeds: School Change Strategies and Readiness for Reform", en S. Netherlands (Ed.), *The Practice and Theory of School Improvement*. Amsterdam, Países Bajos: Springer Netherlands.
- Solow, R. (1957). "Technical change and the aggregate production function". *Review of Economic Statistics*, 39. Cambridge: Harvard University.
- Torche, F. (2005). "Privatization Reform and Inequality of Educational Opportunity: The Case of Chile". *Sociology of Education*, 78 (October). Washington: American Sociological Association
- Treviño, E. y Treviño, G. (2004). *Estudio sobre las desigualdades educativas en México: la incidencia de la escuela en el desempeño académico de los alumnos y el rol de los docentes*. México, D.F.: Instituto Nacional para la Evaluación de la Educación (INEE).
- UNESCO (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura) (1996). *La educación encierra un tesoro: Informe a la UNESCO de la Comisión Internacional sobre la Educación para el Siglo XXI*. París, Francia.
- _____ (2008). *Segundo Estudio Regional Comparativo y Explicativo: Los aprendizajes de los estudiantes de América Latina y el Caribe*. Santiago, Chile: UNESCO/ Oficina Regional de Educación para América Latina y el Caribe (OREALC).
- _____ (2009). *Aportes para la enseñanza de las ciencias del SERCE*. Santiago, Chile.
- _____ (en prensa). *Factores asociados al rendimiento académico de los estudiantes. Estudio de factores asociados del SERCE*. Santiago, Chile.
- Valenzuela, J. P. (2007). *Evolución de la segregación socioeconómica de los estudiantes chilenos y su relación con el financiamiento compartido*. Santiago, Chile: Ministerio de Educación (MINEDUC).
- Wenglinsky, H. (2002). How Schools Matter: The Link Between Teacher Classroom Practices and Student Academic Performance, *Education Policy Analysis Archives*, 10(2). Arizona: Arizona State University.
- _____ (2003). "Using Large-Scale Research to Gauge the Impact of Instructional Practices on Student Reading Comprehension: An Exploratory Study", *Education Policy Analysis Archives*, 11(19). Arizona: Arizona State University.
- Wright, S., Horn, S. y Sanders, W. (1997). "Teacher and classroom context effects on student achievement: Implications for teacher evaluation". *Journal of Personnel Evaluation in Education*, 11, 57-67. Amsterdam, Países Bajos: Springer Netherlands.

Anexo

1. Modelos de regresión y modelos jerárquicos lineales

Respecto del análisis de perfiles escolares, se ajustan modelos de regresión para cada país de América Latina y para la Región en general, donde el aprendizaje promedio por escuela en Ciencias \bar{Y}_j se asocia al índice socioeconómico y cultural promedio de la escuela \overline{ISEC}_j tal como se aprecia en la Ecuación 1.

$$\text{Ecuación 1: } \bar{Y}_j = \beta_0 + \beta_1 \overline{ISEC}_j + \varepsilon_j$$

Donde j representa a cada escuela y β_0 , al intercepto o nivel del perfil escolar. β_1 , representa la pendiente del perfil escolar; ε_j es el error estocástico de la predicción, y la bondad de ajuste o R^2 , representa la fuerza del perfil, es decir, el poder explicativo del modelo.

El modelo general, aplicado a cada uno de los países de América Latina y a la Región como un todo, asocia la puntuación del estudiante¹⁷ i en la escuela j en la prueba de Ciencias Y_{ij} con el índice socioeconómico y cultural del estudiante $ISEC_{ij}$, centrado en la gran media, tal como se observa en la Ecuación 2.

$$\text{Ecuación 2: } Y_{ij} = \beta_{0j} + \beta_{1j} ISEC_{ij} + r_{ij} + u_{1j}$$

Donde β_{0j} representa el intercepto o nivel del gradiente y β_{1j} corresponde a la pendiente o fuerza del gradiente, que indica la relación entre el rendimiento académico y el ISEC de los estudiantes en la escuela j .

Los modelos lineales jerárquicos se organizan como una taxonomía, en la cual primero se consideran los

factores socioculturales, y después se incorporan características de las escuelas y sus docentes. La taxonomía se apega a la siguiente descripción general:

$$\text{Ecuación 3: } Y_{ij} = \beta_{0j} + \beta_{nj} ESTUDIANTE_{ij} + r_{ij}$$

$$\text{Ecuación 4: } \beta_{0j} = \gamma_{0j} + \gamma_{mj} ESCUELA_j + u_{0j}$$

$$\text{Ecuación 5: } \beta_{nj} = \gamma_{n0} + u_{nj}$$

Las ecuaciones 3, 4 y 5 indican la forma general de los modelos jerárquicos lineales utilizados.

La ecuación 3 corresponde al nivel del estudiante, donde el aprendizaje Y_{ij} depende de un vector de variables relativas al estudiante, cuyos coeficientes son β_{nj} , y donde n corresponde a cada uno de los coeficientes asociados a las variables del nivel estudiante, dentro de cada escuela j . La ecuación incluye un error estocástico de predicción para el nivel de estudiantes, representado por r_{ij} .

Las ecuaciones 4 y 5 son una formulación general para el modelo de nivel de escuela. En la ecuación 4 se observa un vector de variables escolares que predicen el rendimiento de los estudiantes sobre el intercepto, considerando un error estocástico representado por u_{0j} . La ecuación 5, por su parte, establece el comportamiento de las variables del nivel estudiante, agrupadas en escuelas, e incluye su correspondiente error estocástico a nivel de escuela, representado por u_{nj} . Es necesario recordar que la n representa a cada uno de los coeficientes asociados a variables del nivel de estudiante.

2. Perfiles escolares y gradientes

Para establecer si los estimadores del modelo de Chile son significativamente distintos a los estimadores del resto de los países considerados en el análisis,

17 La puntuación individual del estudiante que se utiliza, corresponde a los cinco valores plausibles provistos por PISA y sobre los cuales se hacen todas las estimaciones. Las de modelos jerárquicos, a través de las opciones que ofrece el software, y las de regresión lineal, generando una regresión para cada valor plausible y promediando sus coeficientes.

se ajustaron los modelos de perfiles y gradientes, incluyendo una variable *dummy intercepto* (que toma valor 1 cuando se trata de Chile y cero, para el país de contraste) y una variable para la pendiente, que es la interacción entre la *dummy intercepto* y el ISEC.

Para todos los casos se utilizó un nivel de significancia del 0.05. De esta forma, si la *dummy intercepto* es significativa, implica que hay diferencias en el valor de la constante o en el promedio de logro de los países. Con igual lógica, si la interacción entre la variable *dummy* y el ISEC es estadísticamente significativa, implica que hay diferencias en las pendientes del ISEC entre los países en comparación.

Para comparar cada país con el contexto regional, sin embargo, no fue posible proceder de la misma forma, ya que la composición del contexto regional (América Latina) incluye a cada uno de los países con los que se quiere contrastar (no son muestras independientes). Por ello, se utilizaron los intervalos de confianza de cada estimador y así se determinaron las diferencias significativas entre los coeficientes, siempre que estos no se traslaparan entre sí.

Los resultados obtenidos fueron los siguientes:

Tabla A.1
Comparación entre modelos de gradientes y perfiles

País		Gradiente		Perfiles	
		A.L	Chile	A.L	Chile
Argentina	Constante	EDS	EDS	EDS	EDS
	Pendiente	EDS	EDS	EDS	EDS
Brasil	Constante	EDS	EDS	EDS	EDS
	Pendiente	EDS	EDS	EDS	EDS
Chile	Constante	EDS	-	EDS	-
	Pendiente	EDS	-	EDS	-
Colombia	Constante	NDS	EDS	EDS	EDS
	Pendiente	EDS	EDS	EDS	EDS
México	Constante	EDS	EDS	EDS	NDS
	Pendiente	EDS	EDS	EDS	EDS
Uruguay	Constante	EDS	NDS	EDS	NDS
	Pendiente	EDS	EDS	EDS	EDS

EDS: Existen diferencias significativas
NDS: No existen diferencias significativas

3. Factores Escolares

Tabla A.2
Análisis de componentes principales para la autonomía del establecimiento

Componentes de “Autonomía del establecimiento” Alfa de Cronbach= 0,681	Porcentaje acumulado de la varianza explicada	Preguntas del cuestionario del establecimiento, utilizadas para la construcción del factor.
Autonomía del establecimiento para tomar decisiones relacionadas con los profesores y con el presupuesto (*)	38,1%	<p>P.11. En su establecimiento, ¿quiénes tienen la responsabilidad sobre las siguientes tareas? (las opciones de autoridades con responsabilidad son: director o profesores; directorio del establecimiento; SEREMI o alcalde; Ministerio de Educación).</p> <p>a. Contratar profesores. b. Despedir profesores. c. Establecer sueldo inicial de los profesores. d. Determinar aumentos de sueldo de los profesores. e. Elaborar el presupuesto escolar. f. Decidir sobre la asignación de recursos dentro del establecimiento.</p> <p>P.12. ¿Cuál de las siguientes entidades ejerce una influencia directa en la toma de decisiones sobre el personal, el presupuesto, el contenido curricular y las prácticas de evaluación del establecimiento? (las entidades son: autoridades educacionales comunales, regionales y nacionales; directorio del establecimiento; centro de padres y apoderados; grupos de profesores; grupos de estudiantes; consejos examinadores externos).</p> <p>a. Influencia en el personal. b. Influencia en el presupuesto.</p>
Autonomía del establecimiento para tomar decisiones relacionadas con los estudiantes	54,4%	<p>P.11 En su establecimiento, ¿quiénes tienen la responsabilidad sobre las siguientes tareas?</p> <p>g. Establecer las políticas disciplinarias para los estudiantes. h. Establecer las políticas de evaluación de los estudiantes. i. Aprobar la admisión de alumnos en el establecimiento. j. Elegir qué libros de texto usarán. l. Decidir qué cursos se impartirán.</p>
Autonomía del establecimiento para discernir sobre criterios de enseñanza (*)	63,1%	<p>P.12 ¿Cuál de las siguientes entidades ejerce una influencia directa en la toma de decisiones sobre el personal, el presupuesto, el contenido curricular y las prácticas de evaluación del establecimiento? (las entidades son: autoridades educacionales comunales, regionales y nacionales; directorio del establecimiento; centro de padres y apoderados; grupos de profesores; grupos de estudiantes; consejos examinadores externos).</p> <p>c. Influencia en el contenido curricular. d. Influencia en las prácticas de evaluación.</p> <p>P.11 En su establecimiento, ¿quiénes tienen la responsabilidad sobre las siguientes tareas? (las opciones de autoridades con responsabilidad son: director o profesores; directorio del establecimiento; SEREMI o alcalde; Ministerio de Educación).</p> <p>k. Determinar el contenido de ramos.</p>

(*) Indica aquellos factores no utilizados en la especificación del modelo final

Tabla A.3
Análisis de componentes principales para la rendición de cuentas

Componentes de "Accountability o rendición de cuentas del establecimiento" Alfa de Cronbach= 0,584	Porcentaje acumulado de la varianza explicada	Preguntas del cuestionario del establecimiento, utilizadas para la construcción del factor
Rendición de cuentas	44,7%	<p>P.15 Las siguientes preguntas se refieren a los procedimientos a través de los cuales el establecimiento rinde cuenta de su gestión a los padres: Procedimientos de rendición de cuentas a los padres.</p> <p>P.16 ¿Cuál de las siguientes afirmaciones caracteriza mejor las expectativas de los padres sobre el establecimiento?</p> <ul style="list-style-type: none"> - Hay presión constante de muchos padres, que esperan que el establecimiento establezca estándares de educación muy altos y que los alumnos los alcancen. - La presión sobre el establecimiento, para que alcance estándares de educación más altos entre los alumnos, viene sólo de una minoría de los padres. - No existe presión de los padres para que el establecimiento alcance estándares de educación más altos entre los alumnos. <p>P.17 En su establecimiento, ¿se utilizan los datos de rendimiento escolar, en alguno de los siguientes procedimientos de rendición de cuentas de la gestión?</p> <ul style="list-style-type: none"> a. Los datos relativos al rendimiento escolar se informan públicamente (por ejemplo, en los medios de comunicación). b. Los datos relativos al rendimiento escolar se usan para la evaluación del desempeño del director o la directora. c. Los datos relativos al rendimiento escolar se usan para la evaluación del desempeño de los profesores. d. Los datos relativos al rendimiento escolar se usan para tomar decisiones sobre la distribución de los recursos pedagógicos dentro del establecimiento. e) Los datos relativos al rendimiento escolar son monitoreados a través del tiempo, por una autoridad administrativa.

Tabla A.4
Análisis de componentes principales para la calidad de los profesores y la relación con los estudiantes

Componentes de “Calidad de los profesores y relación de los estudiantes con la Ciencia” Alfa de Cronbach= 0,901	Porcentaje acumulado de la varianza explicada	Preguntas del cuestionario de estudiante, utilizadas para la construcción del factor
Capacidad del profesor de Ciencias de hacer partícipes a los estudiantes en el aprendizaje	15,7%	P.34 En el colegio, cuando aprendes temas científicos, ¿con qué frecuencia realizan las siguientes actividades? k. El profesor da a los estudiantes la oportunidad de elegir sus investigaciones. l. El profesor usa la ciencia para ayudar a los estudiantes a entender el mundo más allá del colegio. m. Los estudiantes discuten sobre los temas tratados en clases. n. Los estudiantes realizan sus experimentos, siguiendo las instrucciones del profesor. o. El profesor explica claramente la importancia que tienen los conceptos científicos para nuestras vidas. p. Se les pide a los estudiantes que hagan una investigación para comprobar sus propias ideas. q. El profesor usa como ejemplo, las aplicaciones tecnológicas, para mostrar cuán importante es la ciencia para la sociedad.
Afinidad del estudiante con las ciencias	29,0%	P.16 ¿En qué medida estás de acuerdo o en desacuerdo con las siguientes afirmaciones? a. En general lo paso bien aprendiendo sobre temas científicos. b. Me entretengo leyendo sobre temas científicos. d. Me entretengo cuando aprendo nuevos conocimientos científicos. e. Me interesa aprender Ciencias.
Proyección de los estudiantes en las ciencias	40,8%	P.29 ¿En qué medida estás de acuerdo o en desacuerdo con las siguientes afirmaciones? a. Me gustaría trabajar en una carrera relacionada con las ciencias. b. Me gustaría estudiar Ciencias después de terminar la Enseñanza Media. c. Me gustaría dedicarme a estudiar temas científicos avanzados, durante toda mi vida. d. Cuando sea adulto, me gustaría trabajar en proyectos científicos.
Profesores incentivan la aplicación práctica de las ciencias (a nivel de escuela es “Oportunidad para realizar investigación científica”)	51,3%	P.34 En el colegio, cuando aprendes temas científicos, ¿con qué frecuencia realizan las siguientes actividades? b. Los estudiantes disponen de tiempo para realizar experimentos prácticos en el laboratorio. c. Se pide a los estudiantes que piensen cómo podría investigarse una pregunta científica en el laboratorio. h. Los estudiantes pueden diseñar sus propios experimentos. d. Se pide a los estudiantes que apliquen un tema científico a los problemas cotidianos.
Capacidad del profesor de incentivar la aplicación de las ciencias en la vida diaria (*)	60,9%	P.34 En el colegio, cuando aprendes temas científicos, ¿con qué frecuencia realizan las siguientes actividades? a. Los estudiantes tienen oportunidades de explicar sus ideas. e. Las clases incorporan la opinión que los estudiantes tienen sobre los distintos temas. g. El profesor explica cómo una teoría científica puede aplicarse a distintos fenómenos (por ejemplo, el movimiento de los objetos, las sustancias con propiedades similares o las características de los organismos vivos). f. Se les pide a los estudiantes que saquen conclusiones de un experimento que llevaron a cabo.

(*) Indica aquellos factores no utilizados en la especificación del modelo final

Tabla A.5
Análisis de componentes principales para la autonomía del establecimiento

Componentes de "Disponibilidad de recursos en el establecimiento" Alfa de Cronbach= 0,870	Porcentaje acumulado de la varianza explicada	Preguntas del cuestionario del establecimiento, utilizadas para la construcción del factor
Dotación de recursos didácticos para el aprendizaje (*)	41,0%	P.14 La enseñanza en su establecimiento, ¿se ve afectada por alguno de los siguientes factores? h. Escasez de material de enseñanza o material inadecuado (ej. libros de texto). l. Escasez de material de biblioteca o material inadecuado. m. Escasez de recursos audiovisuales o recursos audiovisuales inadecuados. j. Falta de conexión a internet o conexión inadecuada. k. Escasez de softwares computacionales para la enseñanza o softwares inadecuados.
Dotación de profesores calificados	58,5%	P.14 La enseñanza en su establecimiento, ¿se ve afectada por alguno de los siguientes factores? a. Escasez de profesores de Ciencias calificados. b. Escasez de profesores de Matemáticas calificados. c. Escasez de profesores de Lengua Castellana y Comunicación calificados. d. Escasez de profesores calificados para otros ramos.
Dotación de personal de apoyo (*)	66,3%	P.14 La enseñanza en su establecimiento, ¿se ve afectada por alguno de los siguientes factores? e. Escasez de técnicos de laboratorio. f. Escasez de personal de apoyo. g. Escasez de equipamiento del laboratorio de Ciencias o equipamiento inadecuado.

(*) Indica aquellos factores no utilizados en la especificación del modelo final

¿Qué nos dice PISA sobre la educación de los jóvenes en Chile?

¿Qué explica las diferencias de resultados PISA Matemática entre Chile y algunos países de la OCDE y América Latina?

Por
Juan Pablo Valenzuela
Cristián Bellei
Alejandro Sevilla
Alejandra Osses¹

Resumen

El objetivo de esta investigación fue identificar algunos factores que explican las diferencias de resultados obtenidos por los estudiantes chilenos en Matemática en la prueba PISA 2006, respecto de los obtenidos por los estudiantes de Polonia, España y Uruguay— países que superan a Chile en esta medición. Para cumplir este propósito, el estudio implementó una variante de la metodología de descomposición propuesta por Oaxaca-Blinder. Esto implicó, primero, identificar las características de los estudiantes que rindieron la prueba, de sus familias y de los establecimientos a los que asisten, que fueran explicativas del diferencial observado de puntajes PISA entre Chile y cada uno de los países de referencia. En segundo término, el estudio se centró en estimar si estos factores diferían entre los países, en el modo en que afectaban dichos resultados, pues esta diferencia podría explicar también, parte de la brecha observada.

Los resultados obtenidos indican que no es posible atribuir el menor nivel de desempeño de Chile, a desventajas en condiciones, recursos o características que pudieran tener los estudiantes o los establecimientos; sino a la eficiencia del sistema educacional chileno, el cual aparece significativamente menos eficiente que los de los países comparados, en términos de transformar los recursos y condiciones con que cuenta, en resultados de aprendizaje de sus estudiantes. De particular interés resultan las estimaciones del valor agregado por cada año de escolaridad cursado, claramente menores en Chile que en los países europeos usados como referencia.

¹ Juan Pablo Valenzuela, Cristián Bellei, Alejandro Sevilla y Alejandra Osses, son miembros del Centro de Investigación Avanzada en Educación (CIAE) de la Universidad de Chile.

Los autores agradecen a Kevin Macdonald, consultor del equipo de educación de la Human Development Network, asociada al Banco Mundial, quien facilitó los programas de análisis de PISA para STATA y colaboró en su utilización.

Introducción

¿Qué explica las diferencias de resultados de aprendizaje entre los estudiantes chilenos y sus pares de otros países?, ¿son acaso las diferencias de recursos entre los establecimientos, diferencias en la eficiencia de su utilización, o en el grado de autonomía que los establecimientos tienen para gestionarlos?, ¿es la diferencia en las condiciones socioeconómicas de las familias un factor relevante? Conforme el debate sobre políticas educacionales en Chile “se abre al mundo”, este tipo de preguntas comienza a ocupar un lugar cada vez más preponderante en la discusión pública y académica sobre educación en nuestro país. La presente investigación tuvo como propósito precisamente abordar esta clase de interrogantes.

Utilizando los resultados de aprendizaje en Matemática y Lectura, medidos por PISA 2006, así como la información complementaria provista por esta misma fuente, el presente estudio identificó algunos de los principales factores que están a la base de las diferencias de logro de los estudiantes chilenos en Matemática, en comparación con tres países –dos de la OCDE y uno de América Latina. Ciertamente, el análisis también incluyó algunas de las explicaciones más recurrentes en el debate político y académico sobre esta materia, para las cuales no se encontró sustento en estos datos.

Este artículo se estructura de la siguiente forma. En la sección 1 se provee el contexto más amplio que explica y motiva la realización de este estudio; luego, se describe la metodología empleada, que permite comparar los resultados PISA entre Chile y tres países usados como referencia; en la sección 3 se exponen los principales hallazgos de este estudio y, finalmente, se presentan las principales conclusiones del estudio y se plantean algunas reflexiones en torno suyo. Es importante precisar que los resultados de esta investigación son reportados en dos capítulos: el

presente aborda Matemática y el siguiente, Lectura. Ahora bien, dado que ambos comparten un mismo diseño y que la muestra de estudiantes y escuelas es la misma, las secciones 1, 2, y 3.1 de este capítulo son compartidas con el capítulo siguiente de Lectura.

1. Antecedentes

Los resultados de las pruebas estandarizadas de aprendizaje son una información cada vez más relevante para el proceso de discusión pública y la toma de decisiones en las políticas educativas, por lo cual comprenderlos bien, resulta esencial (Evers y Walberg, Eds., 2002 y Arregui, Ed., 2006). Durante toda esta década, Chile ha estado recibiendo “malas noticias” provenientes de los sistemas de evaluación de aprendizajes, tanto nacionales como internacionales (comenzando por las pruebas SIMCE y TIMSS, aplicadas en 1999, como se expone en Bellei, 2003 y Valenzuela, 2008): malos resultados a nivel nacional, nulo mejoramiento y alta inequidad. Es posible afirmar que esta información (y el uso dado por los medios de comunicación y los encargados de la toma de decisiones) ha contribuido decisivamente a generar un clima de decepción respecto de la reforma educativa y de severa crítica hacia el sistema escolar, cimentando la idea de que, a pesar de las grandes inversiones y reformas, los resultados no han evidenciado las mejoras buscadas.

Los resultados de Matemática, obtenidos por los estudiantes chilenos en la prueba PISA 2006 rompen de cierta manera con esta tendencia. Entre las mediciones PISA realizadas los años 2000 y 2006, Chile aumentó de 384 a 411 puntos. Aunque limitaciones metodológicas de la prueba impidan una comparación directa entre estos puntajes, sí es posible afirmar que la distancia que separa al promedio chileno del promedio de los países de la OCDE disminuyó considerablemente en dicho período, pasando de 1,16 desviaciones estándar en 2000 a 0,87, en 2006. Con

todo, a pesar de este avance, la distancia que separa a Chile de los países desarrollados miembros de la OCDE es aún considerable.

Desde el punto de vista de Chile, el hecho de que PISA 2006 haya sido sensible a cierto mejoramiento, pero continúe significando un parámetro exigente de referencia, representa una enorme oportunidad para los análisis comparados, como el realizado en este estudio. En efecto, la aproximación metodológica utilizada, permite separar dos cuestiones que suelen estar mezcladas en estos análisis: los sistemas escolares nacionales se diferencian tanto porque poseen características distintas (por ejemplo, establecimientos con más o menos recursos educativos) como porque utilizan de modos diferentes dichas características (por ejemplo, más o menos eficiencia en el uso de los recursos educativos). Esto es relevante ya que, generalmente, las discusiones de política educacional tienden a privilegiar la imitación de ciertos recursos con que cuentan otros países, sin advertir la enorme diferencia que hacen los procesos de uso de dichos recursos.

Chile siempre ha mirado más allá de sus fronteras para intentar construir primero, expandir luego, y mejorar ahora, su sistema educacional. El proceso de globalización al que asistimos ha acrecentado esta tendencia en los últimos años. Sin embargo, “importar” políticas educacionales no es tarea fácil, entre otras razones porque la investigación internacional comparada es aún un área de poco desarrollo en nuestro país. Así, es fácil dejarse confundir por el atractivo que causan algunos rasgos salientes de países a los que se quiere imitar, o por la argumentación interesada y poco rigurosa de actores que usan ejemplos de “casos exitosos” para intentar convencer sobre la conveniencia de sus demandas. Este estudio pretende contribuir a informar dicho debate.

2. Metodología

En esta investigación, utilizando la metodología de descomposición conocida como Oaxaca-Blinder, se estimó la proporción del diferencial de los resultados PISA en Matemática, entre Chile y tres países de referencia, que podía ser explicada por el cambio en características observables de la población escolar evaluada (referidas a estudiantes, establecimientos y ciertos aspectos institucionales de los sistemas escolares). Para ello, se analizó un conjunto significativo de variables, las que se seleccionaron, basándose tanto en la amplia literatura existente sobre la materia, como en su disponibilidad en las bases de la prueba PISA 2006 para todos los países en consideración. Complementariamente, dado que, aun si las características de las poblaciones escolares de dos países en comparación fuesen similares, el efecto de dichas características sobre los resultados de aprendizaje podría ser distinto (en un sentido que favoreciese más a los estudiantes de uno u otro país), el estudio también estimó la proporción del diferencial observado en los resultados PISA, atribuible a esta razón. Así, por ejemplo, se pudo determinar si eran sólo aspectos socioeconómicos u otros factores vinculados a los estudiantes y sus familias, los que incidían en las brechas existentes², o si también afectaban las principales características del establecimiento y los aspectos institucionales de los sistemas escolares de cada país.

El estudio se focalizó en los resultados de Chile, España, Uruguay y Polonia. Para los análisis, se usó la base de datos pública de PISA 2006, que contiene todos los resultados de los países participantes en la medición. Específicamente, se utilizaron variables relativas a las características de los estudiantes y de

2 Ciertamente, factores escolares y sociales pueden combinarse para explicar cambios en los resultados medidos por tests entre grupos. Véase, por ejemplo, Grissmer, Flanagan y Williamson, 1998.

los establecimientos, recolectadas a través de los cuestionarios aplicados a cada alumno que rindió la prueba y a los directores de todos los establecimientos participantes.

Cada uno de estos aspectos de la metodología de investigación es explicado con cierto detalle en las secciones siguientes.

2.1 Países de referencia incluidos en el estudio

A continuación se señalan los principales fundamentos que se consideraron para la selección de los países que se usaron en la comparación.

Polonia se consideró de interés porque —junto a Chile y Corea— es uno de los tres países que mejoró sustantivamente sus resultados en Lectura entre el año 2000 y el 2006. En PISA 2006, Polonia obtuvo 508 puntos en Lectura, ubicándose en el 9º puesto de la clasificación elaborada por PISA en base al puntaje promedio de los 57 países participantes, esto equivale a 8 puntos por sobre el promedio de la OCDE. Los resultados de Polonia en Lectura son equivalentes a los de Holanda o Suecia, y superiores a los de Japón, Inglaterra o Alemania. Aunque los resultados en Matemática no son tan destacados, Polonia obtuvo 495 puntos, ubicándose 5 puntos por debajo del promedio de los países de la OCDE, lo que es equivalente a Alemania, Inglaterra o Noruega, y superior a Estados Unidos, Rusia o España.

La selección de Polonia se debe además, a que comparte ciertas características macroeconómicas y sociales con Chile (por ejemplo, es un país que tiene un PIB por habitante similar al chileno en el año 2000) y a que PISA señala que la variación entre los rendimientos de los establecimientos en Lectura, ese mismo año, fue bastante alta, incluso con niveles superiores a los de Chile. Lo que hace a Polonia un país interesante de incluir en el análisis, es que ha

logrado avances significativos y sostenidos en su rendimiento en Lectura (en el año 2000, obtuvo 479 puntos promedio en esta disciplina; en el año 2003 obtuvo 497 y en el año 2006, 508 puntos). Junto con lo anterior, Polonia ha logrado reducir la variación de los rendimientos entre los establecimientos, tanto en Lectura como en Matemática (en Lectura, de un 63% en el año 2000 a un 16% en el año 2006, y en Matemática, de un 54% a un 15% en el mismo período), haciendo así menos desigual su sistema escolar.

En el caso de España, se consideró que, de los países desarrollados de la OCDE, éste es el que comparte con Chile diversos aspectos culturales, como el idioma, especialmente relevantes para el análisis. España obtuvo 461 puntos promedio en Lectura, en 2006, ubicándose 39 puntos por debajo del promedio de los países de la OCDE, pero sobre Rusia e Israel. En tanto, en la disciplina de Matemática, España obtuvo mejores resultados: 480 puntos, ubicándose 20 puntos bajo el promedio de los países de la OCDE, con un resultado equivalente al de Estados Unidos y Rusia, y superior al de Israel, Italia y Portugal. Aunque España está lejos de Chile en los indicadores macroeconómicos y sociales, su vínculo histórico con nuestro país lo hace relevante en términos comparativos.

La selección de Uruguay se debió a que, de los países latinoamericanos participantes, este es el que obtuvo puntajes más cercanos a Chile. Uruguay obtuvo 413 puntos promedio en Lectura, lo que lo sitúa en la clasificación 42 de la escala que PISA elabora en base al puntaje promedio de los países participantes. A nivel latinoamericano, Uruguay se ubica inmediatamente bajo Chile (con una diferencia de 29 puntos) y por sobre los otros 4 países de la región que participaron en esta medición. Al igual que Chile, en esta clasificación, Uruguay se sitúa por debajo de todos los países OCDE, a excepción de México. En la disciplina de Matemática, Uruguay alcanzó 427 puntos promedio, ubicándose en el primer lugar a nivel latinoamericano y 16 puntos por sobre Chile, por lo que la decisión

de incluirlo obedeció justamente a la posibilidad de indagar si las similitudes de puntaje en Lectura y Matemática se producen por la relativa equivalencia de condiciones socioeconómicas y de recursos entre ambos países, pese a que ambos poseen sistemas escolares marcadamente diferentes.

2.2 Variables incluidas en el estudio

El presente estudio incorporó variables de los estudiantes, de los establecimientos e institucionales, las que se presentan en la Tabla 1 (para el detalle de la descripción, véase el Anexo).

En relación con los estudiantes, se incorporaron 6 variables: el género, los años de escolaridad, el retraso escolar, el índice socioeconómico y cultural de sus familias (ISEC), las expectativas laborales futuras y la valoración personal sobre la importancia de tener un buen rendimiento en Matemática y Lenguaje (por separado). Las cuatro primeras variables refieren a características objetivas de los estudiantes, mientras que las dos últimas, miden aspectos subjetivos.

Respecto de la dimensión “establecimiento”, las variables incluidas en este estudio fueron el efecto par, el tamaño del curso, el tamaño del establecimiento, el tiempo de instrucción semanal de Lenguaje (ver próximo capítulo) y de Matemática, la relación alumno-profesor, si el establecimiento es complejo, y el índice de calidad de los recursos educativos. Nótese que en este grupo existen dos variables que no son medidas a nivel del establecimiento, sino del alumno: el “efecto par”³ y el tiempo de instrucción semanal para cada disciplina (es decir, ambas variables tienen valores diferentes para cada estudiante).

Finalmente, la dimensión institucional consideró: la presión que ejercen los padres sobre el establecimiento para que éste alcance altos estándares educativos; la autonomía de los establecimientos en la asignación de recursos y en la toma de decisiones curriculares y de evaluación; la utilización de criterios académicos para seleccionar a los estudiantes que ingresan al establecimiento, y la dependencia de los establecimientos (pública/privada). Se debe aclarar que en esta dimensión, las variables fueron medidas por PISA a nivel del establecimiento, es decir, al interior de un país, los establecimientos varían en estos indicadores que se han considerado “institucionales”.



3 En la literatura, “efecto par” es un concepto genérico referido a las diversas formas de influencia (algunas directas como la posibilidad de hacer tareas juntos, y otras indirectas, como el efecto sobre la motivación de los docentes para enseñar) que las diversas características de los compañeros (género, raza, habilidades, nivel socioeconómico, etc.) pueden ejercer sobre las oportunidades y logros de aprendizaje de un alumno. Como queda claro, en este estudio, el indicador de “efecto par” mide específicamente las condiciones socioeconómicas y culturales de las familias de los compañeros, condiciones que la investigación ha asociado con un conjunto de otras variables relevantes para el proceso escolar. Ciertamente, esta metodología no permite distinguir entre los diferentes canales por los que el efecto par estaría afectando los resultados de aprendizaje de los estudiantes.

Tabla 1.
Variables del estudio

Variable	Definición
Variables de resultado	
Lectura	Puntaje PISA Lectura, considera los 5 valores plausibles.
Matemática	Puntaje PISA Matemática, considera los 5 valores plausibles.
Variables del estudiante	
Género	Variable dummy para el género del estudiante (categoría omitida: hombre).
Años de escolaridad	Variable que indica los años de escolaridad formal del estudiante: escala con rango 7 a 11.
Retraso escolar	Variable dummy que indica si el estudiante ha repetido algún curso (categoría omitida: sin retraso escolar).
Índice socioeconómico y cultural del estudiante (ISEC)	Variable normalizada por PISA (se usaron los valores originales) que considera educación y ocupación de los padres y bienes en el hogar.
Expectativas laborales	Expectativa de los estudiantes sobre su futuro laboral: escala con rango 1 (altas) a 3 (bajas).
Valorización sobre la importancia de un buen rendimiento en Matemática y Lenguaje	Valoración que el estudiante asigna a rendir bien en Lenguaje o Matemática: escala con rango 1 (muy importante) a 3 (de menor importancia).
Variables del establecimiento	
ISEC de los compañeros (efecto par)*	Media del ISEC de los compañeros del establecimiento de cada estudiante (variable normalizada por PISA; se usaron los valores originales).
Tamaño del curso	Variable que indica el tamaño promedio de estudiantes por curso, en el nivel escolar evaluado.
Tamaño del establecimiento	Variable que indica el número de estudiantes matriculados.
Tiempo de instrucción para las asignaturas de Lenguaje y Matemática*	Horas destinadas a las clases regulares de Lenguaje o Matemática en el establecimiento, según reporta el estudiante: escala con rango 1 a 7.
Estudiantes por profesor	Cantidad de estudiantes por profesor al interior del establecimiento.
Establecimiento complejo	Variable dummy que indica si el establecimiento tiene los niveles de educación primaria junto con los de secundaria (variable omitida: establecimiento no complejo).
Índice de calidad de los recursos educativos	Variable normalizada por PISA (se usaron los valores originales) que considera los factores que afectan la enseñanza en el establecimiento, según el director.
Variables Institucionales	
Índice de autonomía en las decisiones sobre asignación de recursos	Variable normalizada por PISA (se usaron los valores originales), que se deriva del número de decisiones en relación a la asignación y administración de los recursos, que son de responsabilidad del establecimiento, según el director.
Índice de autonomía en las decisiones curriculares y de evaluación	Variable normalizada por PISA (se usaron los valores originales), que se deriva del número de decisiones en relación al currículum y la evaluación, que son de responsabilidad del establecimiento, según el director.
Criterios académicos de alta prioridad o prerrequisitos para la selección de los estudiantes postulantes al establecimiento	Variable dummy que indica si el establecimiento aplica criterios académicos de selección, como alta prioridad o requisito, según el director (categoría omitida: no considera estos criterios o considera solo el desempeño o la recomendación).
Presión de los padres sobre el establecimiento, por altos estándares	Variable dummy que indica si los padres ejercen presión por altos estándares sobre el establecimiento, según el director (categoría omitida: poco o nada de presión).
Establecimiento público	Variable dummy que indica si el establecimiento es de administración pública (categoría omitida: establecimiento no público).

* Variables medidas a nivel del estudiante.

Las variables incluidas en el estudio fueron seleccionadas, basándose en la literatura internacional y en los análisis realizados a partir de los resultados de otras pruebas internacionales (que consideran las dos aplicaciones previas de PISA⁴, así como las pruebas TIMSS y PIRLS, administradas por la IEA⁵). A continuación se presenta un breve argumento sobre esta selección, exclusivamente basado en los análisis hechos por los propios autores de estas pruebas.

A nivel del alumno, el género es una de las variables más sistemáticamente relacionadas con el rendimiento, tanto en Lectura como en Matemática. En efecto, los análisis de PISA 2000 (OCDE, 2003) muestran que, en todos los países participantes, las mujeres obtienen mejores puntajes en Lectura y esta diferencia es sustantiva: en promedio existen 32 puntos de diferencia a favor de las niñas. La misma tendencia se ha observado en las evaluaciones sobre Lectura realizadas por la IEA (pruebas PIRLS 2001 y 2006; ver Mullis y otros, 2003 y 2007) y en las evaluaciones PISA 2003 y 2006 (OCDE, 2004 y 2007). En Matemática, aunque el género también es una variable que juega un papel importante en el rendimiento de los estudiantes, la diferencia es a favor de los hombres. En efecto, en los análisis de TIMSS 1995 (Harmon y otros, 1997), ya se había constatado la asociación entre rendimiento en Matemática y el género de los estudiantes, asociación sistemáticamente observada en las siguientes mediciones realizadas por TIMSS (Mullis y otros, 2000). Los resultados de PISA 2000 y 2003 confirmaron este hallazgo, pese a que los análisis indican que la diferencia entre los puntajes de los hombres y las mujeres no sigue un patrón tan universal como en Lectura; siendo además, la diferencia promedio a favor de los hombres en Matemática, comparativamente menor (11 y 14 puntos en PISA 2000 y 2003, respectivamente) (OCDE, 2003 y 2004).

4 Véase OCDE, 2003, 2004 y 2007.

5 Véase Mullis y otros, 2000, 2003 y 2007 y Harmon y otros, 1997.

En términos de la estratificación social, los análisis efectuados por PISA son bastante refinados. En PISA 2000 y 2003, se identificó el estatus ocupacional de los padres como una variable altamente relacionada con el rendimiento de los estudiantes en las tres disciplinas evaluadas (Matemática, Lectura y Ciencias). Por ejemplo, en el año 2000, la diferencia observada en Lectura, entre los estudiantes pertenecientes al cuartil superior del índice PISA de estatus ocupacional y los del cuartil inferior, fue de 81 puntos en promedio para los países OCDE. En PISA 2003, en tanto, esta diferencia fue de 93 puntos en Matemática, equivalente a más de un nivel y medio de desempeño en la escala global de esta disciplina. Aun cuando estas brechas son enormes, es necesario tomar en cuenta que ellas varían sensiblemente según el país que se considere.

En cuanto a los logros educativos de la madre, los análisis de PISA señalan a su vez que la relación entre los logros educativos de las madres y el rendimiento escolar en las tres disciplinas evaluadas por esta prueba es positiva y relevante en todos los países. En PISA Lectura, tanto en 2000 como en 2003, la diferencia de logro promedio entre los estudiantes cuyas madres habían completado la educación secundaria superior y aquellos cuyas madres no lo habían hecho, es de aproximadamente 44 puntos para los países OCDE, mientras que en Matemática, la diferencia en 2003 es de aproximadamente 50 puntos.

Igualmente sólida es la relación que existe entre los resultados escolares y la posesión de bienes culturales en el hogar: la variación en una unidad del índice PISA de posesiones culturales va unida a una diferencia de 68 puntos en la escala de Lectura en PISA 2000 y de 12 puntos en la escala de Matemática en PISA 2003. A su vez, el nivel socioeconómico de la familia (medido en PISA por un índice basado en el reporte que hacen los estudiantes acerca de la posesión de ciertos bienes en el hogar), en PISA 2000, influyó en los resultados de todos los países participantes en la evaluación de ese año (salvo Albania): los estudiantes



con un mayor valor en este índice, tendieron a tener un mejor rendimiento en Lectura, aunque la diferencia varía según el país considerado.

Al combinar los aspectos económicos, sociales y culturales de los estudiantes, en el índice PISA de estatus socioeconómico y cultural (ISEC)⁶, se observó que, en general, los estudiantes con un mayor valor en este índice obtuvieron mejores resultados escolares, tanto en Lectura como en Matemática, aunque la relación no es determinista y muchos estudiantes provenientes de familias socioeconómica y culturalmente menos favorecidas, obtuvieron también buenos resultados (y viceversa). Más específicamente, la relación entre el rendimiento en Lectura o Matemática y este índice, es más fuerte para los estudiantes de menor nivel socioeconómico que para los de los niveles superiores.

En cuanto a las variables del establecimiento, en PISA 2000 se identificó un conjunto de factores relacionados con la brecha existente entre los rendimientos de los diferentes establecimientos educativos al interior de un país. Luego, en PISA 2003 se focalizó el análisis en tres dimensiones: a) la instrucción y la enseñanza

(principalmente referida a aspectos de gestión del aula y a estrategias de enseñanza); b) la eficacia del centro de enseñanza (que considera las características de la gestión y la organización de los establecimientos), y c) factores relacionados con los recursos de los establecimientos (tales como el tamaño, la proporción entre número de estudiantes y de profesores, la calidad de la infraestructura y de sus recursos educativos, y la experiencia del profesorado, su formación y su retribución, y en cómo se traducen estos recursos en resultados educativos). En este estudio, se consideraron solo las dos últimas dimensiones⁷.

Por último, también se consideraron para el análisis algunas variables relevantes en términos de las discusiones de política educativa y de la investigación académica, independientemente de que su efecto neto en los resultados PISA sea discutible. Así, por ejemplo, en cuanto a la gestión y administración de los establecimientos escolares, los datos brutos de PISA 2000 y 2003 sugieren que en aquellos países en los cuales, en promedio, los directores informaban tener mayores grados de autonomía en ciertos aspectos de la gestión del centro educativo (asignación del presupuesto escolar, despido y nombramiento de profesores, contenidos y ofertas de cursos, y políticas disciplinarias), el rendimiento medio en Lectura y Matemática tendía a ser más alto. Sin embargo, estas diferencias se anulaban al controlar por otros factores relevantes, sugiriendo una relación espúrea. En lo referido a la infraestructura de los establecimientos, los resultados de PISA 2000 y 2003 indicaron que no existe relación con el rendimiento de los estudiantes; en cambio, la calidad de los recursos educativos disponibles en el establecimiento sí parece ser un factor explicativo de las diferencias de resultados.

Finalmente, en este mismo sentido, uno de los aspectos más controversiales en la política y la investi-

⁶ El cual está compuesto por el Índice Socioeconómico Internacional de Estatus Ocupacional más alto alcanzado por alguno de los padres, el máximo nivel de educación alcanzado por alguno de los padres (traducido en años de escolaridad), un índice de los recursos educativos disponibles en el hogar, y el número de libros en el hogar.

⁷ Las variables consideradas en la primera dimensión, en PISA 2006, se referían a la enseñanza de Ciencias, disciplina no incluida en nuestro estudio.

gación educacional ha sido la comparación entre los establecimientos públicos y privados. Los resultados de PISA 2000 y 2003 mostraron que los estudiantes de los establecimientos privados obtienen mejores rendimientos que los estudiantes de los establecimientos públicos. Sin embargo, existen muchos factores que afectan la elección del centro escolar. La economía familiar, por ejemplo, puede constituir un impedimento importante para quienes deseen asistir a centros privados que realizan cobros a las familias; incluso, los centros privados gratuitos pueden aplicar prácticas de selección y admisión en cierto modo restrictivas. Con todo, la diferencia de puntajes PISA a favor de los establecimientos privados desaparece, al controlar por el nivel socioeconómico de los estudiantes y del establecimiento, lo que sugiere que los centros privados obtienen parte significativa de su mejor rendimiento por el efecto combinado de las mejores condiciones socioeconómicas y culturales de sus estudiantes, vistas en términos individuales y colectivos (“efecto par”).

2.3 Descomposición de Oaxaca–Blinder

Una vez identificados ciertos países de comparación y definido un conjunto de variables comunes, fue posible proceder con la implementación de la metodología de descomposición de Oaxaca–Blinder (Oaxaca, 1973 y Blinder, 1973). En primer término, para la comparación de los resultados de PISA 2006 entre Chile y otros países, se elaboraron modelos estilizados y equivalentes, de funciones de producción de los desempeños individuales de los estudiantes de cada país⁸.

De esta forma, se estimó la función de producción para cada país, así:

$$[1] \quad Y_{i,s,d}^p = \beta_{0,s,d}^p + \beta_{1,s,d}^p A_{i,s,d}^p + \beta_{2,s,d}^p S_{i,s,d}^p + \beta_{3,s,d}^p I_{i,s,d}^p + v_{i,s,d}^p + \varepsilon_{i,s,d}^p$$

donde Y refleja el resultado individual en la prueba PISA 2006 del estudiante i , en el establecimiento s , en el test d (Matemática), del país p . Adicionalmente, A representa un conjunto de variables a nivel del estudiante, S son variables a nivel del establecimiento o curso en el cual el estudiante i estudia, I son variables institucionales del sistema educativo asociadas al establecimiento S , y β son los coeficientes estimados para las diversas variables de control. Para estimar estos coeficientes asociados al puntaje PISA, se utilizaron por separado los cinco valores plausibles que proporciona PISA del puntaje de cada alumno en cada prueba. El propósito de trabajar con los valores plausibles y no con los puntajes individuales directamente, se basa en las propiedades de consistencia que se encuentran en los estimadores para los parámetros de la población. De esta forma, los coeficientes de regresión (y todo análisis que involucró los valores plausibles) se estimaron cinco veces, siendo el valor del estadístico, el promedio de las cinco estimaciones; su varianza a su vez, fue ajustada por cada estimación y en conjunto (véase el detalle sobre la metodología de trabajo usada con valores plausibles, en el Anexo).

Para considerar que las funciones de producción entregan estimadores no sesgados, teóricamente se requeriría incluir en estas funciones de producción, todos los factores actuales y previos que afectan el desempeño de cada estudiante. Sin embargo, obviamente, esta información no está disponible en la base de datos de los resultados PISA, la cual contiene un conjunto de variables que potencialmente afectan el desempeño de los estudiantes, pero medidas sólo en un momento del tiempo. Además, como se indicó, se requiere que las variables a utilizar estén disponibles para todos los países que están siendo comparados. A pesar de estas restricciones, las

8 Es muy importante tener en cuenta que la aplicación de esta técnica requiere que los modelos de regresión utilizados sean exactamente los mismos para los grupos (en este caso, países) que están siendo comparados. Esto, aunque garantiza la comparabilidad de los resultados, impone una rigidez obvia al análisis, por cuanto no es posible ajustar la combinación de variables explicativas, a las particularidades de cada país. De hecho, se debe proceder con criterio, a fin de estimar un modelo razonablemente adecuado para dar cuenta de los resultados PISA en la diversidad de contextos que están siendo estudiados.

bases de datos PISA 2006 son bastante completas en entregar información relativa a las características de los estudiantes y sus familias, del establecimiento y del curso al que asiste cada estudiante, y también de algunas variables institucionales que la literatura identifica con alguna importancia en el desempeño escolar. En efecto, como se verá, algunas de estas variables han sido frecuentemente omitidas en los estudios previos, incluso en países como Chile, cuyo sistema de evaluación de aprendizajes es comparativamente intensivo en este aspecto.

Ahora bien, dado que los estudiantes comparten la experiencia escolar de pertenecer al mismo establecimiento y el diseño muestral de PISA se basa en la selección de establecimientos y luego de estudiantes al interior de estos, existe una correlación entre las características observadas y no observadas de los estudiantes al interior de cada establecimiento. Por esta razón, la función de producción contempla componentes de errores compuestos, es decir, por una parte existe un término v , que refleja un error a nivel del establecimiento, y por otra, ε , que representa un error asociado a cada estudiante. Dado este diseño, las regresiones realizadas contemplaron estimaciones con *clusters* de los estudiantes en los establecimientos seleccionados.

Por otra parte, debido al alto porcentaje de observaciones sin información en algunas variables (valores ausentes), especialmente aquellas referidas a las expectativas laborales futuras de los estudiantes y a la razón alumno/profesor, fue indispensable aplicar una metodología de imputación que permitiera que las estimaciones fuesen más representativas de la población de estudiantes analizados⁹. Esencialmente, el método aplicado consistió en la imputación de medianas, correspondientes al valor observado de un conjunto de casos, que coincidieran con similares subgrupos asociados a las diversas variables de control, de tal forma que las observaciones con datos

imputados pertenecieran a subgrupos relativamente similares (véase el detalle de la implementación del método de imputación y los indicadores de su efectividad, en el Anexo).

Contando con estas funciones de producción comparables entre países y disciplinas, fue posible analizar los factores que explicaban los diferenciales o brechas entre los resultados individuales de Chile y los de los países comparados. En efecto, dado que las diferencias de resultados podían estar asociadas a diversas causas, se optó por implementar una metodología que permitiera descomponerlas.

La descomposición de Oaxaca-Blinder es la metodología más clásica para entender los factores asociados a diferenciales de resultados entre dos grupos de población¹⁰. Esta metodología permite atribuir las diferencias de resultados entre países a tres componentes distintos: por una lado está la diferencia en la magnitud de los factores que afectan el desempeño de los estudiantes (efecto *endowment* o *características*), por otro, el efecto que tiene un diferencial en la eficiencia en el uso de estos factores entre los países (efecto *eficiencia* o *retorno*) y, finalmente, un efecto combinado de las diferencias de características y las diferencias de retorno (efecto *interacción*).

A modo de ejemplo, la diferencia en los valores esperados de Uruguay respecto de Chile, en Matemática, puede ser expresada como:

$$[2] \quad \Delta \bar{Y}_m = E(Y_m^U) - E(Y_m^C)$$

donde $E(Y)$ indica el valor esperado de los resultados de PISA en el test de Matemática.

Si por un momento, no se considera la interacción de los residuos individuales a nivel escolar, el modelo

9 Ammermüller (2004) estima que para PISA 2000 este problema sobrestimó en 20 puntos el resultado promedio de Alemania.

10 Aunque desarrollada en estudios de economía laboral, también ha sido aplicada a diferencias de resultados medidos por pruebas de aprendizajes, como se hace en este estudio. Aplicaciones recientes en este sentido se encuentran en Ammermüller, 2004; McEwan y Marshall, 2004, y Sakellariou, 2008.

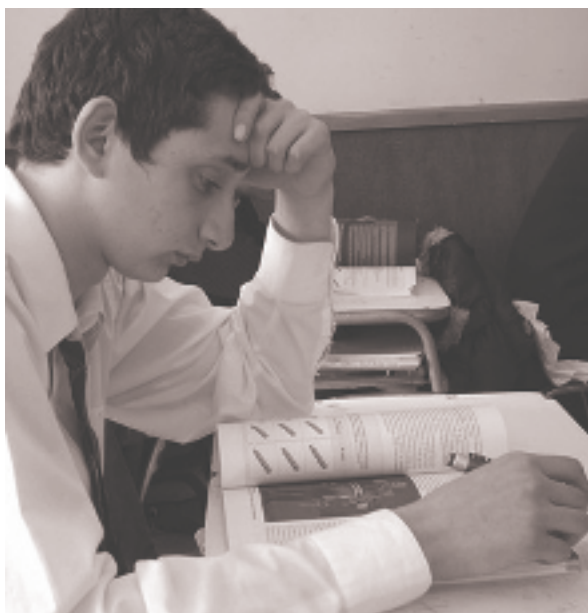
lineal de regresión, para cualquier país, podría ser expresado por:

$$[3] \quad Y_m^p = X_m^p \beta_m^p + \varepsilon_m^p, \text{ donde } E(\varepsilon_m^p) = 0, p \in (U, Ch)$$

En este caso el vector X contiene controles que afectan la producción de resultados educativos, tanto a nivel individual, como a nivel del establecimiento e institucional. Como fue descrito previamente, β contiene los parámetros asociados a estos controles, incluyendo un intercepto, y ε es el término de error. La diferencia de medias entre los resultados de los países puede ser expresada como la diferencia entre las predicciones lineales, evaluadas en los parámetros de uno de los países a analizar.

$$[4] \quad \Delta \bar{Y}_m = [E(X_m^U) - E(X_m^{Ch})] \beta_m^{Ch} + E(X_m^{Ch}) (\beta_m^U - \beta_m^{Ch}) + [E(X_m^U) - E(X_m^{Ch})] (\beta_m^U - \beta_m^{Ch})$$

En el ejemplo, la explicación de las diferencias de medias entre Uruguay y Chile, puede asociarse a tres componentes que, señalados de izquierda a derecha (después del signo igual) corresponden a: (a) la diferencia



en los predictores entre los dos países (efecto de las características); (b) la contribución de las diferencias entre los coeficientes (efecto de los retornos), incluyendo la diferencia entre los interceptos de ambas regresiones (diferencia de constantes), y (c) el efecto interacción, que da cuenta de que existen diferencias en los componentes y coeficientes de ambos grupos, que actúan simultáneamente¹¹. Como se verá más adelante, la diferencia de constantes también puede ser separada para distinguir su aporte a la diferencia observada entre los grupos analizados¹².

Ahora bien, dado que el interés sustantivo de este estudio era avanzar en la comprensión de los resultados de aprendizaje de los estudiantes chilenos, la descomposición [4] se consideró desde el punto de vista de Chile. Así, las diferencias entre predictores fueron ponderadas por los coeficientes de Chile. En otras palabras, este componente midió el cambio esperado en los resultados de Chile, en caso de que éste tuviera los predictores o características de Uruguay. Del mismo modo, las diferencias de coeficientes fueron ponderadas por el valor de las características de Chile, estimando así el cambio esperado de resultados de

11 Su interpretación posterior es relativamente formal: un signo positivo indica que ambos componentes se potencian, ya sea para aumentar o disminuir la brecha de resultados, mientras un signo negativo indica que ambos componentes se contrarrestan en la generación de la brecha que está siendo explicada.

12 Existe cierta discusión en la literatura técnica acerca de la adecuada interpretación del valor de la constante en este tipo de análisis, e incluso de la conveniencia de interpretarla (véase, por ejemplo, Oaxaca y Ransom, 1999). Como se sabe, su valor es afectado no sólo por las variables no observadas (y por tanto ausentes del modelo), sino que, en análisis como el nuestro (que incluye variables dicotómicas "dummy"), su valor es sensible a la categoría omitida en los modelos de regresión. Este hecho hace que la separación entre los componentes de "retorno" y "constante" en la descomposición, sea en cierto modo arbitraria. Con todo, nótese que esto no afecta a la estimación del componente asociado a las diferencias, en el nivel de las "características". Por ello, el modo más conservador de proceder es atribuir a la parte no explicada por las características, la noción de "no explicada", sumando en ese concepto, por defecto, tanto los retornos como las interacciones, la constante y las variables no observadas. En consecuencia, nuestro análisis de los resultados procederá de ambas formas: primero reportaremos los resultados "agregados" (diferencia explicada por las diferencias de características versus el resto) y luego, los "detallados" (distinguiendo la constante, los efectos de los coeficientes, y las interacciones, para cada variable).

Chile, en caso de que éste tuviera los coeficientes o retornos de Uruguay¹³.

Siguiendo con el ejemplo, para describir la descomposición de Oaxaca-Blinder entre uno de los tres países escogidos (en este caso, Uruguay) y compararla con los resultados chilenos, se consideró que:

[5]

$$\Delta \bar{Y}_m = \bar{Y}_m^U - \bar{Y}_m^{Ch}$$

Donde las barras reflejan los promedios de los puntajes obtenidos por los estudiantes de los respectivos países en la prueba de Matemática de 2006. En este caso, tal como se indicó previamente, la metodología de Oaxaca-Blinder permitió explicar la diferencia o brecha total existente entre las medias de los resultados PISA de Uruguay y Chile, por medio de: (a) las diferencias de las medias de las variables utilizadas en las funciones de producción de [1], (b) las diferencias de los coeficientes asociados a estas variables observadas (que corresponden a la productividad o efectividad de estos factores observados) y (c) la interacción de ambos elementos, tal como se describe en [6], donde se han considerado tanto las variables individuales

de los estudiantes como del establecimiento y las institucionales:

[6]

$$\begin{aligned} \Delta \bar{Y}_m = & (\bar{A}_m^U - \bar{A}_m^{Ch}) \hat{\beta}_{1,m}^{Ch} + (\bar{S}_m^U - \bar{S}_m^{Ch}) \hat{\beta}_{2,m}^{Ch} + (\bar{I}_m^U - \bar{I}_m^{Ch}) \hat{\beta}_{3,m}^{Ch} \\ & + (\hat{\beta}_{1,m}^U - \hat{\beta}_{1,m}^{Ch}) \bar{A}_m^{Ch} + (\hat{\beta}_{2,m}^U - \hat{\beta}_{2,m}^{Ch}) \bar{S}_m^{Ch} + (\hat{\beta}_{3,m}^U - \hat{\beta}_{3,m}^{Ch}) \bar{I}_m^{Ch} \\ & + (\bar{A}_m^U - \bar{A}_m^{Ch})(\hat{\beta}_{1,m}^U - \hat{\beta}_{1,m}^{Ch}) + (\bar{S}_m^U - \bar{S}_m^{Ch})(\hat{\beta}_{2,m}^U - \hat{\beta}_{2,m}^{Ch}) \\ & + (\bar{I}_m^U - \bar{I}_m^{Ch})(\hat{\beta}_{3,m}^U - \hat{\beta}_{3,m}^{Ch}) \end{aligned}$$

En definitiva, esta metodología permite descomponer la brecha de resultados a diferentes niveles de desagregación (llegando hasta cada variable por separado). En la sección de resultados, se seguirá este orden de menor a mayor nivel de especificidad.

2.4 Limitaciones de la metodología

A pesar de la capacidad explicativa de la descomposición de Oaxaca-Blinder, es importante considerar sus principales debilidades.

En primer lugar, como se dijo, este estudio buscó identificar algunos factores que “explican” las diferencias de resultados entre Chile y otros países; además, a lo largo del artículo se utiliza frecuentemente la expresión “explicar” para hacer referencia a las asociaciones entre las diferentes variables y los puntajes PISA obtenidos por los estudiantes. Ciertamente, ni el diseño de la prueba PISA, ni la metodología de este estudio permiten realizar inferencias causales acerca de estas asociaciones; sin embargo, en las ciencias sociales, las explicaciones causales no son las únicas admisibles. En el contexto de este diseño de investigación, “explicar” es identificar variables sistemáticamente asociadas con el resultado analizado (puntaje PISA) y poder atribuir a diferencias en dichas asociaciones, parte de las diferencias observadas entre los resultados de los grupos comparados (Chile y los demás países). Así, la técnica de Oaxaca-Blinder permite realizar una simulación, un ejercicio hipotético, que consiste en estimar las diferencias esperadas en el resultado de un grupo, en el supuesto de que éste

13 Como se sabe, los resultados de la descomposición dependen del grupo que se tome como referencia, lo cual ha hecho que algunos autores propongan usar los coeficientes obtenidos de una regresión común para ponderar el efecto estimado de las diferencias de características. Esta solución nos parece inadecuada para un análisis interesado en un país en particular (Chile), a diferencia, por ejemplo, del que realizan McEwan y Marshall (2004), quienes no tienen un interés particular en alguno de los dos países que comparan. En este sentido, nuestro estudio se asemeja más a la aplicación clásica de esta técnica de descomposición: se pregunta “si, y en qué medida, se discrimina a las mujeres (o a los negros) en el mercado laboral”, suponiendo que los hombres (o los blancos) no son el grupo discriminado. En otras palabras, la robustez de los resultados no está dada por la reversibilidad de las estimaciones. Por ejemplo, evaluar la fracción de la diferencia de los resultados de Chile con España en Matemática, implica preguntarse “si los estudiantes chilenos tuvieran los recursos que tienen los estudiantes españoles, cuántos puntos adicionales habrían obtenido en la prueba PISA Matemática, suponiendo que dichos recursos tendrían, en los estudiantes chilenos, el efecto que observamos en Chile (y no en España)”, lo cual es teóricamente adecuado a la pregunta de investigación.

tuviera las propiedades de otro. La contrastación de estas hipótesis exigiría la realización de experimentos evidentemente imposibles de ejecutarse en el contexto de este problema de investigación.

En segundo término, se debe recordar que los resultados de la descomposición de Oaxaca-Blinder se basan en estimaciones de modelos de regresión múltiple, por lo que comparte con ellos todas sus limitaciones (el potencial efecto de variables omitidas, la estimación eventualmente sesgada de algunos coeficientes, etc.), a las cuales –como se explicó– se agrega el hecho de que estos modelos deben ser comunes para los grupos que están siendo comparados. Además, como se sabe, la prueba PISA no tiene un modelo longitudinal de aplicación, a nivel de los estudiantes ni de los establecimientos, de forma que las estimaciones de los “efectos” de las variables explicativas en los resultados de los estudiantes, están basadas en mediciones de corte transversal, lo cual limita enormemente dichas estimaciones. Con todo, la riqueza de los datos recolectados por PISA hace de ésta una de las mejores fuentes de información disponibles en el campo a nivel mundial.

En tercer lugar, las decisiones sobre la elección del tipo de establecimiento (público o privado) entre los países, son sólo controladas por una variable dummy (con valores 1 para establecimientos públicos y 0 para establecimientos privados, en las funciones de producción de cada país y disciplina), que da cuenta de la diferencia de porcentajes entre estudiantes que escogen establecimientos públicos y privados entre países, así como del diferencial de coeficientes asociado al tipo de establecimiento. Sin embargo, es relevante tomar en cuenta que las funciones de producción podrían diferir entre los diversos tipos de oferentes de educación, así como entre países: este tipo de interacciones no son consideradas en este análisis¹⁴.

Finalmente, la principal limitación específica de esta metodología, es que se basa sólo en las diferencias de medias y no considera potenciales efectos diferenciados

a través de la distribución de los resultados entre los diversos países. Es decir, los factores considerados pueden ser más o menos relevantes para explicar las diferencias de resultados de estudiantes con distinto nivel de desempeño en PISA. Ciertamente, tampoco analiza el efecto de los factores no observados en las diferencias entre los países. El que el análisis se concentre en las medias y no considere la distribución completa de resultados, impide abordar preguntas de investigación interesantes, vinculadas, por ejemplo, a los temas de desigualdad. Con todo, tratándose de una línea de investigación muy poco explorada en Chile, es de interés comenzar por el tipo de análisis que de hecho, es el más usado en la literatura internacional sobre educación.

3. Resultados

3.1 Análisis comparativo de las características de estudiantes, establecimientos e institucionales de los países estudiados

Dado que son un componente sustancial del método de descomposición de resultados, a continuación se describen con cierto detalle las características de los estudiantes, de los establecimientos y las características institucionales de los cuatro países estudiados. En la Tabla 2 se presentan estadísticas descriptivas (media y desviación estándar) para cada una de las variables incluidas en el estudio, así como los resultados del test de diferencia de medias entre cada país y Chile¹⁵.

14 Para resolver problemas de tamaño muestral y comparabilidad entre países, se definieron dos tipos de establecimientos escolares en cada país: públicos (administrados directamente por alguna entidad estatal), y privados (de propiedad no estatal), aunque sin diferenciar entre los que reciben o no aportes financieros del Estado (como ocurre en Chile, entre los particulares subvencionados y los particulares pagados, respectivamente).

15 Nótese que la Tabla 2 y su análisis en la sección 3.1, incorporan algunas variables referidas a Lectura. Como se explicó, la descomposición de resultados de Lectura se presenta en el capítulo siguiente.

Tabla 2.
Estadísticas descriptivas de las variables del estudio (ponderadas por factor de expansión de estudiantes)

	Chile		España		Uruguay		Polonia	
	Media	Desv. estánd.	Media	Desv. estánd.	Media	Desv. estánd.	Media	Desv. estánd.
Puntaje PISA Lectura	442,5 (5,01)	95,9	460,8 (2,23)	85,6	412,5 (3,43)	113,0	507,5 (2,79)	94,8
Puntaje PISA Matemática	411,6 (4,59)	83,0	480,0 (2,33)	85,4	426,8 (2,61)	93,4	495,3 (2,44)	82,5
Variables a nivel del alumno								
Mujer (%)	46,10 (1,56)	49,9	49,44* (0,71)	50	51,17*** (0,95)	0,5	50,33** (0,75)	0,5
Años de escolaridad	9,78 (0,02)	0,64	9,53*** (0,01)	62,92	9,473*** (0,03)	1,01	8,954*** (0,01)	0,25
7° Grado (%)	0,90 (0,31)	9,4	0,1** (0,04)	2,9	7,47*** (0,90)	26,29	0,57 (0,16)	7,5
8° Grado (%)	3,10 (0,54)	17,3	7,04*** (0,47)	25,59	9,79*** (0,70)	29,72	3,84 (0,34)	19,2
9° Grado (%)	18,90 (1,00)	39,1	33,0*** (0,79)	47,02	17,31 (1,02)	37,83	95,19*** (0,42)	21,4
10° Grado (%)	71,00 (1,20)	45,4	59,8*** (0,87)	49,02	58,88*** (1,51)	49,21	0,40*** (0,07)	6,3
11° Grado (%)	6,10 (0,46)	23,9	0,03*** (0,02)	1,85	6,56 (0,63)	24,76	n/a	n/a
Retraso escolar (%)	18,70 (1,18)	39	5,1*** (0,37)	21,96	32,75*** (1,46)	46,93	2,78*** (0,30)	16,5
ISEC (nivel socioeconómico y cultural familiar)	-0,70 (0,06)	1,17	-0,31*** (0,03)	1,07	-0,51*** (0,03)	1,18	-0,30*** (0,02)	0,86
Altas expectativas laborales (%)	69,10 (1,72)	46,2	59,86*** (0,83)	49,02	72,33 (1,26)	44,74	54,98*** (1,07)	49,8
Expectativas laborales medias (%)	13,10 (0,55)	33,8	18,35*** (0,54)	38,71	13,98 (0,82)	34,68	22,81*** (0,76)	42
Expectativas laborales bajas (%)	17,80 (1,69)	38,2	21,79** (0,70)	36,22	13,70** (0,94)	34,38	22,21** (0,87)	41,6
Valoración rendimiento Mat.: muy importante (%)	77,30 (0,74)	41,9	55,66*** (0,65)	49,68	57,99*** (0,90)	49,36	35,78*** (0,81)	47,9
Valoración rendimiento Mat.: importante (%)	19,40 (0,64)	39,5	33,15*** (0,50)	47,07	33,91*** (0,77)	47,34	50,69*** (0,75)	50
Valoración rendimiento Mat.: poco/nada importante (%)	3,40 (0,30)	18	11,20*** (0,45)	31,53	8,11*** (0,62)	27,3	13,53*** (0,60)	34,2
Valoración rendimiento Leng.: muy importante (%)	72,51 (1,11)	41,9	47,49*** (0,65)	49,9	44,79*** (1,04)	49,7	41,85*** (0,74)	49,3
Valoración rendimiento Leng.: importante (%)	21,96 (0,94)	39,5	36,98*** (0,60)	48,3	40,14 *** (0,93)	49,0	46,58*** (0,77)	49,9
Valoración rendimiento Leng.: poco/nada importante (%)	5,53 (0,40)	18	15,53*** (0,44)	41,3	15,07*** (0,61)	35,8	11,57*** (0,49)	32,0

(Continuación en página siguiente)

(Continuación Tabla 2)

	Chile		España		Uruguay		Polonia	
	Media	Desv. estánd.	Media	Desv. estánd.	Media	Desv. estánd.	Media	Desv. estánd.
Variables a nivel del establecimiento								
ISEC de los compañeros (efecto par)	-0,70 (0,06)	0,87	-0,29*** (0,03)	0,60	-0,51*** (0,03)	0,76	-0,30*** (0,02)	0,42
Establecimiento complejo (%)	47,6 (2,48)	49,9	18,12*** (1,23)	38,52	21,62*** (1,26)	41,17	1,40*** (0,88)	11,7
Tamaño del curso	38,11 (0,45)	5,82	27,72*** (0,53)	9,61	33,11*** (0,57)	10,57	25,63*** (0,27)	3,72
Tamaño del establecimiento	1048,52 (53,55)	607,65	693,56*** (16,23)	346,12	435,16*** (11,32)	248,9	407,30*** (7,92)	228,01
Horas instrucción Matemática	3,48 (0,06)	2,25	3,42 (0,03)	1,48	3,30** (0,04)	1,79	4,36*** (0,04)	1,75
Horas instrucción Lenguaje	3,44 (0,06)	2,19	3,60** (0,03)	1,53	2,73*** (0,04)	1,69	4,64*** (0,03)	1,73
Alumnos por profesor	25,04 (0,52)	6,92	12,34*** (0,16)	4,31	15,87*** (0,31)	5,53	11,29*** (0,14)	2,25
Recursos educativos	-0,62 (0,10)	1,07	-0,02*** (0,05)	1,00	-0,71*** (0,07)	1,36	-0,09*** (0,07)	0,89
Variables a nivel institucional								
Autonomía uso recursos	-0,25 (0,07)	0,83	0,02*** (0,04)	0,88	-0,89*** (0,01)	0,47	-0,25 (0,04)	0,59
Autonomía en currículum y evaluación	-0,24 (0,07)	0,848	0,24*** (0,05)	0,81	-1,10*** (0,02)	0,46	0,31*** (0,06)	0,86
Selección académica (%)	36,0 (3,77)	48	4,64*** (1,04)	21,04	9,19*** (1,59)	28,89	15,18*** (2,52)	35,9
Mayoría de padres con expectativas de altos estándares (%)	18,60 (2,76)	38,9	8,81*** (1,72)	28,34	6,81*** (1,15)	25,19	22,92 (2,98)	42
Minoría de padres con expectativas de altos estándares (%)	52,80 (4,33)	49,9	31,21*** (2,50)	46,34	36,40*** (3,05)	48,12	51,43 (3,96)	50
Inexistente expectativas de los padres de altos estándares (%)	28,60 (3,79)	45,2	59,98*** (2,99)	49	56,79*** (3,03)	49,54	25,65 (3,71)	43,7
Establecimiento público (%)	43,80 (2,05)	49,6	64,61*** (0,88)	47,82	85,13*** (0,83)	35,58	98,41*** (0,11)	12,5
Número de observaciones	5.224		19.604		4.839		5.528	

Los coeficientes son estadísticamente significativos al *** < 1%; ** < 5%; * < 10%

3.1.1 Características de los estudiantes

Intuitivamente, la primera hipótesis para explicar diferencias de resultados entre dos grupos de estudiantes, es que ambos grupos difieren en características relevantes, asociadas con la adquisición de aprendizajes. Esta hipótesis es aún más plausible en estudios que, como PISA, consisten en comparaciones de muestras de estudiantes provenientes de diferentes países. En consecuencia, la primera tarea del análisis es comparar las características de los estudiantes chilenos con los estudiantes de España, Uruguay y Polonia.

Una simple inspección de la Tabla 2 permite constatar que los estudiantes chilenos efectivamente difieren de sus pares de otros países en varios aspectos potencialmente relevantes para explicar las diferencias de resultados. Así por ejemplo, los estudiantes chilenos tienden a estar cursando grados superiores en relación con sus pares de otros países: aproximadamente 3 de cada 4 estudiantes chilenos evaluados asistía a su 10° ó 11° año de escolaridad al momento de rendir la prueba PISA, mientras que en Polonia, prácticamente no existen estudiantes que a los 15 años de edad hubieran alcanzado esos niveles de estudio (casi la totalidad de los estudiantes polacos participantes estaba matriculado en 9° Grado). La situación de Uruguay y España es intermedia respecto de estos dos extremos: entre la mitad y dos tercios de sus estudiantes se encontraba cursando los grados 10° ó 11°. El efecto acumulado de estas diferencias de grado no es menor: los estudiantes chilenos evaluados por PISA tenían en promedio 9,8 años de escolaridad, mientras que los españoles y uruguayos 9,5 y los polacos sólo 9 años de estudio. Las diferencias en años de escolaridad y grado al que asisten los estudiantes no se explican sólo por diferentes regulaciones de la edad de ingreso a los establecimientos¹⁶, sino por diferencias en las prácticas de reprobación, expresadas

en este caso en los diferentes niveles de retraso escolar observado. Como puede verse, casi un tercio de los estudiantes uruguayos de 15 años presenta retraso escolar, mientras que este fenómeno prácticamente no existe en Polonia y es muy menor en España. Chile tiene una situación intermedia, aunque más cercana al caso uruguayo: aproximadamente un quinto de los estudiantes chilenos evaluados presentaba algún nivel de retraso escolar.

Otra variable individual relevante asociada con los resultados, es el nivel socioeconómico familiar. En la Tabla 2 se muestra que Chile comparte con los demás países de referencia, el tener un nivel socioeconómico promedio inferior al promedio de los países participantes en PISA (valor ISEC negativo); sin embargo, las diferencias son marcadas entre estos cuatro países, especialmente entre los países europeos y latinoamericanos. En este sentido, los estudiantes chilenos se encuentran en una situación desfavorable respecto de sus pares: el ISEC chileno es en promedio 0,2 desviaciones estándar inferior al ISEC uruguayo y 0,4 desviaciones estándar inferior al ISEC español y polaco.

Una última diferencia notoria entre los estudiantes chilenos y los de los otros países, es la mayor valoración que los chilenos asignan a tener un buen rendimiento, tanto en el área de Matemática como de Lenguaje. Así, mientras 3 de cada 4 estudiantes chilenos considera que es muy importante rendir bien en Matemática, sólo 1 de cada 3 estudiantes polacos comparte esta valoración; los estudiantes españoles y uruguayos se encuentran en una situación intermedia entre ambos extremos (poco más de la mitad de los estudiantes de estos países cree que es muy importante). En el caso de Lenguaje la situación es similar: mientras la gran mayoría de los estudiantes chilenos considera muy importante rendir bien en esta asignatura (73%), menos de la mitad de los estudiantes españoles, uruguayos y polacos comparte esta valoración (48%, 45%, y 42%, respectivamente).

¹⁶ Diferencias que efectivamente existen. Véase la descripción de la variable "años de escolaridad", en el Anexo.

Aunque menos notorias, en la Tabla 2 también se pueden observar diferencias entre Chile y los demás países estudiados en otras características potencialmente relevantes, tales como la distribución por género (menos mujeres rindieron la prueba en Chile) o las expectativas laborales que poseen (más altas en Chile que en los países europeos, aunque similares a Uruguay).

En síntesis, las diferencias observadas en todas las variables individuales entre Chile y los otros tres países, son estadísticamente significativas (al menos al nivel $\alpha < 0.1$, aunque casi siempre al nivel $\alpha < 0.01$), con la única excepción de las expectativas laborales, en la comparación con Uruguay.

3.1.2 Características de los establecimientos e institucionales

La investigación comparada muestra que, a la base de las brechas de resultados escolares entre países, puede haber diferencias en las características y condiciones de los establecimientos, así como en otros aspectos institucionales que diferencian a los sistemas escolares. Los resultados sobre algunas de estas características y condiciones medidas por PISA, se presentan en la Tabla 2. Allí se puede observar que, en efecto, Chile presenta varias diferencias importantes con los países de referencia.

A nivel de las condiciones institucionales, por ejemplo, Chile posee un sistema escolar altamente privatizado en comparación con los demás países: menos de la mitad de los estudiantes chilenos que rindieron la prueba PISA, estudia en establecimientos públicos, lo que representa una proporción muy inferior a la de Polonia, donde prácticamente todos los estudiantes asisten a establecimientos públicos; pero también inferior a Uruguay, donde la gran mayoría asiste a este tipo de establecimientos, y a España, donde dos de cada tres estudiantes evaluados estudian en el sistema público.



Chile también se diferencia sensiblemente de los demás países en el grado de relevancia que tiene la selección de los estudiantes, basada en criterios académicos, en los procesos de admisión. Aproximadamente un tercio de los estudiantes chilenos evaluados en PISA 2006 asiste a establecimientos en los que se aplican criterios académicos en la selección de estudiantes, proporción que es más del doble de la observada en Polonia, cuatro veces mayor que la de Uruguay, y ocho veces superior a la de España. Es decir, los estudiantes chilenos asisten, en mayor medida que sus pares de los otros países, a establecimientos académicamente segregados.

Otras características de contexto también muestran diferencias entre los países, aunque de modo menos sistemático. Así, según lo reportado por los directores, los padres chilenos tienden a ejercer una mayor presión sobre los establecimientos por lograr altos estándares que los padres españoles y uruguayos, aunque la presión de los chilenos es semejante a la de los padres polacos. Del mismo modo, los establecimientos chilenos tienen menor autonomía en materias curriculares que

los establecimientos polacos y españoles, pero mayor que los uruguayos. Finalmente, los establecimientos chilenos tienen mayor autonomía que los uruguayos en el manejo de recursos, inferior a los españoles, y semejante a los polacos.

En cuanto a las características más directamente vinculadas a las condiciones en que se realizan los procesos de enseñanza-aprendizaje, se observa que casi la mitad de los estudiantes chilenos que rindió PISA en 2006 estudia en establecimientos educacionales que imparten cursos de enseñanza primaria y secundaria (“establecimientos complejos”); esto es, más del doble que en España y Uruguay, y significativamente mayor que en Polonia, donde este tipo de establecimientos prácticamente no existe. Esto se asocia además con el hecho de que, en promedio, los establecimientos chilenos tienen una matrícula notoriamente mayor a la de los establecimientos de los demás países.

Otra característica que diferencia marcadamente a los establecimientos chilenos del resto, es la comparativamente menor disponibilidad de docentes, medida tanto por el tamaño del curso como por el promedio de estudiantes por profesor. Esto implica que, en promedio, los docentes chilenos tienen a su cargo (tanto en la sala de clases como en el establecimiento en general) una mayor cantidad de estudiantes que sus colegas polacos, españoles y uruguayos.

Además de los docentes, existen otros dos recursos críticos para la enseñanza escolar: los materiales educativos y el tiempo de instrucción. En cuanto a los recursos educativos, se aprecia una clara distinción entre los países europeos y latinoamericanos: Chile y Uruguay poseen establecimientos menos equipados que Polonia y España. En cuanto al tiempo de instrucción, los estudiantes chilenos (al igual que los españoles y uruguayos) tienen en promedio prácticamente una hora menos de clases de Matemática a la semana, que sus pares polacos. En relación al tiempo de instrucción de Lenguaje (ver próximo capítulo), los estudiantes polacos nuevamente superan a los chilenos (en 1,2

horas semanales), pero éstos a su vez tienen 0,7 horas más de instrucción semanal en Lenguaje que los uruguayos¹⁷.

De acuerdo con lo anterior, se puede afirmar que, en términos de las variables asociadas al establecimiento y al sistema educativo, Chile difiere en forma estadísticamente significativa (nivel $\alpha < 0.01$) de los otros tres países, en prácticamente todas estas variables. Las únicas excepciones son que no hay diferencias con España en las horas de instrucción de Matemática, y que no hay diferencias con Polonia en el nivel de autonomía de los establecimientos relativa al uso de los recursos y al nivel de presión de los padres por lograr altos estándares.

En síntesis, considerando el análisis de las variables individuales, de los establecimientos y las relativas a las condiciones institucionales de los sistemas escolares, se puede señalar que efectivamente existen diferencias relevantes entre Chile y los demás países. Particularmente, destaca el hecho de que, en promedio, los estudiantes chilenos poseen el más bajo nivel socioeconómico y cultural familiar y asisten a establecimientos que cuentan con menos recursos educativos y una menor dotación docente, características todas teóricamente asociadas a menores oportunidades de aprendizaje. Todo lo anterior da sustento a la hipótesis de que una cierta fracción de las diferencias de resultados de la prueba PISA 2006 entre los estudiantes chilenos y los de los tres países en comparación, puede estar asociada a las mencionadas diferencias en características individuales, escolares e institucionales.

17 En la Tabla 2, se incluye el valor agregado a nivel del establecimiento, del índice de nivel socioeconómico y cultural ISEC, el cual se ha considerado una medición aproximada (aunque obviamente parcial) del “efecto par”. En la Tabla 2 se presentan los promedios nacionales de las variables: estos valores son similares a los presentados a nivel individual. Sin embargo, para comprender los análisis posteriores, debe recordarse que esta variable es diferente al ISEC individual, pues se trata en este caso del ISEC promedio de los compañeros del mismo establecimiento (valor que presenta, por tanto, pequeñas variaciones entre compañeros de un mismo establecimiento).

3.2 Factores explicativos de los resultados

Atribuir las diferencias en los resultados entre los estudiantes chilenos y sus pares de otros países solamente a diferencias en las características que poseen los estudiantes y sus establecimientos, es una hipótesis difícil de defender. En efecto, esto supondría afirmar que si se igualasen dichas condiciones (por ejemplo, el nivel socioeconómico familiar, la cantidad de estudiantes por curso, etc.), los estudiantes de los diferentes países obtendrían en promedio el mismo puntaje en las pruebas estandarizadas del tipo PISA. La evidencia demuestra que esto, de hecho, no sucede¹⁸.

Una hipótesis alternativa, por tanto, es que los países difieren, para decirlo en términos generales, en el modo en que educan a su población estudiantil. Más específicamente, los procesos institucionales de gestión escolar y de enseñanza-aprendizaje, que se dan en cada contexto nacional, pueden diferir de forma tal que, características de los estudiantes e insumos educativos semejantes, resulten en logros de aprendizaje diferentes. En otras palabras, en el contexto de nuestro estudio, el hecho de que los estudiantes chilenos logren, en promedio, resultados más bajos que los estudiantes de los países comparados, puede explicarse no sólo porque poseen características y condiciones diferentes, sino porque dichas características y condiciones son “transformadas” en resultados de aprendizaje de un modo diferente en Chile, que en los demás países.

Una forma de indagar acerca de la verosimilitud de esta hipótesis alternativa es utilizar la metodología de

18 Países de condiciones similares obtienen en promedio resultados muy diferentes en pruebas como TIMSS y PISA. Asimismo, al interior de los países, se ha demostrado que, por ejemplo, los establecimientos efectivos, obtienen resultados notablemente superiores a los que cabría esperar, basándose en las condiciones en las que ejercen su trabajo.



regresiones múltiples, para saber si las características analizadas se asocian sistemáticamente, y en qué magnitud, a los resultados PISA de Matemática obtenidos por los estudiantes, y comparar luego si dichas asociaciones son comunes o difieren entre los países estudiados. En la Tabla 3 se muestran los resultados de este análisis. El primer panel presenta los coeficientes estimados para las variables a nivel del estudiante, el segundo, a nivel del establecimiento, y el tercero, a nivel institucional. Los resultados de Chile se presentan en la segunda columna, y los de los demás países, en las columnas 3, 4, y 5. A continuación, se procederá, en primer lugar, con un breve análisis de los resultados de Chile, para luego continuar con un análisis comparativo.

Como puede apreciarse, en Chile, la mayoría de las variables que miden características de los estudiantes tienen una asociación estadísticamente significativa con los resultados obtenidos por esos estudiantes en la prueba PISA Matemática. Así, por ejemplo, en promedio, las mujeres chilenas obtienen 30 puntos menos que sus compañeros hombres. El grado que se encuentra cursando el alumno es también un factor

asociado a los resultados: cada grado adicional está asociado a una diferencia positiva de puntajes que va desde 17 hasta 41 puntos (8° versus 7°, categoría omitida, y 10° versus 9°, respectivamente). También son factores relevantes el nivel socioeconómico y cultural de la familia, y aspectos más subjetivos, como la valoración que el alumno tiene acerca de la importancia de rendir bien en Matemática y su nivel de expectativas laborales.

En contraste, la mayoría de las variables que miden características de los establecimientos a los que asisten los estudiantes chilenos, no parecen ser factores relevantes para explicar sus resultados. Tal es el caso del tamaño del establecimiento, el que posea los niveles básico y medio, la cantidad de alumnos por profesor, el tamaño del curso, y el nivel de recursos educativos. La excepción la constituyen el nivel socioeconómico promedio de los compañeros del establecimiento (“efecto par”) y la cantidad de horas semanales de clases de Matemática planificadas. Es especialmente fuerte el efecto del nivel socioeconómico de los compañeros: es casi cuatro veces mayor al efecto asociado al mismo indicador, pero medido en términos individuales (es decir, de la propia familia del alumno).

Finalmente, tres de las cinco variables que miden aspectos institucionales, presentan una asociación estadísticamente significativa con los resultados de los estudiantes chilenos. Así, los estudiantes obtienen en promedio puntajes más altos, cuando los padres tienen expectativas por altos estándares educativos y presionan al establecimiento para lograrlos, cuando los establecimientos aplican procedimientos de selección académica en los procesos de admisión de estudiantes, y cuando son públicos¹⁹.

Es importante notar que el conjunto de factores analizados permite explicar poco más de la mitad (53%) de la variación de resultados de los estudiantes chilenos en la prueba PISA-Matemática, lo cual constituye una fracción muy elevada para este tipo de metodologías, especialmente considerando que la unidad de análisis es el alumno y no el establecimiento.



19 El que, una vez controlando por otras características relevantes, los estudiantes chilenos que estudian en establecimientos públicos obtengan mejores resultados en la prueba PISA, no debería sorprendernos. Porejemplo, Bellei (2009) demuestra que los resultados de los estudios que comparan la efectividad de los establecimientos públicos y privados en Chile son muy sensibles a las variables de control que se utilicen; en particular, una fracción relevante de la ventaja de resultados obtenidos por los establecimientos privados se explicaría por los procedimientos de selección de estudiantes y por las mejores condiciones socioeconómicas y culturales de las familias: cuando ambos factores son controlados, los establecimientos privados resultan ser, en promedio, menos efectivos que los públicos. Afortunadamente, las bases de datos de PISA permiten controlar también por ambos factores.

Tabla 3.

Factores a nivel de estudiante, establecimiento e institucionales, explicativos de los resultados PISA Matemática ^a

	Coeficiente			
	Chile	España	Uruguay	Polonia
VARIABLES A NIVEL DEL ESTUDIANTE				
Mujer	-30,11*** (2,87)	-23,36*** (1,83)	-26,41*** (3,44)	-21,61*** (2,59)
8° Grado	17,31* (16,39)	138,17** (47,80)	-1,10 (8,38)	14,84 (15,92)
9° Grado	52,06*** (15,11)	185,43*** (49,82)	28,06*** (7,31)	86,20*** (16,71)
10° Grado	93,26*** (15,07)	254,72*** (50,00)	63,67*** (11,68)	158,76*** (19,02)
11° Grado	118,05*** (14,51)	395,57*** (54,37)	78,59*** (12,54)	N/A
Retraso escolar	-2,74 (5,11)	-4,99 (7,49)	-26,37*** (8,55)	-16,82 (10,64)
ISEC (nivel socioeconómico y cultural familiar)	8,70*** (1,13)	8,55*** (1,15)	13,28*** (1,69)	22,95*** (1,19)
Altas expectativas laborales	16,81*** (3,98)	22,39*** (3,13)	9,81** (4,82)	46,03*** (2,70)
Expectativas laborales medias	7,99* (4,38)	9,76** (3,44)	6,27 (5,59)	21,42*** (2,77)
Valoración rendimiento Mat.: muy importante	14,19** (5,77)	36,56 (3,47)	28,72*** (6,11)	51,48*** (3,64)
Valoración rendimiento Mat.: importante	5,95 (6,24)	18,35*** (3,27)	16,58** (5,61)	19,50*** (3,36)
VARIABLES A NIVEL DEL ESTABLECIMIENTO				
ISEC de los compañeros (efecto par)	31,07*** (4,34)	15,50*** (3,60)	24,70*** (4,83)	13,44** (5,85)
Establecimiento complejo	6,39 (5,85)	6,39 (4,97)	14,96*** (4,41)	-0,10 (10,66)
Tamaño del curso	0,24 (0,39)	-0,01 (0,13)	0,50** (0,20)	0,11 (0,70)
Tamaño del establecimiento	0,00 (0,00)	-0,01 (0,00)	0,00 (0,01)	-0,01 (0,01)
Horas instrucción Matemática	6,31*** (0,48)	4,55*** (0,82)	5,81*** (0,81)	5,54*** (0,62)
Alumnos por profesor	-0,20 (0,33)	0,62 (0,71)	-0,39 (0,49)	1,25 (1,01)
Recursos educativos	2,13 (2,24)	2,39 (1,80)	-0,63 (2,00)	0,35 (2,16)

(Continuación en página siguiente)

(Continuación Tabla 3)

	Coeficiente			
	Chile	España	Uruguay	Polonia
Variables a nivel institucional				
Autonomía uso recursos	4,26 (3,07)	1,08 (3,04)	5,85 (4,35)	-1,09 (3,33)
Autonomía en currículum y evaluación	-2,59 (2,26)	-2,46 (2,08)	-3,42 (5,36)	-3,53 * (2,05)
Selección académica	14,79*** (4,37)	-3,66 (9,22)	11,09 (7,00)	14,76*** (5,50)
Expectativas de padres por altos estándares	17,22*** (6,66)	0,35 (6,44)	7,05 (5,80)	-1,09 (3,98)
Establecimiento público	11,56** (5,44)	11,56 (7,47)	15,81 (11,70)	-3,53 (17,91)
Constante	301,43*** (20,08)	208,28*** (52,71)	342,00*** (16,02)	339,41*** (24,39)
Número de observaciones	5.224	19.604	4.839	5.528
R ² Promedio ^b	0,53	0,43	0,43	0,33

^a: Coeficientes estimados en regresiones separadas para cada país

^b: Corresponde al valor promedio pues se realizan 5 regresiones, una por cada valor plausible (*plausible value*)

Los coeficientes son estadísticamente significativos a *** < 1%; ** < 5%; * < 10%

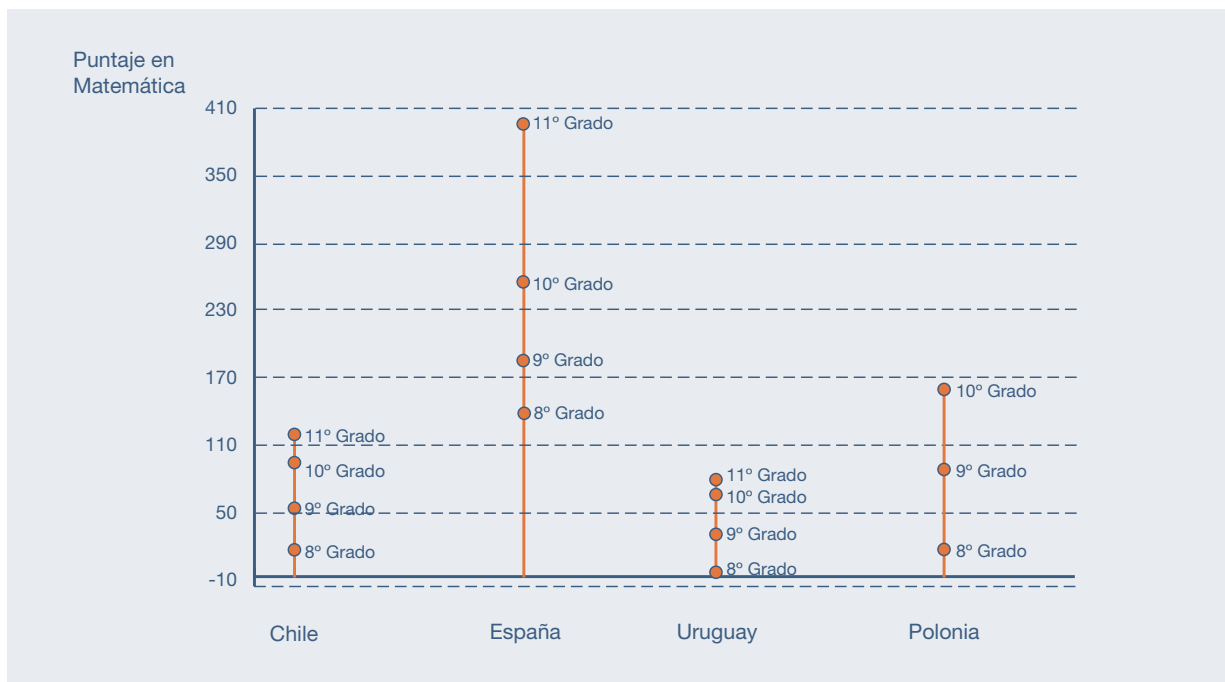
A partir de los datos de la Tabla 3 también se pueden comparar, entre Chile y los países de referencia, los factores asociados a los resultados de los estudiantes y la magnitud de los coeficientes estimados para dichos factores.

En general, el patrón de comportamiento de las variables de caracterización individual de los estudiantes es relativamente similar en los cuatro países, tanto en términos de su sentido (signo positivo o negativo, es decir, mayor o menor puntaje PISA, respectivamente) como de sus tendencias (estructura en variables como el grado, las expectativas laborales y la valoración de las matemáticas). Sin embargo, un análisis más detallado revela diferencias importantes entre los países. Tal es el caso, por ejemplo, del grado o nivel escolar al que asisten los estudiantes al momento de ser evaluados por PISA.

Para mayor claridad, en el Gráfico 1 se muestran los coeficientes estimados para cada grado, en los cuatro

países (recuérdese que todos ellos están en referencia al 7° Grado). Como se ve, en todos los países, los estudiantes de los grados 8° a 11° obtienen en promedio mejores puntajes que sus compañeros de 7° Grado (salvo los uruguayos de 8°), y esta diferencia aumenta sistemáticamente grado a grado, en todos los casos. Sin embargo, la magnitud del efecto asociado a las diferencias de grado entre los estudiantes que rindieron la prueba, es notoriamente diferente entre los cuatro países. El caso de España constituye un extremo: la diferencia de puntaje entre los estudiantes de 8° y 7° grados es aproximadamente ocho veces superior a la diferencia observada entre esos dos grupos, en Chile. En los cursos superiores (11° y 10°), la diferencia de puntaje entre los estudiantes es casi seis veces mayor en España que en Chile. Uruguay es el país en donde las diferencias de resultados entre los estudiantes de diferentes grados son menores y en Polonia, dichas diferencias tienen valores intermedios entre los de España y Chile.

Gráfico 1: El grado como factor explicativo de los resultados PISA Matemática en Chile, España, Polonia y Uruguay*.



* Diferencia neta estimada entre estudiantes de diferentes grados (ver Tabla 3 para valores exactos; categoría de referencia: alumnos de 7° Grado)

Un patrón similar al de los grados escolares, se observa para el caso de las dos variables subjetivas consideradas a nivel del alumno (las expectativas laborales y la valoración del buen rendimiento en Matemática): los cuatro países comparten la estructura de los efectos estimados, pero la magnitud de dichos efectos varía enormemente entre ellos, siendo en este caso Polonia, el país en que las diferencias asociadas a estas variables son mayores. Finalmente, el efecto estimado que tiene el nivel socioeconómico familiar del alumno en el rendimiento, es en Chile muy similar al de España, pero claramente menor al de Uruguay y Polonia. En este último país, el efecto estimado del ISEC individual es aproximadamente 2,5 veces mayor que en Chile.

Los resultados asociados a las variables del establecimiento, en cambio, no son tan heterogéneos entre los países: al igual que en el caso chileno,

en España y Polonia sólo el nivel socioeconómico promedio de los compañeros y las horas semanales de instrucción en Matemática, presentan una asociación estadísticamente significativa con los resultados. En Uruguay estos dos factores también son relevantes, pero a ellos se suman el tamaño del curso y el que el establecimiento imparta clases en primaria y secundaria. Más aún, el efecto estimado asociado a las horas semanales de clases de Matemática es bastante similar en los cuatro países.

De este conjunto de variables, el factor escolar más relevante es el “efecto par” y en Chile es donde mejor se aprecia esta asociación: la magnitud de su efecto estimado es más del doble que en Polonia, y el doble que en España. A este respecto, conviene hacer notar que el efecto estimado en los resultados de aprendizaje del nivel socioeconómico familiar, medido a nivel del establecimiento, es relativamente independiente

del efecto de esta misma variable medida a nivel individual. Así, mientras Polonia tiene el más fuerte efecto entre estos cuatro países a nivel individual, tiene a su vez el más pequeño efecto a nivel agregado; es decir, en Polonia los resultados de aprendizaje de los estudiantes están más asociados al ISEC de la propia familia que al de los compañeros. Esta situación es opuesta a la de Chile, donde el ISEC de los compañeros es comparativamente más relevante para predecir los resultados, que el del propio alumno. El caso español demuestra que esto no es simplemente un juego de suma-cero, por cuanto en ese país ambos efectos del ISEC (propio y de los compañeros) son comparativamente más bajos que en el resto de los países, demostrando que es posible desvincular en mayor medida, los resultados de aprendizaje de los estudiantes, de su origen socioeconómico y cultural y del contexto socioeconómico y cultural de su establecimiento.

Finalmente, los efectos asociados a las variables institucionales son comparativamente más heterogéneos entre países, que los de las variables individuales y escolares.

Así, tanto en España como en Uruguay, ninguna de las cinco variables incluidas resultó ser un factor asociado significativamente con las diferencias de resultados. A su vez, Polonia es el único país en que la autonomía de los establecimientos en materias curriculares y de evaluación posee una asociación estadísticamente significativa con los puntajes PISA en Matemática, pero inversa (en promedio, a mayor autonomía, peores resultados de los estudiantes)²⁰. Por otro lado, en Chile y en Polonia, la selección académica de los estudiantes está asociada con mayores puntajes promedio y en una magnitud similar. Por último, recuérdese que en Chile hay dos factores adicionales explicativos de los resultados: las expectativas de los padres por altos estándares de calidad educativa

(que tiene una asociación positiva con los puntajes, que no se advierte en los demás países, salvo en Uruguay, aunque en una magnitud claramente menor y estadísticamente no significativa) y la dependencia o propiedad del establecimiento: los establecimientos públicos chilenos son en promedio más efectivos que los establecimientos privados, en lograr resultados en Matemática, asociación que se repite tanto en España como en Uruguay, con magnitudes similares al caso chileno, aunque en estos países, sin alcanzar el nivel de significación estadística.

En síntesis, el análisis anterior permite concluir que muchas de las características individuales, escolares e institucionales estudiadas son factores relevantes para explicar las diferencias de resultados PISA en Matemática en los cuatro países considerados; adicionalmente, se puede afirmar que, aunque en la mayoría de los casos el sentido y estructura de la asociación estimada entre estos factores y los resultados de los estudiantes es similar entre los cuatro países, ellos varían sensiblemente en la magnitud de los efectos estimados para las diferentes variables. Finalmente, es importante considerar que existen algunas variables cuya asociación con los resultados de aprendizaje de los estudiantes es poco sistemática, encontrándose sólo en algunos de los países, y a veces con signos opuestos entre sí. Más aún, como quedó claro, Chile presenta varias particularidades en términos de los factores asociados a los logros de aprendizaje de los estudiantes.

En el contexto de nuestro estudio, estos hallazgos son relevantes, pues dan sustento a la hipótesis de que los países no sólo difieren en sus características y condiciones, sino también en los efectos que dichas características y condiciones tienen sobre los resultados de aprendizaje de los estudiantes, todo lo cual hace viable y relevante, el análisis tendiente a identificar cuánto del menor logro de resultados en la prueba PISA Matemática de los estudiantes chilenos, es atribuible a una u otra fuente.

20 Nótese, sin embargo, que en los demás países, el signo y magnitud de este coeficiente es similar, aunque no alcanzó el umbral de significatividad estadística establecido.

3.3 Descomposición de resultados de PISA Matemática

3.3.1 Descomposición general de la diferencia de resultados entre Chile y los tres países de referencia

Como se indicó antes, Chile tiene una diferencia negativa en sus resultados en la prueba PISA Matemática 2006, respecto de los tres países de referencia: en promedio, los estudiantes españoles superan a los chilenos por 68 puntos, los uruguayos, por 15 puntos, y los polacos, por 84, pero ¿cuánto de esa brecha es atribuible a diferencias en las características de los países?, y, complementariamente, ¿cuánto de esa brecha puede atribuirse al hecho de que en Chile el efecto de algunas características en los resultados de aprendizaje es diferente que en los otros países?

En el Gráfico 2 (ver Tabla 4 para los valores exactos), se muestran los resultados básicos del análisis que permite responder ambas preguntas. El gráfico contiene las comparaciones de Chile con España, Uruguay y Polonia, respectivamente²¹. La barra celeste indica la brecha de resultados que separa a Chile de cada país analizado, cuya descomposición se está realizando. Las barras naranja y azul indican la fracción de esta brecha que es atribuible a las diferencias en características y a las diferencias en efectos, respectivamente²².

Como puede apreciarse, al comparar Chile con España se observa que, tomadas como un todo, las diferencias en las características entre ambos países no explican ninguna fracción de la brecha de resultados. Más aún, el valor asociado a este conjunto de variables, aunque muy pequeño, es de signo negativo (-4 puntos), lo

que sugiere que las diferencias en las características deberían favorecer a los estudiantes chilenos. En otras palabras, si los estudiantes y los establecimientos chilenos tuviesen las características observadas en los españoles, el promedio chileno habría sido 4 puntos menor al que se obtuvo efectivamente y, por tanto, la brecha con España habría sido aún mayor. De acuerdo con este análisis, no es posible explicar el menor rendimiento de Chile respecto de España, porque los estudiantes o los establecimientos chilenos posean (en promedio) características y condiciones menos conducentes al aprendizaje.

Este mismo patrón se repite para las comparaciones con Uruguay y Polonia: en ambos casos, no es posible atribuir la brecha de resultados a favor de estos países, a que ellos posean características o condiciones más favorables para producir mejores logros de aprendizaje en sus estudiantes. Aunque en términos absolutos el mayor contraste de Chile se da con Polonia (menos 20 puntos), en términos relativos la diferencia más marcada es con Uruguay: si Chile tuviese las características de Uruguay, el promedio nacional habría sido 12 puntos menor al obtenido, es decir, la brecha con Uruguay habría aumentado en un 78%.

En los análisis de descomposición ha sido frecuente atribuir el total de la brecha no explicada por las diferencias en características observadas entre los grupos, a una explicación “residual”, coherente con la hipótesis del estudio: así, por ejemplo, en los estudios sobre las diferencias salariales en el mercado de trabajo entre hombres y mujeres, o entre blancos y negros, esta fracción se asimila a la discriminación de que son objeto los grupos desfavorecidos. Siguiendo esta lógica, en el contexto de esta investigación, se debería concluir que el total de la diferencia de resultados en la prueba PISA Matemática, entre Chile y los tres países de referencia, se puede atribuir a diferencias en los efectos o retornos que las distintas características (observadas y no observadas) tienen sobre los resultados de aprendizaje de los estudiantes

21 Todos los resultados son reportados en la misma unidad original de puntajes PISA Matemática 2006.

22 La interpretación de la barra azul (diferencia atribuible a diferencias en “efectos”) es un poco más compleja que lo aquí dicho. Se retomará este punto en la próxima sección, cuando se presenten los resultados de la descomposición detallada.

(barra azul del Gráfico 2). En este caso, Chile tendría, por así decirlo, una menor capacidad que España, Polonia y Uruguay, para “transformar” las condiciones y características de sus estudiantes y establecimientos en resultados de aprendizaje. En otras palabras, el sistema escolar chileno sería el menos eficiente de este grupo.

A continuación, se profundizará en este aspecto.

Gráfico 2: Descomposición de la diferencia de resultados PISA Matemática entre Chile y los tres países de referencia

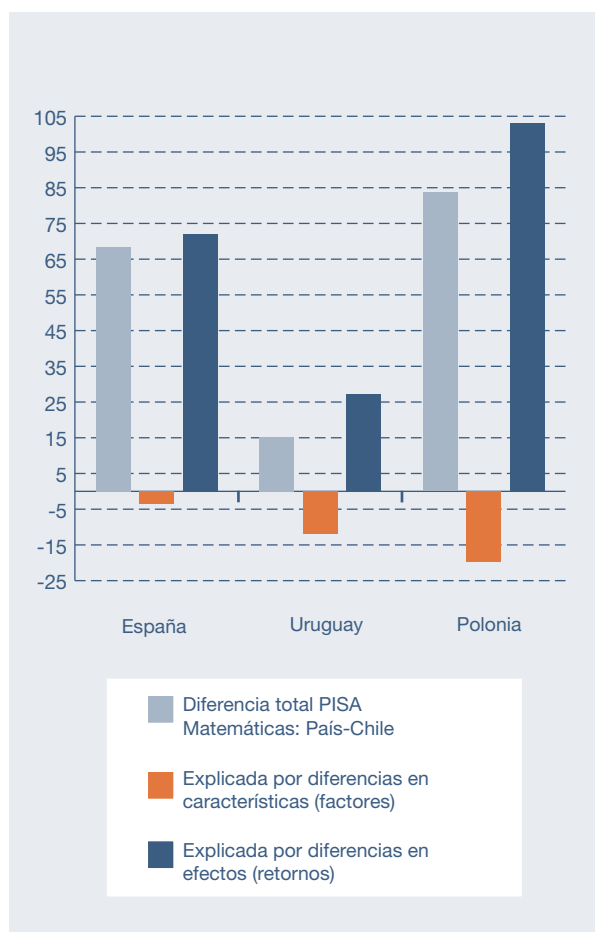


Tabla 4.

Descomposición general de la diferencia entre Chile y los tres países de referencia en PISA Matemática

	España	Uruguay	Polonia
Diferencia total PISA Matemática: País - Chile	68,4	15,2	83,7
Explicada por diferencias en características (factores)	-3,5	-11,9	-19,6
Explicada por diferencias en efectos (retornos)	71,9	27,1	103,2
Proporción diferencia explicada (factores)	-5%	-78%	-23%
Proporción diferencia no explicada (retornos)	105%	178%	123%

3.3.2 Descomposición detallada de la diferencia de resultados entre Chile y los tres países de referencia

Como se explicó, la metodología de descomposición de las brechas de resultados permite no sólo estimar la fracción de la diferencia atribuible a las distintas características y la fracción atribuible a las diferencias de efectos o productividad de esas características, sino también desagregar y detallar aún más ambos componentes. Así, se puede estimar la contribución específica que realiza a la brecha observada de resultados, cada variable incorporada al modelo de análisis, distinguiendo si este aporte lo realiza por las diferencias en el nivel (“características”) o en el efecto que tiene (“retorno”) entre los dos países bajo comparación. También es posible estimar el aporte simultáneo de las diferencias de características y de efectos o retornos, lo cual es técnicamente identificado como un componente de interacción entre ambos términos²³. Las diferencias de constante también

²³ En los análisis de descomposición que no realizan esta distinción, los efectos de interacción se consideran parte de la diferencia asociada a diferentes “retornos”; tal como fue presentado en el Gráfico 2 y en la Tabla 3.

pueden separarse para el análisis. Los resultados de esta descomposición detallada, para los tres análisis realizados, se presentan en la Tabla 5²⁴.

En la primera columna, se listan las variables incorporadas en los análisis (las mismas de las regresiones presentadas previamente por cada país), agrupándolas según si se refieren principalmente a los estudiantes, a los establecimientos o a cuestiones institucionales. Las columnas 2 a 4 contienen los resultados de la comparación de Chile con España, las columnas 5 a 7, con Uruguay, y las columnas 8 a 10, con Polonia. Las columnas 2, 5 y 8, presentan las estimaciones de la descomposición atribuidas a las diferencias observadas entre Chile y los demás países, en el nivel de cada una de las variables analizadas, incluyendo subtotalet para los grupos de variables. Las columnas 3, 6 y 9, presentan los resultados asociados a las diferencias de retornos (coeficientes de las regresiones presentadas en la Tabla 3), entre Chile y los demás países, para cada una de las variables estudiadas. Las columnas 4, 7 y 10, presentan las estimaciones de la descomposición, atribuidas al efecto simultáneo de las diferencias de nivel y retorno, de cada una de las variables. Nótese que en la columna de los retornos, se ha agregado en

la primera línea, el valor correspondiente a la constante o intersección²⁵.

Una primera constatación que se evidencia de la Tabla 5, es que, como conjunto, las diferencias en el nivel de las variables que se han denominado “institucionales”, no juegan un papel relevante para comprender las diferencias de resultados entre Chile y los demás países: la magnitud total de la brecha asociada con las diferencias observadas en este componente es, en todos los casos, entre 2 y 4 puntos.

Esto no implica, sin embargo, que algunas variables por separado no sean relevantes. Así por ejemplo, dado que los establecimientos públicos son más efectivos en promedio y que Chile tiene proporcionalmente menos estudiantes que asisten a establecimientos públicos que los demás países, una parte de la brecha de resultados en los tres casos, se explica por este factor. De esta forma, por ejemplo, si Chile tuviese la proporción de estudiantes en establecimientos públicos que tiene Polonia, la brecha en los resultados de la prueba PISA Matemática con este país, se habría reducido en 6 puntos. Otro factor institucional sistemáticamente asociado a la brecha de resultados, es la selección académica de estudiantes, por parte

24 Nótese que los resultados de la descomposición detallada reportan los errores estándar de la estimación hecha para cada variable del modelo (para lo cual hemos utilizado el método propuesto por Jann, 2008). La posibilidad de hacer inferencia estadística fortalece nuestro análisis; sin embargo, dada la naturaleza del método (basado precisamente en “descomponer”, es decir, en separar en fracciones la brecha observada) hemos aplicado adicionalmente en nuestro análisis el criterio de la magnitud, para ponderar los resultados. Así, hemos dado menos relevancia a variables que, aun siendo estadísticamente significativas, permiten explicar una fracción menor de la brecha observada, y no hemos dejado de mencionar algunas otras que, aunque su aporte estimado no superó el umbral de la significación estadística, dan cuenta sin embargo de una proporción importante de dicha brecha. Por lo demás, este proceder es el regular en los estudios publicados que utilizan el método de Oaxaca-Blinder (la inmensa mayoría de los cuales, de hecho, no reporta los errores estándar o el nivel de significancia de sus estimaciones).

25 Aunque algunos estudios de descomposición interpretan el valor de las diferencias de constantes como un indicador de las diferencias globales de “pertenencia al grupo” o, en estudios de este tipo, como diferencias en la eficiencia general del sistema escolar; lo cierto es que en este contexto, su interpretación es más compleja y menos clara. La constante subsume no sólo los efectos de variables no observadas para el estudio, sino que incorpora los valores de referencia de las categorías excluidas para todas las variables dummy utilizadas en el análisis, cuestión que, como sabemos, es completamente arbitraria. La sensibilidad de la constante a estas decisiones metodológicas ha sido demostrada y documentada. En este estudio, por ejemplo, basta tener en cuenta el tremendo efecto que tiene considerar 7° Básico como la categoría excluida en la variable grado, lo cual por sí solo hace cambiar de signo la diferencia entre Chile y, por ejemplo, España y Polonia, sin que esto tenga una interpretación sustantiva clara. Así, en nuestro análisis preferimos adoptar la decisión metodológicamente más conservadora y considerar, como en el análisis clásico, que la constante es parte de la diferencia “no explicada” de la brecha de resultados entre Chile y los demás países de referencia.



de los establecimientos, en los procesos de admisión: en la medida que los establecimientos chilenos que aplican esas prácticas obtienen mejores resultados en promedio y que en Chile la selección académica está más extendida, una fracción negativa (entre 3 y 5 puntos PISA) de la brecha con los demás países, está asociada a ese factor. En todo caso, como se ve, se trata siempre de magnitudes pequeñas. Es interesante constatar que, factores como la autonomía de los establecimientos en materias curriculares, de evaluación y de manejo de recursos, no juegan un papel significativo en la explicación de la diferencia de resultados entre Chile y los demás países analizados: éstos no obtienen mejores resultados que Chile porque sus establecimientos sean más o menos autónomos que los chilenos.

En otro ámbito, en la Tabla 5 también se puede observar que las diferencias en las características de los estudiantes, tienen un efecto negativo sobre la brecha de resultados entre Chile y los demás países. En otras palabras, tomadas como conjunto, si los estudiantes chilenos tuviesen las características observadas de los estudiantes de España, Uruguay o Polonia, la distancia entre los resultados de Chile y de estos países, hubiese sido aun mayor (entre 10 y 36 puntos). En este sentido apuntan, en las tres comparaciones, tanto las expectativas laborales de los estudiantes como su valoración respecto de la importancia de tener un buen rendimiento en Matemática, aunque la magnitud estimada de ambos efectos es pequeña (entre 1 y 4 puntos).

El factor más relevante en este análisis a nivel de los estudiantes, es el grado que éstos se encuentran cursando al momento de dar la prueba. Como se recordará, los estudiantes chilenos acumulaban en promedio más años de escolaridad que sus pares de los otros países (especialmente comparados con los polacos), es decir, cursaban grados superiores. De este modo, como se ve en la Tabla 5, si los estudiantes chilenos hubiesen estado en los grados que cursaban sus pares de los otros países, los resultados de Chile habrían sido más bajos, aumentando la brecha en 10 puntos con España o Uruguay, y en 33 puntos, con Polonia.

En cuanto al efecto asociado a las diferencias en características de los establecimientos, es interesante notar que éste va en el sentido opuesto a los resultados del análisis global presentado hasta aquí, es decir, parte de la brecha observada de resultados (entre 2 puntos en la comparación con Uruguay y 14 puntos en la comparación con Polonia) es atribuible a que, en promedio, los establecimientos chilenos cuentan con condiciones menos favorables para el proceso de enseñanza-aprendizaje.

Es preciso aclarar, sin embargo, que el factor más relevante en este aspecto se refiere al nivel socioeconómico y cultural de los estudiantes del establecimiento (ISEC de los compañeros). Como se observa en la Tabla 5, si el nivel socioeconómico y cultural promedio de los establecimientos chilenos fuese equivalente al de los establecimientos españoles o polacos, los resultados de Chile en PISA Matemática habrían sido más altos, disminuyendo la brecha con estos países en 13 y 12 puntos, respectivamente. En verdad, el hecho de que las condiciones socioeconómicas y culturales de las familias de los estudiantes chilenos sean comparativamente más desfavorables que las de los países de referencia, tiene un efecto aún mayor en la explicación de la brecha observada en PISA, puesto que la misma variable medida a nivel individual también explica parte de esta brecha. De

esta forma, tomadas ambas estimaciones en conjunto (es decir, ISEC individual e ISEC de los compañeros), es posible afirmar que 16 puntos de la brecha en los resultados de PISA Matemática con España y Polonia, y 8 puntos de la brecha con Uruguay, son explicados por las condiciones socioeconómicas y culturales, comparativamente más desfavorables de las familias de los estudiantes chilenos.

El análisis desagregado de las demás variables a nivel del establecimiento (tamaño del establecimiento, tamaño del curso, razón de estudiantes por profesor, recursos educativos y horas de instrucción de Matemática) indica que el resto de la diferencia en los resultados, no se explica por diferentes condiciones de los establecimientos, puesto que ninguna de ellas es un factor que resulte significativo (más aún, casi siempre el efecto asociado a estas variables es de signo negativo). La única excepción a esta afirmación la constituyen las horas de enseñanza de Matemática en la comparación con Polonia, la cual sugiere que 6 puntos PISA de la brecha de Chile con ese país son explicados por la mayor cantidad de horas de instrucción semanal que tienen los estudiantes polacos, en comparación con los chilenos.

Hasta aquí se han analizado con cierto detalle las estimaciones sobre cuánto cambiaría la diferencia entre los resultados PISA Matemática de Chile y los de España, Uruguay o Polonia, si los estudiantes o los establecimientos chilenos tuviesen las características de aquellos países. Sin embargo, los resultados de los análisis contenidos en la Tabla 5 también permiten estimar cuánto se modificaría dicha diferencia, si Chile adquiriese los retornos (o efectos asociados a las características estudiadas, en términos de puntos PISA) que obtienen los otros países en sus procesos educacionales (ciertamente, manteniendo Chile sus actuales características).

En términos generales, el análisis permite concluir que las diferencias de resultados entre Chile y los demás países, no se explican por diferencias en los retornos

asociados a los factores institucionales, cuyos efectos en general son estadísticamente no significativos y/o de magnitudes pequeñas.

La relevancia, en cambio, de las diferencias de retorno, asociadas a las características de los establecimientos varía según el país. Así, en la comparación con Uruguay, ninguno de los coeficientes es estadísticamente significativo (ni los coeficientes principales de retorno ni los de interacción), de forma que es posible afirmar que las diferencias de eficiencia entre Chile y Uruguay a nivel de los establecimientos no juegan un rol importante para explicar la brecha de resultados. Las comparaciones con España y Polonia, en cambio, muestran un patrón similar. En ambos casos, el ISEC de los compañeros (la medición del “efecto par”) es un factor estadísticamente significativo, pero en el sentido opuesto, es decir, si Chile tuviese los retornos de estos países, la brecha con ellos aumentaría²⁶. En este caso, también el coeficiente de interacción es significativo, indicando que la relativa desventaja de Chile respecto de España y Polonia en el nivel socioeconómico de los establecimientos, tiende a ser “compensada” por la mayor productividad de esta variable en términos de resultados en Chile²⁷. En esta dimensión, sin embargo, el factor con la mayor contribución para explicar la brecha de resultados con España y Polonia es el número de estudiantes

26 En este caso, el signo del coeficiente puede inducir a error, puesto que la descomposición detallada de “retornos” es sensible a los signos usados en la construcción de la variable original. En términos intuitivos: Chile tiene un coeficiente de regresión asociado al ISEC de los compañeros mayor que el de España y Polonia (como se observa en la Tabla 2), por lo que este factor debería aportar un valor negativo a la descomposición (es decir, Chile siendo “más eficiente” que los países europeos). El signo positivo, en cambio, se produce porque el valor de Chile en el ISEC es negativo, dada la forma de estandarización de esta variable construida por PISA. Ciertamente, lo mismo es válido para el coeficiente asociado al retorno del ISEC a nivel individual, el cual también debe ser leído con el sentido opuesto a su signo.

27 Ciertamente, a modo de hipótesis, esto puede indicar que, aunque Chile como país es más pobre que España y Polonia, los establecimientos que en Chile se benefician del “efecto par”, tienen comparativamente una retribución mayor en términos de logros de aprendizaje; resultado consistente con un sistema escolar en que los establecimientos son socioeconómica y académicamente más segregados que en los países europeos.

por profesor, sugiriendo una menor eficiencia en el uso de este recurso en el caso de Chile, aunque en ambos casos los coeficientes estimados no son estadísticamente significativos.

El principal hallazgo de este análisis, visto en conjunto, es que los resultados presentados en la Tabla 5, indican que son las diferencias de efectos de las variables medidas a nivel de los estudiantes, las más importantes para explicar la brecha de resultados entre Chile y estos países.

Como puede apreciarse, los resultados sugieren que la educación chilena es más efectiva que la uruguaya en el “uso” de los factores medidos a nivel individual. En otras palabras, si en Chile éstos tuviesen la productividad (retorno) que en Uruguay, Chile habría obtenido, en promedio, 26 puntos menos en la prueba. Sin embargo, en este caso, los factores con el mayor peso estimado no son estadísticamente significativos.

El análisis comparativo con España y Polonia, en cambio, tiene el signo contrario, indicando que la educación chilena es menos efectiva en la transformación de estos factores individuales en resultados de aprendizaje de los estudiantes. De hecho, las magnitudes estimadas atribuibles a estas diferencias de “retorno” son enormes: si en Chile las variables agrupadas en esta dimensión tuviesen los mismos efectos —en términos de los aprendizajes de los estudiantes— que en España, el promedio PISA Matemática de Chile habría aumentado en 187 puntos, y si la comparación es hecha con Polonia, el aumento habría sido menor, pero aún enorme: 90 puntos.

Tanto en la comparación con España como con Polonia, el factor más relevante para explicar esta diferencia de eficiencia entre Chile y los países europeos es el grado que cursan los estudiantes: 161 puntos en España y 46 en Polonia. En otras palabras, si aprobar un grado adicional de escolaridad tuviese en

Chile el efecto que tiene en esos países, las brechas observadas de resultados se reducirían ostensiblemente o desaparecerían. En efecto, como se recordará de la sección previa, una simple inspección del Gráfico 1 mostró la enorme diferencia entre Chile y los dos países europeos (especialmente con España), en cuanto a la relevancia que tiene en esos países las diferencias de grado escolar como factor explicativo de los resultados. En ambos casos, además, el coeficiente de interacción indica que características y retornos apuntan en el sentido contrapuesto (aunque es sólo estadísticamente significativo en el caso de España); es decir, aunque Chile tiene una ventaja en el nivel de escolaridad logrado por los estudiantes, su menor eficiencia en el uso de esos mayores años de escolaridad, resta parte de la contribución potencial de este factor en la comparación. Comparado con Polonia, por ejemplo, mientras la “ganancia” chilena estimada por mayor escolaridad es de 33 puntos PISA, la “pérdida” por ineficiencia es de 46; con España este contraste es aún más marcado.

Finalmente, es interesante observar que las variables “subjetivas” de los estudiantes son también factores importantes para entender las diferencias de resultados de Chile con España y con Polonia. Como se pudo apreciar en la Tabla 3, en los dos países europeos, el hecho de que los estudiantes tengan una alta valoración de la importancia de tener un buen rendimiento en Matemática, está asociado con diferencias de puntaje mucho mayores que en Chile, de forma que si éstas se replicaran en Chile, la brecha de resultados se disminuiría en 20 puntos con España y en 31 puntos, con Polonia. En este último país ocurre algo similar con el hecho de que los estudiantes tengan altas expectativas laborales, factor que podría disminuir la brecha en 22 puntos, si Chile obtuviese los mismos retornos que Polonia. Más aún, los coeficientes de interacción negativos indican que, aunque Chile posee en promedio estudiantes con mayor motivación y más altas expectativas que Polonia y España, sus retornos

en términos de resultados son menores. Ciertamente, no es fácil argumentar que estas asociaciones tengan un carácter causal, toda vez que es posible que los estudiantes más talentosos sean los que, por ejemplo, más valoren los resultados en Matemática. El punto es que, por razones que están fuera del alcance de este estudio, el poseer esta valoración hace una diferencia mayor en términos de resultados de aprendizaje en Polonia y España que en Chile.

Resumiendo. En primer lugar, el análisis muestra que las variables que se ha denominado institucionales, prácticamente no juegan ningún rol para explicar las diferencias entre los resultados de Chile y los demás países de referencia, tanto desde el punto de vista de su nivel como de su retorno. En segundo lugar, las variables que se han agrupado como factores de los establecimientos sí explican una fracción de la brecha de resultados, aunque con importantes matices, siendo más relevantes las diferencias de características en la comparación con Polonia y España, y las diferencias de retorno, en la comparación con Uruguay. Con todo, el factor más importante en esta dimensión, no se refiere a los recursos materiales ni al equipamiento educacional, sino a las condiciones socioeconómicas y culturales de los propios estudiantes (agregadas a nivel del establecimiento). Finalmente, las variables más relevantes para comprender la brecha de resultados de Chile, especialmente con España y Polonia, son las agrupadas a nivel del alumno: los estudiantes chilenos tienen mejores condiciones que los de España y Polonia, pero el sistema escolar chileno es notoriamente menos eficiente. En particular, este análisis se aplica paradigmáticamente al nivel de escolaridad



alcanzado por los estudiantes: los estudiantes chilenos tienen más años de escolaridad que los españoles y los polacos, pero el avance en resultados asociado a esa mayor escolaridad es comparativamente menor. En otras palabras, lo que los chilenos “obtienen” por escolaridad, lo “pierden” por una menor efectividad de dicha escolaridad. Este mismo patrón se repite, aunque en magnitudes menores, respecto de las mayores expectativas laborales y la mayor valoración del buen rendimiento en Matemática.

Tabla 5.

Descomposición de los efectos sobre el puntaje PISA Matemática, asociados a las características, los retornos y la interacción entre ambos.

Análisis separados para la comparación de Chile con España, Uruguay y Polonia

	España			Uruguay			Polonia		
	Efectos asociados a			Efectos asociados a			Efectos asociados a		
	Características	Retornos	Interacción (caract. x retorno)	Características	Retornos	Interacción (caract. x retorno)	Características	Retornos	Interacción (caract. x retorno)
Constante		-93,15 (59,88)			40,57 (25,94)			37,98 (33,47)	
Variables a nivel del estudiante									
Mujer	-1,01* (0,55)	3,11** (1,49)	0,23 (0,19)	-1,53*** (0,58)	1,70 (2,07)	0,19 (0,24)	-1,28*** (0,53)	3,92** (1,80)	0,36 (0,24)
Grado	-9,57*** (1,32)	160,54*** (52,06)	-11,31*** (2,31)	-10,46*** (1,90)	-28,53 (18,18)	2,56 (1,87)	-33,22*** (3,72)	45,69* (22,92)	-13,02 (8,38)
Retraso escolar	0,37 (0,69)	-0,42 (1,80)	0,31 (1,32)	-0,39 (0,73)	-4,42** (1,78)	-3,32** (1,35)	0,44 (0,81)	-2,63 (2,16)	2,24 (1,85)
ISEC (nivel socioeconómico familiar)	3,37*** (0,73)	0,11 (1,22)	-0,06 (0,68)	1,65*** (0,61)	-3,19** (1,29)	0,87* (0,45)	3,44*** (0,71)	-9,94*** (1,40)	5,64*** (1,07)
Expectativas laborales	-1,14*** (0,41)	4,09 (4,04)	-0,42 (0,32)	0,61* (0,32)	-5,07 (5,60)	-0,24 (0,31)	-1,60*** (0,50)	21,95*** (3,56)	-2,83*** (0,75)
Valoración de un buen rendimiento en Mat.	-2,25*** (0,61)	19,69*** (6,45)	-3,12*** (0,75)	-1,87*** (0,51)	13,29 (8,62)	-1,26 (0,84)	-4,02*** (1,08)	31,44*** (6,41)	-11,23*** (1,40)
Subtotal estudiantes	-10,22*** (2,42)	187,11*** (56,39)	-14,38*** (3,43)	-11,99*** (2,56)	-26,22 (22,23)	-1,20 (1,73)	-36,25*** (4,81)	90,42*** (25,59)	-18,84** (9,17)
Variables a nivel del establecimiento									
ISEC de los compañeros (efecto par)	12,59*** (2,51)	10,85*** (3,62)	-6,31*** (2,25)	5,89*** (2,08)	4,44 (4,68)	-1,21 (1,31)	12,35*** (2,35)	12,28** (5,26)	-7,00** (3,12)
Establecimiento complejo	-1,88 (1,75)	0,00 (4,01)	0,00 (2,49)	-1,66 (1,55)	4,07 (3,37)	-2,22 (1,82)	-2,95 (2,72)	-3,09 (5,79)	3,00 (5,62)
Tamaño del curso	-2,44 (4,06)	-9,30 (16,00)	2,53 (4,36)	-1,17 (1,97)	10,16 (16,90)	-1,33 (2,23)	-2,93 (4,89)	-4,59 (31,46)	1,50 (10,28)
Tamaño del establecimiento	-1,00 (1,28)	-10,92* (6,15)	3,70* (2,11)	-1,73 (2,21)	1,46 (11,40)	-0,86 (6,66)	-1,81 (2,31)	-9,41 (12,48)	5,75 (7,68)
Horas de instrucción Matemática	-0,39 (0,44)	-6,13* (3,27)	0,11 (0,15)	-1,15** (0,48)	-1,74 (3,51)	0,09 (0,21)	5,55*** (0,54)	-2,67 (2,77)	-0,67 (0,69)
Alumnos por profesor	2,57 (4,14)	20,61 (19,20)	-10,45 (9,80)	1,86 (3,00)	-4,69 (15,26)	1,72 (5,58)	2,79 (4,49)	36,44 (26,80)	-20,01 (14,71)
Recursos educativos	1,29 (1,42)	-0,16 (1,79)	0,16 (1,73)	-0,18 (0,32)	1,73 (1,93)	0,24 (0,43)	1,15 (1,25)	1,11 (2,09)	-0,96 (1,76)
Subtotal establecimientos	10,75* (5,54)	4,94 (21,80)	-10,26 (10,03)	1,86 (3,91)	15,42 (22,01)	-3,57 (7,65)	14,14** (6,42)	30,08 (34,74)	-18,39 (16,61)

(Continuación en página siguiente)

(Continuación Tabla 5)

	Características	Retornos	Interacción (caract. x retorno)	Características	Retornos	Interacción (caract. x retorno)	Características	Retornos	Interacción (caract. x retorno)
VARIABLES A NIVEL INSTITUCIONAL									
Autonomía en el uso de los recursos	1,12 (0,87)	0,78 (1,17)	-0,83 (1,27)	-2,76 (1,98)	-0,39 (1,40)	-1,03 (3,59)	-0,02 (0,38)	1,32 (1,18)	0,03 (0,48)
Autonomía en currículum y evaluación	-1,25 (1,14)	-0,03 (0,75)	0,07 (1,52)	2,24 (1,98)	0,20 (1,38)	0,71 (5,00)	-1,41 (1,23)	0,22 (0,73)	-0,51 (1,68)
Selección académica	-4,63*** (1,49)	-6,64* (3,60)	5,78* (3,17)	-3,96*** (1,27)	-1,33 (2,72)	0,99 (2,05)	-3,08** (1,17)	-0,01 (2,54)	0,01 (1,44)
Expectativas de padres por altos estándares	-1,69** (0,80)	-3,15* (1,77)	1,66 (1,01)	-2,04** (0,90)	-1,90 (1,69)	1,20 (1,10)	0,74 (0,78)	-3,41** (1,47)	-0,78 (0,87)
Establecimiento público	2,40** (1,15)	0,00 (4,08)	0,00 (1,94)	4,78** (2,25)	1,86 (6,05)	1,75 (5,65)	6,31** (2,96)	-6,61 (7,65)	-8,24 (9,48)
Subtotal institucionalidad	-4,06 (2,59)	-9,04* (5,17)	6,67 (4,59)	-1,74 (3,50)	-1,56 (6,42)	3,63 (5,49)	2,54 (3,55)	-8,50 (8,69)	-9,50 (9,50)
TOTAL	-3,53 (6,86)	89,87*** (8,56)	-17,97* (9,75)	-11,86* (6,35)	28,21*** (9,31)	-1,14 (10,90)	-19,57** (8,82)	149,98*** (21,59)	-46,74** (23,03)

Los coeficientes son estadísticamente significativos al: *** < 1%; ** < 5%; * < 10%

4. Conclusiones

Este estudio buscó avanzar en la comprensión de los resultados obtenidos por los estudiantes chilenos en la prueba PISA 2006, en Matemática²⁸. Para ello, se realizó un análisis comparativo con dos países de la OCDE (España y Polonia) y un país de América Latina (Uruguay). En particular, se intentó identificar los principales factores que explican las diferencias en los resultados promedio obtenidos por los estudiantes chilenos y sus pares de estos tres países, todos los cuales superan a los chilenos. La investigación utilizó los datos disponibles de PISA 2006 en Matemática, para los cuatro países mencionados, incluyendo los puntajes de logro a nivel de los estudiantes, así como variables descriptivas de los estudiantes y sus familias, y diferentes aspectos de sus establecimientos. Como se sabe, en términos del promedio nacional, la brecha que separa a Chile del rendimiento de los países de

la OCDE es significativamente mayor en Matemática que en Lectura y Ciencias (las otras dos asignaturas evaluadas por PISA).

La investigación implementó una variante de la metodología de Oaxaca-Blinder, la cual permite básicamente descomponer dichas diferencias de resultados, estimando la parte de esta brecha que puede ser atribuida a diferencias en el nivel de ciertas características relevantes entre los grupos que están siendo comparados (en este caso, Chile y cada uno de los países de referencia); la parte que puede ser atribuida a diferencias en el retorno (productividad) de dichas características; así como el efecto combinado o interacción de ambas diferencias, de nivel y retorno. La hipótesis básica del estudio es que las diferencias de logro entre Chile y los demás países se deben a una combinación de diferencias de características con diferencias en la capacidad de transformar dichas características en logros de aprendizaje.

28 En el artículo siguiente, estos hallazgos son comparados con los obtenidos en Lectura y son además analizados desde la perspectiva de sus implicancias para las políticas educacionales.

Analizando un listado amplio de variables (agrupadas en aspectos del alumno, del establecimiento y de la institucionalidad) consideradas por la literatura como relevantes para explicar los resultados de los estudiantes en las pruebas PISA, se constató que efectivamente Chile difiere significativamente de los demás países en prácticamente todas estas características. Luego, a partir de los modelos de regresión múltiple para explicar los puntajes PISA de los estudiantes a partir de las variables seleccionadas, la mayor parte de las variables estudiadas resultó ser un factor relevante de los logros de aprendizaje. En el caso chileno, estos modelos permitieron explicar el 53% de la variación de resultados en PISA Matemática a nivel de los estudiantes. Adicionalmente, se pudo constatar que el efecto estimado de la mayoría de estos factores en los logros de aprendizaje de los estudiantes difiere en su intensidad entre Chile y los países sometidos a comparación. Finalmente, a través de la descomposición de las diferencias de resultados entre Chile y estos países, fue posible estimar la fracción atribuible a diferencias en las características, sus retornos y la interacción entre ambos.

Tomados en su sentido más general, los hallazgos de este estudio indican que los peores resultados de Chile en PISA-Matemática 2006 no pueden ser atribuidos a desventajas en las características o recursos de los estudiantes y establecimientos chilenos; al contrario, en varios aspectos relevantes para el aprendizaje, los estudiantes y los establecimientos de Chile se encontrarían incluso “mejor equipados” que sus pares de los tres países comparados. Dicho de otro modo, tomando en cuenta sus condiciones, Chile tiene buenas razones para estar insatisfecho con los niveles de aprendizaje que logra generar en sus estudiantes.

Ciertamente, el análisis muestra que existen algunos aspectos en los cuales la educación chilena está en desventaja y que, de ser mejorados, permitirían a Chile mejorar su posición relativa respecto de estos países.

El más importante de ellos, sin embargo, no se refiere a los establecimientos, sino al nivel socioeconómico-cultural de las familias chilenas; de hecho, según nuestras estimaciones, si Chile alcanzara los niveles socioeconómicos promedio de los estudiantes de los países de comparación (todos superiores a Chile, pero inferiores a la media OCDE) el promedio chileno aumentaría entre 8 y 16 puntos en Matemática, lo que sugiere un cierto potencial de mejoramiento educativo por la vía del mejor desarrollo socioeconómico. Del resto de las variables analizadas, sólo la diferencia con Polonia en cuanto a las horas de instrucción en Matemática y la proporción de estudiantes que asiste a establecimientos públicos (mayores para Polonia en ambos casos), tienen una incidencia relativamente importante, aunque comparativamente menor en términos de su magnitud.

¿Qué explica entonces la brecha de resultados entre Chile y los países de referencia? Según los análisis realizados, la totalidad de la brecha de resultados en Matemática es susceptible de ser explicada por las diferencias en la eficiencia de la educación; en otras palabras, la razón de la desventaja chilena se encontraría en la menor productividad de las características y recursos, en términos de logros de aprendizaje.

Como se sabe, en educación, el proceso por el cual los recursos y las condiciones se “transforman” en logros de aprendizaje es enormemente complejo. Una dimensión clave de este proceso son las prácticas de gestión escolar y las características institucionales de los establecimientos, las cuales, en teoría, afectan decisivamente la eficiencia de los sistemas escolares. Esta investigación analizó la posible incidencia de algunas de las decisiones más relevantes para la política educacional en este campo (por ejemplo, la propiedad (pública o privada) de los establecimientos, el grado de autonomía en las decisiones curriculares y de recursos, por parte de los establecimientos, entre otros), sin embargo, la evidencia indica que ellas no son relevantes para explicar los peores resultados obtenidos por Chile en Matemática.

La otra dimensión clave para comprender el proceso de generación de aprendizajes en el sistema escolar son las decisiones de gestión y práctica pedagógicas. Desafortunadamente la información provista por PISA es más pobre en este ámbito, por lo que la comprensión de cómo afecta lo pedagógico a la mayor ineficiencia chilena emerge como una línea de investigación a profundizar. Con todo, el hallazgo más relevante del estudio en este aspecto es la enorme incidencia que tienen en los resultados PISA, las diferencias en los años de escolarización entre los estudiantes evaluados.

En efecto, dado que la muestra seleccionada por PISA incluye estudiantes de la misma edad, pero que están cursando diferentes grados, y dado que las políticas de edad/grado, así como de promoción/repitencia, varían entre los países; se produce una enorme heterogeneidad en este aspecto entre los países estudiados. Las comparaciones con España y Polonia son elocuentes en este sentido. Sacando partido a la heterogeneidad referida, se pudo estimar que toda la diferencia de resultados con España y la mayor parte de la diferencia con Polonia, podría ser atribuida a la mayor efectividad que cada grado cursado tiene en esos países, respecto de Chile.

Lo que está a la base de este hallazgo son las enormes diferencias en lo que se podría denominar el “valor agregado” por cada curso aprobado en uno y otro país. Así, mientras en Chile (en promedio y controlando por otras diferencias) los estudiantes de 10° grado superan a sus compañeros de 9° grado en 41 puntos, esta diferencia es en España de 70 puntos y en Polonia, de 73 puntos. Por cierto, esta desventaja es compensada en cierta medida por el hecho de que los estudiantes chilenos han alcanzado a los 15 años de edad una mayor escolarización que sus pares españoles y polacos. Este reemplazo de “insumos” por “eficiencia” no alcanza, sin embargo, para compensar la mayor efectividad polaca y es, en todo caso, marginal en la comparación con España.

Dado que se trata de diferencias netas estimadas (controlado por las diferencias socioeconómicas y culturales de los estudiantes y sus familias, así como por un conjunto relevante de variables a nivel de los establecimientos), es posible formular dos potenciales explicaciones alternativas: la presencia de diferencias curriculares entre los países y la existencia de diferencias en la efectividad y calidad de los docentes. En efecto, si el patrón identificado para los años de escolaridad se repitiese en referencia a otros recursos y condiciones asociados a las prácticas de enseñanza y la gestión pedagógica, en general, reforzaría la noción de que los grandes desafíos contemporáneos de la educación chilena se encuentran comparativamente más cerca de las salas de clase que de las oficinas de la dirección o de las decisiones sobre la magnitud de los recursos.

Finalmente, un hallazgo interesante es que, comparados con sus pares, los estudiantes chilenos tienden a manifestar una mayor valoración de la importancia de lograr buenos resultados en Matemática, sin embargo, Chile es el país en donde esta mayor valoración tiene un efecto estimado más pequeño en los resultados. ¿Es acaso que los establecimientos y docentes chilenos sacan menos provecho en términos de enseñanza-aprendizaje a esta positiva actitud de sus estudiantes?

Referencias

- Ammermüller, A. (2004). "PISA: What makes the difference? Explaining the gap in PISA test scores between Finland and Germany". *Discussion paper*, 04-04. Mannheim, Alemania: Centre for European Economic Research.
- Arregui, P. (Ed.) (2006). *Sobre estándares y evaluaciones en América Latina*. Santiago, Chile: Programa de Promoción de la Reforma Educativa de América Latina y el Caribe (PREAL).
- Bellei, C. (2003). "¿Ha tenido impacto la Reforma Educativa Chilena?", Cox, C. (Ed.), *Políticas educacionales en el cambio de siglo. La reforma del sistema escolar de Chile*. Santiago, Chile: Editorial Universitaria.
- _____ (2009). "The public-private school controversy in Chile", Chakrabarti, R. y Peterson, P.E. (Eds.), *School Choice International*. Cambridge, MA: MIT Press.
- Blinder, A. (1973). Wage discrimination: reduced form and structural estimates. *The Journal of Human Resources*, 4 (8). Wisconsin, Estados Unidos: The University of Wisconsin Press.
- Evers, W. y Walberg, H. (Eds.) (2002). *School Accountability*. Stanford, California: Hoover Institution Press.
- Fuchs, T. y Woessmann L. (2004). "What accounts for International differences in student performance? A re-examination using PISA data", *IZA Discussion Paper*, 1287. Bonn, Alemania: Institute for the Study of Labor.
- Grissmer, D., Flanagan, A. y Williamson, S. (1998). *Why did the black-white score gap narrow in the 1970s and the 1980s?*. Washington DC: Brookings Institution Press.
- Harmon, M. y otros (1997). *Performance assessment in IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, Massachusetts: TIMSS International Study Center, Boston College.
- Jann, B. (2008). "A Stata implementation of the Blinder – Oaxaca decomposition", *ETH Zurich Sociology Working Paper*, 5. Zurich, Suiza: Swiss Federal Institute of Technology Zurich (ETH).
- McEwan, P. y Marshall, J. (2004). "Why does academic achievement vary across countries? Evidence from Cuba and Mexico". *Education Economics*, 12 (3). Oxford, Reino Unido: Taylor & Francis.
- Mullis, I. y otros (2000). *Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. Chestnut Hill, Massachusetts: TIMSS International Study Center, Boston College.
- _____ (2003). *PIRLS 2001 International Report: IEA's Study of Reading Literacy Achievement in Primary Schools*. Chestnut Hill, Massachusetts: Boston College.
- _____ (2007). *IEA's Progress in International Reading Literacy Study in Primary School in 40 Countries*. Chestnut Hill, Massachusetts: TIMSS & PIRLS International Center, Boston College.
- Oaxaca, R. (1973). "Male – Female wage differentials in urban labor markets". *International Economic Review*, 14 (3). Philadelphia PA: University of Pennsylvania.
- Oaxaca, R. y Ransom, M. (1999). "Identification in detailed wage decompositions". *The Review of Economics and Statistics*, 81 (1). Cambridge, Massachusetts: MIT Press.
- OCDE (Organización para la Cooperación y el Desarrollo Económico) (2003). *Literacy Skills for the World of Tomorrow - Further Results from PISA 2000*. París, Francia.
- _____ (2004). *Learning for Tomorrow's World, First Results from PISA 2003*. París, Francia.
- _____ (2007). *PISA 2006 Science Competencies for Tomorrow's World*. París, Francia.
- _____ (2009) *PISA Technical Report 2006*. París, Francia.
- Sakellariou, C. (2008). "Peer effects and the indigenous / non indigenous early test-score gap in Peru". *Education Economics*, 16 (4). Oxford, Reino Unido: Taylor & Francis.
- Valenzuela, J. P. (2008). "Segregación en el sistema escolar chileno: en la búsqueda de una educación de calidad en un contexto de extrema desigualdad", en *Transformaciones del Espacio Público*. Santiago, Chile: Universidad de Chile.

Anexo

1. Descripción de variables

1.1 Variables de resultados

a) Puntaje PISA Lectura.

Variable continua elaborada por PISA, normalizada con una media de 500 puntos y desviación estándar de 100 puntos para los países OCDE. En este estudio, el rango de valores está entre 413 (Uruguay) y 508 (Polonia).

b) Puntaje PISA Matemática.

Variable continua elaborada por PISA, normalizada con una media de 500 puntos y desviación estándar de 100 puntos para los países OCDE. En este estudio, el rango de valores está entre 411 (Chile) y 495 (Polonia).

1.2 Variables de control

Variables a nivel individual - familiar

a) Género

Variable dicotómica construida por PISA a partir de la pregunta del cuestionario de estudiantes: *¿Eres mujer u hombre? (Q4)*

Categorías de la variable:

0 = hombre

1 = mujer

b) Años de escolaridad

Variable categórica de elaboración propia, construida sobre la base de la pregunta del cuestionario de estudiantes: *¿En qué curso estás? (Q1)*

Categorías de la variable:

7 = 7°

8 = 8°

9 = 9°

10 = 10°

11 = 11°

Los cursos del sistema educativo de los países utilizados en la comparación se estructuran de la siguiente forma:

Uruguay

- Enseñanza Primaria a partir de los 6 años (son 6 años).
- Educación Media inicial común y obligatoria hasta los 14 años.
- Segundo Ciclo de 3 años (bachillerato) o cursos técnico-profesionales (2 a 7 años).

España

- Escuela Primaria de 1° a 6° grado, de 6 a 12 años.
- Educación Secundaria obligatoria (*Compulsory Secondary Education*) 4 cursos, de 12 a 16 años.
- Bachillerato (*Baccalaureat*) 2 cursos, con 4 modalidades de especialización en 2006 (actualmente son 3).

Polonia²⁹

- Escuela Elemental de 0° a 6° grado, de 7 a 12 años.
- *Gymnasium*, de 7° a 9° grado, de 13 a 15 años.
- *Postsecondary Schools Lyceum General*, 3 años.

c) Retraso escolar

Variable dicotómica de elaboración propia, construida sobre la base de las preguntas sobre el curso en que se encuentra el alumno (Q1) y su edad.

Edad: variable calculada por PISA que corresponde a la diferencia entre el año y mes en que se aplica

²⁹ En Polonia no se encuestaron estudiantes de 11° grado, por la forma en que se distribuyen los cursos en su sistema educativo.

PISA en el país, y el año y mes de nacimiento del estudiante. Los valores son de 15 a 16 años, con algunas variaciones en los meses entre países.

En Chile y Uruguay, se establece que los estudiantes encuestados tienen retraso escolar si están cursando entre 7° y 8° grado, al momento de rendir PISA, o si están cursando el 9° grado y tienen una edad de 15 años y 6 meses o más. El resto, se considera que no presenta retraso escolar (la fecha de aplicación de PISA fue agosto de 2006).

En España y Polonia, se establece que los estudiantes encuestados tienen retraso escolar si están cursando 7° grado, o si cursan 8° grado, teniendo una edad de 15 años y 6 meses o más. El resto se considera que no presenta retraso escolar (la fecha de aplicación de PISA fue abril de 2006).

Categorías de la variable:

0 = sin retraso escolar

1 = con retraso escolar

d) ISEC (Índice de Estatus Socioeconómico y Cultural) familiar

Variable continua calculada por PISA, que se deriva de las siguientes variables:

- HISCEI: índice internacional sobre la más alta ocupación alcanzada por el padre o la madre.
- PARED: índice sobre el más alto nivel educacional de los padres, expresado en años de escolaridad.
- HOMEPOS: índice de bienes en el hogar, que incluye los ítems de *wealth* (riqueza familiar), *cultpos* (bienes culturales) y *hedres* (recursos educacionales).

La variable está normalizada, para los países del estudio y llega a tomar un valor mínimo de -5.0691, en el caso de España y un valor máximo de 2.9472, en Polonia.

e) Expectativas laborales

Es una variable categórica de elaboración propia, construida a partir de la pregunta del cuestionario

de estudiantes: *¿Qué tipo de trabajo esperas tener cuando tengas alrededor de 30 años? (Q30)*

Las respuestas a esta pregunta fueron codificadas por PISA, según los códigos ocupacionales ISCO, usados por la Organización Internacional del Trabajo. Las categorías ISCO fueron recodificadas para este estudio en tres categorías:

1 = Expectativas laborales altas: se consideran las categorías profesionales (códigos 1120 hasta < 3000).

2 = Expectativas laborales medias: se consideran las categorías de técnicos y profesionales asociados (códigos 3000 hasta < 5000).

3 = Expectativas laborales bajas: se consideran las categorías de técnicos y administrativos, trabajadores de servicios (códigos 4000 hasta < 9501).

Las respectivas variables dicotómicas, se obtienen a partir de cada una de las alternativas de la variable.

f) Valorización sobre la importancia de un buen rendimiento en Matemática y Lenguaje

Variable categórica de elaboración propia, derivada a partir de la pregunta del cuestionario de estudiantes: *En general, ¿qué tan importante crees tú que es tener buen rendimiento en los siguientes ramos? (Q36)*

Matemática:

(1) Muy Importante, (2) Importante, (3) Poco Importante, (4) No es importante.

Lenguaje:

(1) Muy Importante, (2) Importante, (3) Poco Importante, (4) No es importante.

La variable fue recodificada para obtener 3 categorías.

Categorías de la variable:

1 = muy importante (1)

2 = importante (2)

3 = de menor importancia (3 y 4)

A partir de éstas, se obtienen las variables dicotómicas para las 3 categorías.

Variables del establecimiento

a) Efecto par

Variable continua de elaboración propia, que mide (separadamente para cada alumno) el efecto par del ISEC del establecimiento. El valor del efecto-par está medido a nivel del alumno, y corresponde al promedio del ISEC del establecimiento, excluyendo al alumno en cuestión (“ISEC promedio de mis compañeros”).

La variable está normalizada, para los países del estudio y llega a tomar un valor mínimo de -3,171 en el caso de Chile y un valor máximo de 1,789, en Polonia.

b) Tamaño del curso

Variable continua construida por PISA (*CLSIZ*), a partir de la pregunta del cuestionario a los directores del establecimiento: *En su establecimiento, ¿cuál es el tamaño promedio de los cursos de Lenguaje en el grado que corresponde a los estudiantes de 15 años? (Q6)*

Se presentan nueve alternativas de respuesta:

(1) 15 estudiantes o menos; (2) entre 16 y 20 estudiantes; (3) entre 21 y 25 estudiantes; (4) entre 26 y 30 estudiantes; (5) entre 31 y 35 estudiantes; (6) entre 36 y 40 estudiantes; (7) entre 41 y 45 estudiantes; (8) entre 46 y 50 estudiantes; (9) más de 50 estudiantes.

La variable *CLSIZ* toma el punto medio de cada alternativa, asignando un valor mínimo de 13 en la primera alternativa y un valor máximo de 53, para la última.

Para los países del estudio llega a tomar un valor mínimo de 13 en el caso de España, Uruguay y Polonia y un valor máximo de 53, en España y Uruguay.

c) Tamaño del establecimiento

Variable continua construida por PISA (*SCHLSIZ*), contiene el total de la matrícula del establecimiento, basado en la respuesta del director del establecimiento

a la pregunta: *¿Cuál era la matrícula total del establecimiento, al (fecha) del 2006? (Q1)*

– Número de estudiantes hombres.

– Número de estudiantes mujeres.

Las fechas correspondieron a: 1° de febrero (países de la OCDE), 20 de julio (en Chile).

La variable se obtiene del total de ambas alternativas, y para los países del estudio, llega a tomar un valor mínimo de 20 en el caso de Polonia y un valor máximo de 2834, en Chile.

d) Tiempo de instrucción para las asignaturas de Matemática y Lenguaje

Variable categórica de elaboración propia, construida a partir de la pregunta del cuestionario de estudiantes: *Usualmente, ¿cuánto tiempo semanal, dedicas al estudio de los siguientes ramos? (Q31)*

Indica por separado para cada ramo:

– *El tiempo que dedicas a las clases regulares en tu establecimiento.*

– *El tiempo que dedicas a clases extras (en tu mismo establecimiento, en tu casa, o en cualquier otro lugar).*

– *El tiempo que dedicas a estudiar o a hacer tareas tú solo.*

Contesta considerando que una hora se refiere a 60 minutos y no a un período de clases de 45 minutos.

Categorías de la variable:

(1) No le dedico tiempo; (2) Menos de 2 horas semanales; (3) 2 o más, pero menos de 4 horas semanales; (4) 4 o más, pero menos de 6 horas semanales; (5) 6 o más horas semanales.

Para este estudio, se consideró solo las respuestas en relación a las clases regulares en el establecimiento. La variable fue recodificada, tomando el punto medio de cada alternativa, y asignando un valor de 0 en la primera alternativa y un valor máximo de 7 para la última.

e) Alumnos por profesor

Variable continua construida por PISA (*STRATIO*), que se obtiene dividiendo el tamaño del establecimiento por el número total de profesores. El número de profesores con media jornada, está pesado por 0.5 y el número de profesores con jornada completa esta pesado por 1.0.

La variable para los países del estudio va de un valor mínimo de 1,19 en España, a un valor máximo de 43,42 en Chile.

f) Establecimiento complejo

Variable dicotómica de elaboración propia, construida a partir de los grados (cursos) que se imparten en el establecimiento, sobre la base de la pregunta: *¿Cuáles de los siguientes niveles o cursos se imparten en el establecimiento? (Q4)*

Los cursos van de 1° a 13° grado en el cuestionario internacional.

Aquellos establecimientos que tienen los niveles de Educación Primaria o Inicial junto con Educación Media o Secundaria, son considerados establecimientos complejos; el resto de los establecimientos que sólo cuenta con un nivel, es considerado como no complejo. Aquellos establecimientos que solo tienen algunos grados de Educación Primaria junto con los de Educación Media, son también considerados como establecimientos no complejos.

Categorías de la variable:

0 = establecimiento no complejo.

1 = establecimiento complejo.

g) Índice de calidad de los recursos educativos

Variable continua construida por PISA (*SCMATEDU*), que se deriva de siete ítemes que miden las percepciones de los directores frente a los potenciales factores que afectan la enseñanza en el establecimiento. Los ítemes se encuentran en la pregunta del cuestionario del

establecimiento: *La enseñanza en su establecimiento, ¿se ve afectada por alguno de los siguientes factores? (Q14)*

Categorías de la variable:

g) Escasez de equipamiento del laboratorio de ciencias o equipamiento inadecuado.

h) Escasez de material de enseñanza o material inadecuado (por ejemplo, libros de texto).

i) Escasez de computadores para la enseñanza o computadores inadecuados.

j) Falta de conexión a internet o conexión inadecuada.

k) Escasez de software computacional para la enseñanza o software inadecuado.

l) Escasez de material de biblioteca o material inadecuado.

m) Escasez de recursos audiovisuales o recursos audiovisuales inadecuados.

La variable está normalizada y tiene un valor mínimo de -3,4335 en el caso de Chile y Uruguay, y un valor máximo de 2.1351 en todos los países, sin excepción.

Variables institucionales

a) Índice de autonomía escolar en decisiones sobre la asignación de recursos

Índice construido por PISA (*RESPRES*) que se deriva del número de decisiones en relación a la asignación y administración de los recursos del establecimiento que son de responsabilidad del establecimiento. Se utiliza la pregunta del cuestionario del establecimiento: *En su establecimiento, ¿quién tiene la responsabilidad sobre las siguientes tareas? Marque todas las alternativas que sean necesarias (Q11)*

– Contratar profesores.

– Despedir profesores.

– Establecer los sueldos iniciales de los profesores.

– Determinar aumentos de sueldo de los profesores.

– Elaborar el presupuesto escolar.

– Decidir sobre la asignación de recursos dentro del establecimiento.

Alternativas de respuesta:

- 1) Director o profesores.
- 2) Directorio del establecimiento.
- 3) Autoridad educativa regional o local.
- 4) Autoridad educativa nacional (Ministerio).

El índice fue calculado sobre la base del ratio de las respuestas “sí” en las 4 alternativas anteriores³⁰. Los valores altos de la escala indican mayores niveles de autonomía del establecimiento en esa área. El índice fue estandarizado teniendo una media de los países de la OCDE de 0 y una desviación estándar de 1 (para el conjunto de datos con muestras por país igualmente pesadas). Para los países del estudio llega a tomar un valor mínimo de -1,103 en el caso de Uruguay y un valor máximo de 2,018 en todos los países, sin excepción.

b) Índice de autonomía escolar en decisiones sobre currículum y evaluación

Es un índice construido por PISA (*RESPCURR*) que se deriva del número de decisiones en relación al currículum, que son responsabilidad del establecimiento. Se utiliza la pregunta del cuestionario del establecimiento: *En su establecimiento, ¿quién tiene la responsabilidad sobre las siguientes tareas? Marque todas las alternativas que sean necesarias (Q11)*

- Establecer las políticas de evaluación de los estudiantes.
- Elegir qué libros de texto se usarán.
- Determinar el contenido de los ramos.
- Decidir qué cursos se impartirán.

Alternativas de respuesta:

- 1) Director o profesores.
- 2) Directorio del establecimiento.
- 3) Autoridad educativa regional o local.
- 4) Autoridad educativa nacional (Ministerio).

³⁰ Una respuesta sí, significa que el director ha seleccionado la alternativa correspondiente.

El índice fue calculado sobre la bases del ratio de las respuestas “sí” en las 4 alternativas anteriores³¹. Los valores altos de la escala indican mayores niveles de autonomía del establecimiento en esa área. El índice fue estandarizado, teniendo una media de los países de la OCDE de 0 y una desviación estándar de 1 (para el conjunto de datos con muestras por país igualmente pesadas). Para los países del estudio llega a tomar un valor mínimo de -1,399 en el caso de Chile y Uruguay y un valor máximo de 1,269 en todos los países, sin excepción.

c) Criterios académicos de alta prioridad o prerequisites para la selección de los estudiantes postulantes al establecimiento

Variable dicotómica de elaboración propia, construida a partir de la recodificación que PISA realiza (*SELSCH*) de dos categorías de la pregunta del cuestionario del establecimiento: *Al admitir estudiantes en su establecimiento, ¿qué importancia se da a los siguientes factores? (Q19)*

- Desempeño académico del alumno (incluidas las pruebas y exámenes de admisión).
- Recomendación del establecimiento del cual el alumno proviene.

Alternativas

- (1) No se consideran.
- (2) Al menos una se considera.
- (3) Al menos una es de alta prioridad.
- (4) Al menos una es prerequisite.

Categorías de la variable:

0 = 1 y 2

1 = 3 y 4

³¹ Una respuesta sí, significa que el director ha seleccionado la alternativa correspondiente.

d) Presión de los padres sobre el establecimiento por altos estándares

Variable dicotómica de elaboración propia, construida a partir de la pregunta del cuestionario del establecimiento: *¿Cuál de las siguientes afirmaciones caracteriza mejor las expectativas de los padres sobre el establecimiento? (Q16)*

(1) *Hay presión constante de una mayoría de padres, que esperan que el establecimiento establezca estándares de educación muy altos y que los estudiantes los alcancen.*

(2) *La presión sobre el establecimiento, para que alcance estándares de educación más altos entre los estudiantes, viene sólo de una minoría de los padres.*

(3) *No existe presión de los padres para que el establecimiento alcance estándares de educación más altos entre los estudiantes.*

Categorías de la variable:

0 = Alternativas 2 y 3

1 = Alternativa 1.

e) Establecimiento público

Variable dicotómica construida por PISA, que se deriva de la pregunta del cuestionario al establecimiento: *Su establecimiento es (Q2):*

1) *Público (establecimiento administrado directa o indirectamente por la autoridad educacional pública, una agencia gubernamental, o un consejo designado por el gobierno o elegido por franquicia pública).*

2) *Privado (establecimiento administrado directa o indirectamente por una organización no gubernamental; i.e. congregaciones religiosas, organizaciones de comercio, instituciones privadas).*

Para el caso de Chile, esta última categoría incluye los establecimientos particulares subvencionados y particulares pagados.

Categorías de la variable:

0 = establecimiento no público

1 = establecimiento público

2. Estimación con valores plausibles (plausible values)

Como se describe en el *Technical Report 2006* (OCDE, 2009, pp. 156-160), los valores plausibles son números aleatorios obtenidos a partir de una distribución de puntajes asignados a cada individuo de la muestra, es decir, corresponden a una post distribución. En términos de PISA estos valores son asignados en un número de cinco por observación, por cada test de la encuesta (en nuestro caso, *mathematical literacy* y *reading literacy*).

El propósito de trabajar con los valores plausibles —y no con los puntajes individuales, directamente— se basa en lo deseables que son las propiedades de consistencia que se encuentran en los estimadores para los parámetros de la población.

Cualquier análisis que involucre los valores plausibles deberá realizarse cinco veces, por el número de los *vp* (*valores plausibles*); el valor del estadístico será el promedio de las cinco estimaciones, y su varianza será ajustada por cada estimación y en conjunto.

A manera de ilustración, supóngase que se quiere calcular el estadístico $r^*(X, Y)$. Este deberá ser calculado como el promedio de las cinco estimaciones con los valores plausibles:

$$\begin{aligned} r^*(X, Y) &= \frac{1}{M} \sum_{m=1}^M r(\theta_m, Y) \\ &= \frac{1}{M} \sum_{m=1}^M \hat{r}_m \end{aligned}$$

Si U_m es la varianza muestral de \hat{r}_m , entonces la varianza muestral de r^* será:

$$V = U^* + (1 + M^{-1})B_M$$

$$\text{donde } U^* = \frac{1}{M} \sum_{m=1}^M U_m \text{ y } B_M = \frac{1}{M-1} \sum_{m=1}^M (\hat{r}_m - r^*)^2$$

Las estimaciones de los errores estándar incluyen los pesos de las 80 Replicaciones Balanceadas (*BRR*) proporcionadas en la base de datos, con la corrección de *Fay* (0.5). Estos errores estándar proporcionan el grado de variación de cualquier estadístico respecto del parámetro poblacional. Estas replicaciones consideran el diseño complejo de la encuesta, realizado en dos etapas, con un diseño de muestra estratificado. No incluir este componente de encuesta compleja en las estimaciones, es subestimar los errores estándar, llevando a inferir significancia estadística en las estimaciones, cuando en realidad no la hay (error tipo I).

3. Método de imputación

El método general para resolver problemas de imputación consiste en identificar un conjunto de variables explicativas que tengan todas las observaciones de la muestra, con el fin de utilizarlos como controles para la imputación, es decir en base a estas variables ordenadas se procede con la imputación.

En este estudio, el método aplicado consistió en la imputación de medianas, la cuales correspondían al valor observado de un conjunto de casos que coincidieran con similares subgrupos asociados a las diversas variables de control, de tal forma que las observaciones con datos imputados pertenecieran a subgrupos relativamente similares. El proceso implicaba que en aquellos casos en que no se podía aplicar la imputación, paulatinamente el número de controles utilizados para definir los subgrupos se iba reduciendo.

La mejor forma de testear la efectividad del método es calcular el porcentaje de aciertos entre la variable observada y la variable imputada, por cada iteración. Los porcentajes de aciertos varían de acuerdo a la cantidad de valores ausentes (*missings*) en cada variable, pero no necesariamente es válido pensar que

serán más efectivas las variables con menor cantidad de valores ausentes.

La efectividad de las imputaciones realizadas con este método oscila entre el 90% y 60% de aciertos entre la variable observada y la variable imputada, para un promedio del 70% de los casos de todas las variables imputadas. En contraste, la capacidad de predicción de un modelo de regresión construido sobre la base de las variables fundamentales³², en promedio no superó una predicción del 40% de los casos de la variable observada para los datos de Chile.

Es importante señalar que la imputación de variables categóricas tiende a alcanzar un mayor porcentaje de aciertos y una menor cantidad de iteraciones para completar todos los casos, a diferencia de las variables continuas, como es el caso de los índices construidos por PISA.

Las variables seleccionadas como explicativas en el común de los países, corresponden a:

- Género.
- Edad.
- Curso.
- Número de estudiantes muestreados por establecimiento.
- Establecimiento público/privado, cuando el país tiene todas las observaciones.
- *Stratum* (clasificación de los establecimientos de acuerdo con sus características).

Entre las variables particulares de cada país se utilizaron:

- Chile
Región y establecimiento Técnico/Profesional
- Uruguay
Proxy según las variables de área (Montevideo/Interior) y comunidad donde se localiza el establecimiento, y establecimiento complejo, los cursos que se imparten en el establecimiento.

³² Como sugieren Fuchs y Woessmann, 2004.

- Polonia

Comunidad donde se localiza el establecimiento, establecimiento complejo, índice de responsabilidad en la asignación de recursos, índice de responsabilidad en el currículo.

- España

Código único nacional de Programa de Estudios y sub-región adjudicada.

Las variables de carácter individual (curso, género y edad) se utilizaron solo para la imputación de variables de la misma característica. Por otra parte, en la imputación de las variables del establecimiento no se las consideró, con el fin de que la imputación asignara el mismo valor para todos los estudiantes del establecimiento.

Con todo, en los cuatro países, para la gran mayoría de las variables, los casos con datos perdidos no superaron el 5%. La variable con mayor cantidad de datos perdidos (alrededor del 20%) se refiere a las expectativas laborales de los estudiantes. Las tablas con la información precisa acerca de la cantidad de datos imputados por variable para cada país, pueden ser solicitadas a los autores.

¿Qué explica las diferencias de resultados PISA Lectura entre Chile y algunos países de la OCDE y América Latina?

Por
Cristián Bellei
Juan Pablo Valenzuela
Alejandra Osses
Alejandro Sevilla¹

Resumen

El objetivo de esta investigación fue identificar algunos factores que explican las diferencias de resultados obtenidos por los estudiantes chilenos en Lectura en la prueba PISA 2006, respecto de los obtenidos por los estudiantes de Polonia, España y Uruguay. Para cumplir este propósito, el estudio implementó una variante de la metodología de descomposición propuesta por Oaxaca-Blinder. Esto implicó, primero, identificar las características de los estudiantes que rindieron la prueba, de sus familias, y de los establecimientos a los que asisten, que fueran explicativas del diferencial observado de puntajes PISA entre Chile y cada uno de los países de referencia. En segundo término, el estudio se centró en estimar si estos factores diferían entre los países, en el modo en que afectaban dichos resultados, pues esta diferencia podría explicar también parte de la brecha observada.

Los hallazgos mostraron que, en la comparación con España y Uruguay, aproximadamente un tercio de la brecha de resultados (negativa y positiva, respectivamente) se explica por diferencias en las características de los estudiantes y los establecimientos. Los dos tercios restantes de la brecha con estos países, así como el total de la brecha con Polonia, tendrían como causa una menor eficiencia de la educación chilena. En términos generales, los resultados obtenidos sugieren que, para avanzar en el mejoramiento de la educación, Chile debe poner renovado énfasis en la calidad de los procesos al interior de las aulas, más que en el equipamiento o las condiciones de administración de los establecimientos.

¹ Cristián Bellei, Juan Pablo Valenzuela, Alejandra Osses y Alejandro Sevilla, son miembros del Centro de Investigación Avanzada en Educación (CIAE) de la Universidad de Chile. Los autores agradecen a Kevin Macdonald, consultor del equipo de educación de la *Human Development Network*, asociada al Banco Mundial, quién facilitó los programas de análisis de PISA para STATA y colaboró en su utilización.

Introducción

Este artículo presenta los hallazgos del análisis sobre los resultados de Lectura, en la prueba PISA 2006, complementarios a los análisis sobre los resultados de Matemática, expuestos en el artículo anterior. En este sentido, es importante precisar que, dado que ambos artículos comparten un mismo diseño y que la muestra de estudiantes y escuelas es la misma, las secciones 1, 2, y 3.1 del capítulo anterior son compartidas con éste, y deben por tanto ser leídas si se quiere conocer en detalle los fundamentos y el diseño de la presente investigación. En este caso particular, utilizando los resultados de aprendizaje en Lectura, medidos por PISA 2006, así como la información complementaria provista por esta misma fuente, este estudio tuvo por objeto identificar algunos de los principales factores que están a la base de las diferencias de logro de los estudiantes chilenos, en comparación con España, Polonia y Uruguay.

El presente artículo se estructura de la siguiente forma. En la primera sección se proveen antecedentes básicos de la investigación; en la sección siguiente se exponen y analizan los principales hallazgos del estudio y por último, se señalan las principales conclusiones del estudio. Adicionalmente, como parte de las conclusiones, se comparan los resultados obtenidos en el análisis de Lectura, con los obtenidos en el análisis de Matemática (artículo anterior) y se plantean algunas reflexiones e implicancias en torno a las políticas educacionales.

1. Antecedentes

Entre las pruebas PISA de los años 2000 y 2006, el promedio de Chile en Lectura, aumentó de 410 a 442 puntos, lo que equivale a 0,32 desviaciones estándar: un cambio muy significativo para este tipo de comparaciones. De hecho, en 2006, de todos los países participantes, Chile fue el que experimentó el mayor avance en sus puntajes, en relación a sí mismo; lo que lo situó en el primer lugar de América

Latina. Con todo, el puntaje obtenido por Chile fue significativamente inferior al promedio de 500 puntos de los países OCDE (la diferencia fue de 0,58 desviaciones estándar en contra de Chile).

En este contexto, como se señaló, el propósito de este estudio fue identificar algunos de los principales factores que explican las diferencias de resultados entre Chile y tres países usados como referencia (Polonia, España y Uruguay), en la prueba PISA de Lectura, aplicada en 2006 (en adelante, “PISA Lectura”). Al respecto, se plantearon dos hipótesis potencialmente complementarias:

1. Las diferencias de resultados entre países se pueden atribuir en parte, a diferencias en características relevantes de dichos países (por ejemplo, estudiantes con más o menos recursos educacionales en sus familias).
2. Las diferencias de resultados se deben en parte, a que los sistemas escolares de dichos países procesan de modo diferente las características de sus estudiantes y establecimientos educacionales, de forma que su “transformación” en resultados de aprendizaje, varía entre países (por ejemplo, más o menos eficiencia en el uso del tiempo escolar).

La metodología del estudio consistió básicamente en implementar una variante de la técnica de descomposición propuesta por Oaxaca-Blinder (Oaxaca, 1973 y Blinder, 1973), la cual permite estimar la fracción de la diferencia observada entre los resultados PISA Lectura de Chile y de los demás países, atribuible a diferencias en las medias de características o factores relevantes, y a diferencias en los retornos o efectos estimados de dichas características o factores². La información utilizada para realizar estas estimaciones fue provista por las bases de datos a nivel de estudiantes, para Chile, España, Polonia y Uruguay, por

2 Esta técnica, desarrollada en estudios sobre economía laboral, ha sido también aplicada a estudios sobre pruebas de aprendizajes. Referencias recientes se encuentran en Ammermüller, 2004; McEwan y Marshall, 2004, y Sakellariou, 2008.



lo que se trabajó con los cinco valores plausibles que PISA proporciona para cada estudiante (véase el Anexo del artículo anterior para más detalle). Estas bases de datos incluyen información complementaria acerca de diversas características de los estudiantes que rinden la prueba y de los establecimientos en que estudian (para una descripción detallada de las variables usadas, véase el Anexo antes mencionado).

Como se sabe, la técnica Oaxaca-Blinder requiere que las mismas variables sean consideradas para todos los grupos que están siendo comparados (en este caso, países). Por ello, se debió implementar un método de imputación de valores para algunos casos con información faltante (para detalles sobre cómo se procedió con las imputaciones véase también el Anexo). En la Tabla 1 se presenta un resumen de las variables incluidas en el estudio. Una explicación detallada de la metodología utilizada se encuentra en el artículo anterior.

Tabla 1.
Variables del estudio

Variable	Definición
VARIABLES DE RESULTADO	
Lectura	Puntaje PISA Lectura, considera los 5 valores plausibles.
Matemática	Puntaje PISA Matemática, considera los 5 valores plausibles.
VARIABLES DEL ESTUDIANTE	
Género	Variable dummy para el género del estudiante (categoría omitida: hombre).
Años de escolaridad	Variable que indica los años de escolaridad formal del estudiante: escala con rango 7 a 11.
Retraso escolar	Variable dummy que indica si el estudiante ha repetido algún curso (categoría omitida: sin retraso escolar).
Índice socioeconómico y cultural del estudiante (ISEC)	Variable normalizada por PISA (se usaron los valores originales) que considera educación y ocupación de los padres y bienes en el hogar.
Expectativas laborales	Expectativa de los estudiantes sobre su futuro laboral: escala con rango 1 (altas) a 3 (bajas).
Valorización sobre la importancia de un buen rendimiento en Matemática y Lenguaje	Valoración que el estudiante asigna a rendir bien en Lenguaje o Matemática: escala con rango 1 (muy importante) a 3 (de menor importancia).

(Continuación en página siguiente)

(Continuación Tabla 1)

Variable	Definición
VARIABLES DEL ESTABLECIMIENTO	
ISEC de los compañeros (efecto par) *	Media del ISEC de los compañeros del establecimiento de cada estudiante (variable normalizada por PISA; se usaron los valores originales).
Tamaño del curso	Variable que indica el tamaño promedio de estudiantes por curso, en el nivel escolar evaluado.
Tamaño del establecimiento	Variable que indica el número de estudiantes matriculados.
Tiempo de instrucción para las asignaturas de Lenguaje y Matemática *	Horas destinadas a las clases regulares de Lenguaje o Matemática en el establecimiento, según reporta el estudiante: escala con rango 1 a 7.
Estudiantes por profesor	Cantidad de estudiantes por profesor al interior del establecimiento.
Establecimiento complejo	Variable dummy que indica si el establecimiento tiene los niveles de educación primaria junto con los de secundaria (variable omitida: establecimiento no complejo).
Índice de calidad de los recursos educativos	Variable normalizada por PISA (se usaron los valores originales) que considera los factores que afectan la enseñanza en el establecimiento, según el director.
VARIABLES INSTITUCIONALES	
Índice de autonomía en las decisiones sobre asignación de recursos	Variable normalizada por PISA (se usaron los valores originales), que se deriva del número de decisiones en relación a la asignación y administración de los recursos, que son de responsabilidad del establecimiento, según el director.
Índice de autonomía en las decisiones curriculares y de evaluación	Variable normalizada por PISA (se usaron los valores originales), que se deriva del número de decisiones en relación al currículum y la evaluación, que son de responsabilidad del establecimiento, según el director.
Criterios académicos de alta prioridad o prerrequisitos para la selección de los estudiantes postulantes al establecimiento	Variable dummy que indica si el establecimiento aplica criterios académicos de selección, como alta prioridad o requisito, según el director (categoría omitida: no considera estos criterios o considera solo el desempeño o la recomendación).
Presión de los padres sobre el establecimiento, por altos estándares	Variable dummy que indica si los padres ejercen presión por altos estándares sobre el establecimiento, según el director (categoría omitida: poco o nada de presión).
Establecimiento público	Variable dummy que indica si el establecimiento es de administración pública (categoría omitida: establecimiento no público).

* Variables medidas a nivel de estudiante.

Finalmente, cabe señalar que, como se explicó en el artículo anterior, el análisis comparativo de las características de los estudiantes y sus establecimientos (incluyendo los aspectos institucionales de éstas), mostró importantes diferencias entre Chile y los demás países, lo cual es, de hecho, el primer paso (y en cierta forma un prerrequisito) para la aplicación del método de descomposición de Oaxaca-Blinder. Dado que el set de variables utilizadas en el análisis, es común para ambas asignaturas, los resultados de dicho análisis deben ser tenidos en cuenta al momento de interpretar los hallazgos de la descomposición de las diferencias en Lectura. Las únicas excepciones (ya analizadas en el artículo anterior) son el nivel de valoración que los estudiantes asignan a obtener un buen rendimiento y la cantidad de horas semanales de instrucción; ambas variables fueron medidas por separado para cada una de las asignaturas. Del análisis, se encontró que, en promedio, los estudiantes chilenos: (a) poseen una mayor valoración del hecho de obtener un buen rendimiento en Lenguaje, que sus pares de los otros tres países, y (b) tienen menos horas semanales promedio de instrucción en Lenguaje, que los estudiantes polacos y españoles, pero más horas que los uruguayos.

2. Resultados

2.1 Factores explicativos de los resultados de PISA Lectura

Constatadas diferencias en las características entre Chile y los demás países, corresponde indagar acerca de eventuales diferencias en los efectos que dichas características tienen sobre los resultados de PISA Lectura. Lo anterior, constituye el fundamento para atribuir a dichas diferencias, parte de la brecha observada entre los resultados de Chile y los del resto de los países analizados.

En la Tabla 2 se presentan los resultados de cuatro regresiones múltiples que asocian, separadamente

para cada uno de los países, los puntajes de PISA Lectura, con el conjunto de variables incluidas en este estudio. El propósito de este análisis es, primero, identificar un conjunto de factores explicativos de los resultados obtenidos por los estudiantes en Lectura y, a continuación, verificar si esta asociación entre factores y resultados difiere en alguna medida entre Chile y los países de referencia.

En términos generales, las estimaciones muestran que la mayor parte de las variables que miden características de los estudiantes, son factores relevantes en la explicación de los resultados de Lectura de Chile; en contraste, sólo dos de las variables que miden aspectos del establecimiento y una de las referidas a cuestiones institucionales, poseen una asociación estadísticamente significativa con el puntaje PISA Lectura de este país.

Específicamente, en cuanto a las características de los estudiantes, las mujeres obtienen 15 puntos más que los hombres, y los estudiantes que manifiestan tener altas expectativas laborales, obtienen 18 puntos más que quienes dicen tener bajas expectativas³. El nivel socioeconómico y cultural de la familia, por su parte, también se asocia positivamente con el puntaje (el efecto estimado es de 8,2 puntos, por cada desviación estándar de diferencia en el ISEC). El grado que se encuentran cursando los estudiantes se asocia de un modo sistemático: a mayor escolaridad, mayor puntaje (la diferencia estimada por cada año de escolaridad, entre 7° y 10° grados, es de alrededor de 40 puntos). Finalmente, no se encontró una relación estadísticamente significativa entre los resultados y la valoración que los estudiantes manifiestan respecto de la importancia de obtener buenos logros en Lenguaje.

En cuanto a las variables que miden características de los establecimientos, las estimaciones presentadas en la Tabla 2 indican que sólo la medida del efecto par

3 Recuérdese que se habla de "efectos netos", puesto que estas estimaciones son hechas controlando simultáneamente, por el efecto del resto de las variables incorporadas en la Tabla 2.

(“ISEC de los compañeros”) y las horas semanales de instrucción de Lenguaje, son factores asociados, de manera estadísticamente significativa, a los resultados; en ambos casos, de manera positiva.

Por último, de las variables que describen características institucionales de los establecimientos, sólo aquella referida a la selectividad académica en la admisión de los estudiantes, está asociada de manera estadísticamente significativa a los resultados: en promedio, los estudiantes chilenos que asisten a establecimientos donde se aplican procedimientos de selección académica, obtuvieron 17 puntos más que sus pares que no asisten a este tipo de establecimientos. Así, ni la autonomía en el uso de los recursos ni la autonomía en materias curriculares y evaluativas,

son factores relevantes para explicar los resultados de los estudiantes chilenos. Tampoco lo es el grado de presión de los padres hacia el establecimiento por lograr altos estándares ni (una vez controladas las demás diferencias) la dependencia administrativa de los establecimientos (públicos o privados).

En comparación con los análisis de los resultados de Matemática (expuestos en el artículo anterior), en Lectura, las variables explicativas significativas son menos en número y tienen en conjunto una menor capacidad explicativa de los resultados a nivel individual. Así, en Lectura, el modelo de análisis logró dar cuenta del 40% de la varianza observada en los resultados entre estudiantes chilenos (en Matemática, 53%).

Tabla 2.
Factores a nivel de estudiante, establecimiento e institucionales, explicativos de los resultados PISA Lectura.
Coeficientes de regresión para Lectura^(a)

	Coeficiente			
	Chile	España	Uruguay	Polonia
VARIABLES A NIVEL DE ESTUDIANTE				
Mujer	14,59*** (3,72)	20,19*** (1,84)	23,69*** (5,14)	22,22*** (2,78)
8° Grado	39,88* (20,77)	176,38** (38,36)	14,22 (12,75)	73,59*** (23,68)
9° Grado	81,54*** (24,15)	219,24*** (38,45)	46,55*** (10,44)	163,65*** (25,90)
10° Grado	123,26*** (25,13)	280,77*** (38,24)	94,59*** (17,25)	216,66*** (31,49)
11° Grado	137,6*** (26,22)	394,72*** (48,11)	107,61*** (19,86)	N/A
Retraso escolar	-6,66 (6,03)	-6,08 (9,12)	-19,40 (14,58)	-10,05 (13,95)
ISEC (nivel socioeconómico y cultural familiar)	8,20*** (2,08)	8,02*** (1,26)	10,88*** (1,86)	26,13*** (1,55)
Altas expectativas laborales	18,39*** (5,10)	22,48*** (2,83)	33,03*** (6,85)	54,92*** (3,19)
Expectativas laborales medias	3,22 (5,96)	6,41** (3,14)	25,55** (8,62)	23,76*** (3,32)
Valoración rendimiento Leng.: muy importante	9,05 (6,88)	5,11* (2,80)	7,81 (5,61)	10,53* (5,25)
Valoración rendimiento Leng.: importante	2,19 (7,01)	2,51 (3,32)	1,86 (6,26)	0,31 (5,10)

(Continuación en página siguiente)

(Continuación Tabla 2)

	Coeficiente			
	Chile	España	Uruguay	Polonia
VARIABLES A NIVEL DEL ESTABLECIMIENTO				
ISEC de los compañeros (efecto par)	37,39*** (5,23)	13,56*** (4,05)	22,95*** (6,79)	15,28** (7,17)
Establecimiento complejo	-0,18 (6,53)	0,89 (5,62)	11,634* (6,03)	20,08* (10,20)
Tamaño del curso	0,30 (0,62)	0,21 (0,20)	0,44 (0,29)	1,05 (0,75)
Tamaño del establecimiento	0,00 (0,01)	-0,013** (0,01)	0,01 (0,01)	0,00 (0,01)
Horas instrucción Lenguaje	6,18*** (0,68)	6,59*** (0,66)	6,67*** (1,15)	7,93*** (0,79)
Estudiantes por profesor	-0,29 (0,58)	0,14 (0,83)	0,39 (0,79)	-1,94 (1,32)
Recursos educativos	2,78 (2,79)	0,69 (1,94)	-0,77 (2,24)	1,84 (2,34)
VARIABLES A NIVEL INSTITUCIONAL				
Autonomía en el uso de los recursos	5,85 (4,18)	5,03 (4,30)	7,09 (7,03)	1,90 (3,98)
Autonomía en currículum y evaluación	-4,97 (3,12)	-1,57 (2,14)	-0,85 (8,62)	-6,98** (2,69)
Selección académica	17,41** (6,73)	9,88 (6,87)	9,30 (10,79)	9,33 (6,85)
Expectativas de padres por altos estándares	6,81 (9,83)	2,69 (6,67)	4,79 (10,11)	-1,75 (4,91)
Establecimiento público	6,49 (6,97)	2,99 (9,62)	-5,58 (14,72)	10,18 (23,84)
Constante	305,38*** (28,77)	162,09*** (43,22)	281,49*** (21,69)	259,83*** (36,01)
Número de observaciones	5.224	19.604	4.839	5.528
R ² promedio (b)	0,40	0,39	0,38	0,33

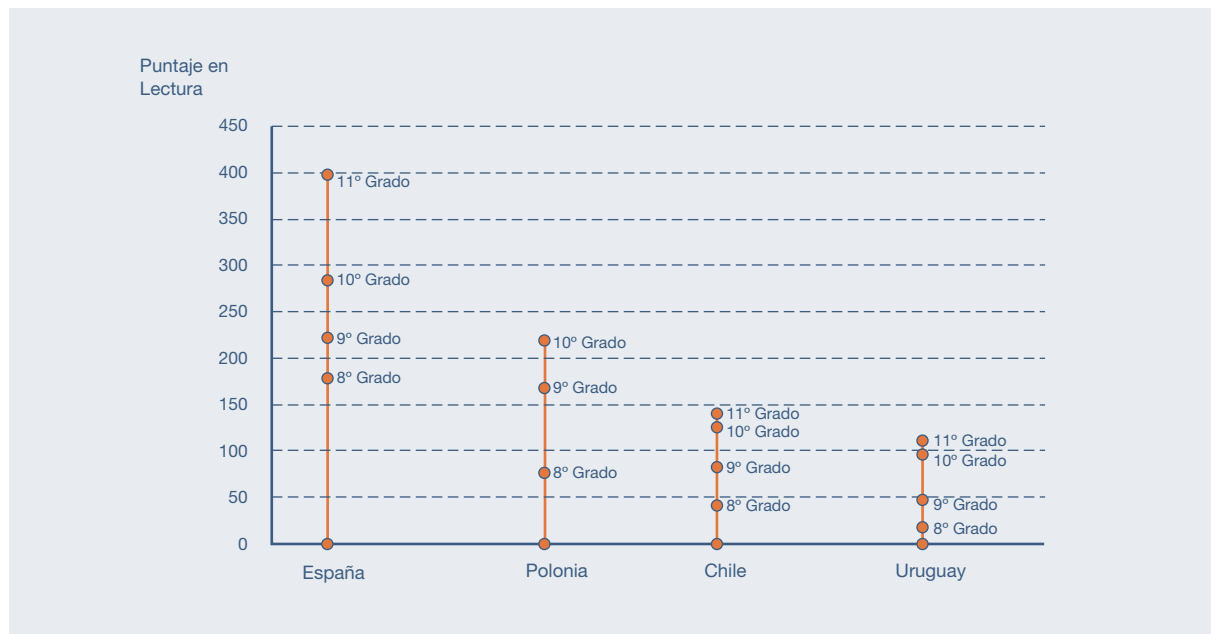
Los coeficientes son estadísticamente significativos al: *** < 1%; ** < 5%; * < 10%

(a) Coeficientes estimados en regresiones separadas para cada país.

(b) Corresponde al valor promedio, pues se realizan 5 regresiones, una por cada valor plausible (*plausible value*).

Gráfico 1:

El grado como factor explicativo de los resultados PISA Lectura en Chile, España, Polonia y Uruguay.



Diferencia neta estimada entre estudiantes de diferentes grados (ver Tabla 2 para valores exactos; categoría de referencia: alumnos de 7° Grado).

En términos generales, de la Tabla 2 se puede concluir que los resultados de este análisis son similares para los demás países de referencia, en cuanto a la estructura del modelo de análisis (factores estadísticamente significativos —con algunas excepciones— y sentido de las relaciones con los resultados), pero que existen importantes diferencias en la magnitud de los coeficientes estimados para las variables analizadas. El factor donde estas diferencias son más marcadas, es el grado que se encuentran cursando los estudiantes. En el Gráfico 1 se presentan los coeficientes asociados a esta variable, para los cuatro países analizados.

Como puede apreciarse, al igual que en Chile, en todos los países, los estudiantes de grados superiores obtienen mejores puntajes; sin embargo, la magnitud de las diferencias estimadas entre estudiantes de

diversos grados, al interior de un mismo país, varía enormemente⁴.

Así por ejemplo, mientras los estudiantes chilenos de 9° Grado obtienen en promedio 42 puntos más que sus pares de 8° Grado, los estudiantes polacos de 9° Grado superan en promedio a sus pares de 8° en 90 puntos. Del mismo modo, la brecha estimada entre los estudiantes de 10° y 9° Grado en Chile es de 42 puntos, mientras que en España, es de 62 puntos. Es decir, el diferencial de resultados entre los estudiantes de diferentes grados es mayor en los países europeos que en Chile. Lo contrario sucede en la comparación con Uruguay, donde la ventaja estimada para los estudiantes de grados superiores tiende a ser más pequeña que en Chile.

4 Recuérdese que los coeficientes para cada grado, están estimados según su diferencia con los estudiantes de 7° Grado, y controlando por todos los demás factores.

Aunque menos pronunciadas que las asociadas al grado escolar, también se observan diferencias en otras variables. Por ejemplo, entre las variables que miden características de los estudiantes, se observa que las mujeres chilenas son las que obtienen, comparativamente, la menor ventaja respecto de los hombres. Asimismo, el efecto estimado del nivel socioeconómico y cultural individual en Polonia, es tres veces superior mayor al estimado en Chile. Lo mismo ocurre con las expectativas laborales de los estudiantes: Chile es el país en que esta asociación (siempre positiva) es más débil, siendo particularmente notorias las diferencias con Polonia y Uruguay.

En cuanto a los efectos estimados de las características de los establecimientos, la diferencia más marcada es la del ISEC de los compañeros, siendo Chile el país en que este “efecto par” es más fuerte: más del doble que en España y Polonia, y aproximadamente, 60% mayor que en Uruguay. Otra diferencia importante a nivel del establecimiento está asociada a si éste es o no complejo: en Polonia y Uruguay, los estudiantes que asisten a establecimientos complejos (es decir, que imparten Enseñanza Básica y Media) obtienen en promedio 20 y 12 puntos más, respectivamente, que sus pares que asisten a establecimientos no complejos. Esta diferencia no se observa en Chile ni en España.

Por último, en lo referido a los efectos estimados de las variables que miden aspectos institucionales, las diferencias entre Chile y los demás países son muy menores. En España y Uruguay, estas variables no representan un factor relevante para explicar los resultados de sus estudiantes, y en Polonia, sólo la autonomía curricular está asociada de un modo estadísticamente significativo, aunque en sentido negativo: a mayor autonomía, menores resultados. En Chile, por su parte, la única variable institucional que resulta explicativa de los resultados es la selección académica de los estudiantes: el coeficiente es casi el

doble que para los demás países (en los cuales, además, no es una variable estadísticamente significativa).

En resumen, es posible afirmar que existen diferencias importantes entre Chile y los tres países analizados, en el efecto estimado sobre los resultados de un conjunto relevante de características, especialmente de los estudiantes, aunque también de los establecimientos e institucionales. Como se dijo anteriormente, una de las hipótesis que sustenta este estudio es que estas diferencias de “productividad” (efectos) de los factores asociados al aprendizaje, que se dan entre Chile y los demás países, explican una fracción de la brecha observada de resultados en Lectura. Comprobar esta hipótesis y estimar la magnitud de dicha fracción es el objetivo de la siguiente sección.

2.2 Descomposición de resultados de PISA Lectura

2.2.1 Descomposición general de la diferencia de resultados entre Chile y los tres países de referencia

La lógica básica de este análisis es estimar cuánto de la brecha observada entre los resultados de Chile y cada país de referencia, es posible de ser explicada por diferencias en las características de los países. De allí que, por ahora, el diferencial “no explicado” de la brecha se puede atribuir como un todo, a diferencias entre los países en la eficiencia en el uso de los recursos disponibles (en un sentido amplio). La descomposición detallada que se presenta más adelante, permitirá hacer distinciones más finas al respecto.

En el Gráfico 2, donde se muestran los resultados de este análisis⁵, la barra celeste indica la brecha de resultados entre Chile y cada país analizado, y las barras naranja y azul, indican la fracción de esta brecha que es atribuible a las diferencias en características

⁵ En la Tabla 3 se encuentran los valores exactos.

o factores, y a las diferencias en efectos o retornos, respectivamente. Como puede apreciarse, cada comparación presenta un patrón diferente de resultados, por lo que conviene analizarlos por separado.

En la prueba PISA Lectura, España obtuvo en promedio 18,3 puntos más que Chile y, como se observa en el gráfico, aproximadamente un tercio de esa diferencia (5,9 puntos) puede ser atribuida a diferencias en las características de los estudiantes y de los establecimientos de ambos países. Esto sugiere que, en promedio, los estudiantes de España poseen condiciones más propicias para el aprendizaje; dicho de otro modo, si los establecimientos y estudiantes chilenos tuviesen las características observadas de los establecimientos y estudiantes españoles, el puntaje promedio de Chile habría sido 5,9 puntos superior al obtenido. Aunque las diferencias en las características observadas que separan a Chile de España juegan un papel relevante, es preciso notar que dos tercios de la brecha entre estos dos países, no puede ser explicada por dichas diferencias.

La comparación con Uruguay es, en cierto modo, un “espejo” de la comparación con España: Chile supera en promedio a Uruguay, por 30 puntos, y un tercio de esa brecha se explica por diferencias en las características de los estudiantes y de los establecimientos. Así, si los estudiantes y establecimientos chilenos tuviesen las características de los uruguayos, el promedio de Chile habría sido 10 puntos más bajo que el obtenido. Ahora bien, al igual que con España, dos tercios de la ventaja de Chile sobre Uruguay no está asociado a diferencias entre ambos países, en los factores observados.

Finalmente, los resultados del análisis comparativo con Polonia indican que las diferencias entre ambos países no juegan ningún rol en la explicación de la brecha de resultados; más aún, estos afectan en el sentido inverso. En otras palabras, si los estudiantes y los establecimientos chilenos tuviesen las características de

los polacos, la brecha de resultados habría aumentado en 9,1 puntos (aproximadamente 14%). Esto implica que la explicación de toda la ventaja que muestra Polonia sobre Chile, debe buscarse en la productividad de esos factores (es decir, en sus retornos) o en otras variables no observadas en este estudio.

En resumen, los análisis de la descomposición general de la brecha de puntajes señalan que, tanto para explicar los mejores logros de Chile frente a Uruguay, como los peores logros frente a España, una fuente relevante son las diferencias en las características de los estudiantes y de los establecimientos. En ambos casos, estas diferencias explican un tercio de la brecha de resultados en la prueba PISA Lectura. La comparación con Polonia, sin embargo, muestra que la brecha no puede ser atribuida a las características de los establecimientos ni de los estudiantes.

Gráfico 2: Descomposición de la diferencia de resultados PISA Lectura 2006 entre Chile y los tres países de referencia

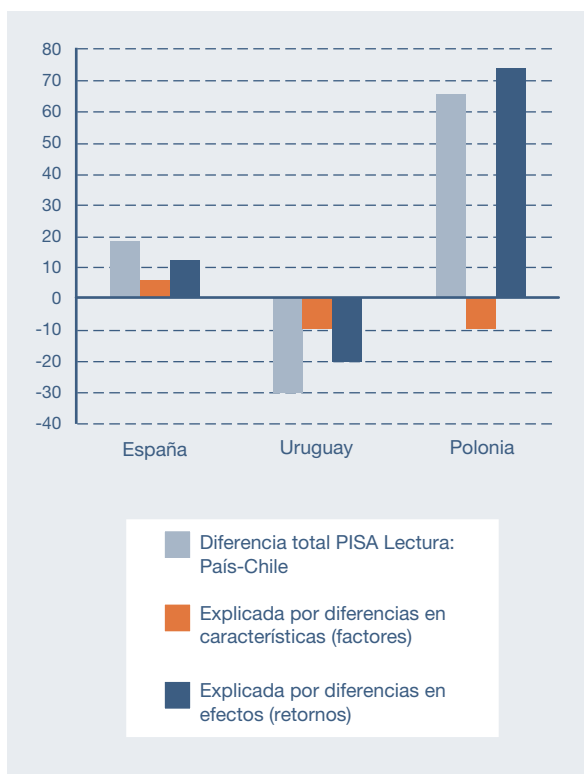


Tabla 3.
Descomposición general de la diferencia entre Chile y tres países de referencia. PISA Lectura

	España	Uruguay	Polonia
Diferencia total PISA Lectura: Chile	18,3	-30,0	65,0
Explicada por diferencias en características (factores)	5,9	-10,0	-9,1
Explicada por diferencias en efectos (retornos)	12,5	-20,0	74,1
Proporción diferencia explicada (factores)	31,9 %	33,2 %	-14,1 %
Proporción diferencia no explicada (retornos)	68,1 %	66,8 %	114,1 %

2.2.2 Descomposición detallada de la diferencia de resultados entre Chile y los tres países de referencia

A continuación se descomponen en forma detallada las fuentes de la diferencia de puntajes entre Chile y los tres países de referencia, distinguiendo, en cada una de las variables estudiadas, entre el efecto asociado a su nivel, el asociado a su retorno y el efecto de la interacción entre nivel y retorno, así como el aporte específico de la constante en los modelos de regresión empleados. Los resultados de la descomposición por país, se presentan en la Tabla 4, en la cual se organizan las variables, según la dimensión que describen (estudiante, establecimiento o nivel institucional). Primero se analizará el efecto estimado asociado a las diferencias de características, para luego proceder al análisis del efecto estimado asociado a las diferencias de retorno e interacción.

Como se observa en la Tabla 4, tomadas en conjunto, las diferencias con España y Polonia en el nivel de las variables institucionales no permiten explicar nada de la brecha que los separa con Chile (más aún, su efecto, aunque pequeño, apunta en la dirección opuesta, pues tendería a aumentar en lugar de disminuir la brecha). En el caso de Uruguay, sólo 2 puntos de los 30 de ventaja que obtuvo Chile, son atribuibles a esta dimensión.

Si se observa más detenidamente, se comprueba que, del conjunto de variables institucionales, sólo

aquella referida a la selectividad académica de los establecimientos, posee un efecto estadísticamente significativo en la explicación de las diferencias de resultados entre Chile y el resto de los países de referencia. La estimación sugiere que, si el porcentaje de estudiantes chilenos que asiste a establecimientos académicamente selectivos fuese equivalente al de los demás países (todos ellos menos selectivos), los resultados promedio de Chile disminuirían entre 4 y 6 puntos, aumentando así la brecha con Polonia y España, y disminuyendo la ventaja sobre Uruguay. El análisis de las variables referidas a características de los establecimientos, en cambio, proporciona evidencia mucho más relevante para comprender la distancia que separa a Chile de Polonia y España. Tomadas en conjunto, las estimaciones indican que, si los establecimientos chilenos tuviesen las características de los establecimientos españoles o polacos, el promedio de los estudiantes chilenos habría sido 20 y 26 puntos superior, respectivamente. Este efecto equivale al total de la brecha de resultados con España y a un 40% de la brecha con Polonia.

Ahora bien, contrario a lo que comúnmente se piensa, no son los recursos educativos, el tamaño del curso o la cantidad de docentes disponibles, las características que explican la diferencia de resultados con los países europeos; sino el nivel socioeconómico y cultural de los compañeros y las horas semanales de instrucción que reciben los estudiantes.

Tabla 4. Descomposición de los efectos sobre el puntaje PISA Lectura, asociados a las características, los retornos y la interacción entre ambos. Análisis separados para la comparación de Chile con España, Uruguay y Polonia

	España			Uruguay			Polonia		
	Efectos asociados a			Efectos asociados a			Efectos asociados a		
	Características	Retornos	Interacción (caract. x retorno)	Características	Retornos	Interacción (caract. x retorno)	Características	Retornos	Interacción (caract. x retorno)
Constante		-143,29** (49,84)			-23,89 (36,68)			-45,55 (46,73)	
Variables a nivel del estudiante									
Mujer	0,49* (0,26)	2,58 (1,87)	0,19 (0,20)	0,74*** (0,27)	4,19 (3,03)	0,46 (0,38)	0,62** (0,27)	3,52 (2,32)	0,32 (0,28)
Curso	-9,06*** (1,55)	157,80*** (39,40)	-8,40*** (2,90)	-12,97*** (2,90)	-29,59 (30,26)	2,18 (3,09)	-32,95*** (4,37)	74,49* (37,75)	5,34 (11,97)
Retraso escolar	0,91 (0,83)	0,11 (1,99)	-0,08 (1,46)	-0,94 (0,86)	-2,38 (2,94)	-1,79 (2,20)	1,06 (0,96)	-0,63 (2,84)	0,54 (2,42)
ISEC (nivel socioeconómico familiar)	3,17*** (1,00)	0,13 (1,75)	-0,07 (0,98)	1,55** (0,68)	-1,87 (2,15)	0,51 (0,63)	3,24*** (0,99)	-12,51*** (1,98)	7,09*** (1,44)
Expectativas laborales	-1,53*** (0,47)	3,25 (4,96)	-0,21 (0,39)	0,62 (0,39)	13,05* (6,72)	0,66 (0,48)	-2,29*** (0,57)	27,94*** (4,70)	-3,17*** (0,98)
Valoración de un buen rendimiento en Lenguaje	-1,94* (0,97)	-2,79 (7,16)	1,03 (1,21)	-2,11** (1,03)	-0,97 (8,43)	0,29 (1,42)	-2,24** (1,06)	0,66 (7,73)	-0,92 (1,27)
Subtotal estudiantes	-7,96*** (2,65)	161,08*** (41,83)	-7,54** (3,74)	-13,10*** (3,39)	-17,57 (31,46)	2,30 (3,28)	-32,55*** (5,29)	93,47** (38,96)	9,21 (12,58)
Variables a nivel del establecimiento									
ISEC de los compañeros (efecto par)	15,15*** (3,05)	16,60*** (4,73)	-9,66*** (2,96)	7,09*** (2,60)	10,06 (6,32)	-2,74 (1,94)	14,86*** (2,95)	15,40** (6,17)	-8,78** (3,61)
Establecimiento complejo	0,05 (1,90)	0,51 (4,07)	-0,31 (2,53)	0,05 (1,68)	5,62 (4,29)	-3,06 (2,35)	0,08 (3,00)	9,64* (5,08)	-9,35* (4,89)
Tamaño del curso	-3,09 (6,36)	-3,35 (25,39)	0,91 (6,90)	-1,49 (3,07)	5,42 (25,53)	-0,71 (3,37)	-3,71 (7,66)	28,60 (36,16)	-9,36 (11,76)
Tamaño del establecimiento	1,02 (2,03)	-10,73 (7,86)	3,63 (2,70)	1,75 (3,52)	16,06 (14,71)	-9,40 (8,70)	1,83 (3,68)	1,01 (13,09)	-0,62 (8,02)
Horas de instrucción Lenguaje	0,96** (0,43)	1,40 (3,30)	0,06 (0,15)	-4,40*** (0,67)	1,67 (5,09)	-0,35 (1,06)	7,40*** (0,86)	6,03 (3,67)	2,10 (1,30)

(Continuación en página siguiente)

(Continuación Tabla 4)

	España			Uruguay			Polonia		
	Efectos asociados a			Efectos asociados a			Efectos asociados a		
	Características	Retornos	Interacción (caract. x retorno)	Características	Retornos	Interacción (caract. x retorno)	Características	Retornos	Interacción (caract. x retorno)
Variables a nivel del establecimiento									
Estudiantes por profesor	3,68 (7,44)	10,70 (25,19)	-5,43 (12,77)	2,66 (5,39)	17,07 (23,87)	-6,25 (8,78)	3,98 (8,06)	-41,24 (36,85)	22,64 (20,23)
Recursos educativos	1,69 (1,65)	1,30 (2,09)	-1,27 (2,01)	-0,24 (0,48)	2,22 (2,04)	0,30 (0,61)	1,49 (1,43)	0,59 (2,55)	-0,50 (2,15)
Subtotal establecimientos	19,45*** (7,28)	16,44 (26,86)	-12,06 (11,18)	5,42 (5,68)	58,12* (30,38)	-22,20* (11,30)	25,94*** (8,47)	20,04 (41,57)	-3,89 (20,61)
Variables a nivel institucional									
Autonomía en el uso de los recursos	1,53 (1,15)	0,20 (1,57)	-0,22 (1,67)	-3,79 (2,76)	-0,30 (1,94)	-0,80 (5,18)	-0,03 (0,52)	0,97 (1,57)	0,02 (0,39)
Autonomía en currículum y evaluación	-2,39 (1,59)	-0,81 (0,93)	1,64 (1,81)	4,30 (2,71)	-0,98 (2,19)	-3,56 (7,50)	-2,71 (1,83)	0,48 (0,88)	-1,10 (1,97)
Selección académica	-5,46** (2,11)	-2,71 (3,14)	2,36 (2,72)	-4,67** (1,79)	-2,92 (4,67)	2,17 (3,51)	-3,62** (1,57)	-2,91 (3,30)	1,68 (1,89)
Expectativas de padres por altos estándares	-0,67 (0,95)	-0,77 (2,29)	0,41 (1,19)	-0,81 (1,16)	-0,38 (2,49)	0,24 (1,58)	0,29 (0,51)	-1,60 (2,02)	-0,37 (0,64)
Establecimiento público	1,35 (1,45)	-1,53 (5,27)	-0,73 (2,46)	2,68 (2,86)	-5,29 (7,17)	-4,99 (6,78)	3,54 (3,79)	1,62 (10,98)	2,02 (13,80)
Subtotal institucional	-5,64 (3,65)	-5,62 (5,96)	3,46 (5,17)	-2,28 (4,58)	-9,87 (8,49)	-6,94 (9,03)	-2,53 (4,83)	-1,44 (12,24)	2,26 (13,65)
TOTAL	5,85 (8,29)	28,61*** (9,61)	-16,14 (10,30)	-9,95 (8,01)	6,80 (15,25)	-26,84 (16,61)	-9,14 (10,37)	66,52** (26,67)	7,58 (29,00)

Los coeficientes son estadísticamente significativos al: *** < 1%; ** < 5%; * < 10%

Así, por ejemplo, si los estudiantes chilenos recibiesen la cantidad de instrucción en Lenguaje, que reciben sus pares polacos (esto es, 1,2 horas semanales más que los chilenos), habrían obtenido alrededor de 7 puntos más en la prueba. Por su parte, el efecto estimado del nivel socioeconómico y cultural de las familias, a nivel del establecimiento (“efecto par”) es aun mayor: aproximadamente 15 puntos de la brecha de resultados entre Chile y Polonia y España, puede ser atribuida a este factor. Esto representa alrededor de 80% de la brecha con España y 20% de la brecha con Polonia.

Los resultados de la comparación con Uruguay, en cambio, no son tan nítidos. Aunque estas mismas dos variables son las únicas cuya contribución a la descomposición es estadísticamente significativa, ambas tienden a neutralizarse, pues poseen sentidos opuestos. Es interesante, sin embargo, constatar nuevamente la relevancia que tiene el tiempo de instrucción en los resultados. En efecto, alrededor de 4 puntos de la ventaja de los estudiantes chilenos sobre los uruguayos, pueden ser atribuidos a las 0,7 horas adicionales de enseñanza semanal que los chilenos reciben en esta asignatura.

La última dimensión a considerar en el análisis de los efectos asociados a las diferencias de características, se refiere a las variables que describen a los estudiantes en términos individuales. De acuerdo con la información de la Tabla 4, se puede afirmar que, si los estudiantes chilenos tuviesen las características de los estudiantes de los países de referencia (tomadas en conjunto), sus resultados promedio habrían sido más bajos (con un rango de entre 8 y 33 puntos menos, dependiendo del país de referencia). En ese contexto, habría aumentado la brecha con España y Polonia, y habría disminuido la diferencia con Uruguay. En el caso de este último país, las estimaciones indican que, un poco menos de la mitad de la ventaja lograda por Chile, puede ser explicada por las características más favorables de los estudiantes chilenos; en particular, por el hecho

de que los estudiantes chilenos evaluados están mayoritariamente cursando grados superiores que los uruguayos (13 de los 30 puntos de la ventaja chilena, pueden ser explicados por este factor).

La comparación con España y Polonia, en cambio, sugiere que la superioridad en los resultados de esos países sobre Chile, no se relaciona con diferencias en las características de los estudiantes evaluados. Por el contrario, los estudiantes chilenos tienden a mostrar ventajas sobre sus pares europeos en esta dimensión. Similar a la comparación con Uruguay, la más importante de estas ventajas, es el hecho de que los estudiantes chilenos estuvieran cursando grados superiores, de forma que, de igualarse con los españoles o polacos en este aspecto, habrían obtenido 9 y 33 puntos menos, respectivamente. Aunque en una menor magnitud, también aportan en este mismo sentido las dos variables “subjetivas” consideradas en el estudio: las expectativas laborales y la valoración del buen rendimiento en Lenguaje, por parte de los estudiantes. A estas variables, tomadas en conjunto, se asocian negativamente entre 3 y 5 puntos de la brecha con España y Polonia, respectivamente, es decir, los resultados de los estudiantes chilenos tienden a ser comparativamente mejores debido a estas características.

Finalmente, el nivel socioeconómico de las familias de los estudiantes sí es un factor explicativo de la brecha de resultados a favor de España y Polonia, aunque tomado aisladamente de magnitud modesta. Sin embargo, dado que esta misma variable es luego considerada colectivamente como una característica del establecimiento (en el ISEC de los compañeros), es posible afirmar que aproximadamente 18 puntos de la ventaja de España y Polonia sobre Chile se explican por las mejores condiciones socioeconómicas y culturales de los estudiantes europeos.

Ahora bien, como se ha dicho, las brechas de resultados pueden explicarse no sólo por diferencias

en la “cantidad” de las variables observadas, sino por el “modo” en que estas variables se relacionan con los resultados, es decir, por la diferente eficiencia con que los factores son “transformados” en resultados de aprendizaje. Estas diferencias de eficiencia de los sistemas educativos, son estimadas en este análisis, usando los efectos de retorno e interacción, también incluidos en la Tabla 4.

Al respecto, en primer lugar, destaca el hecho de que, en la comparación con los tres países estudiados, los efectos de retorno e interacción no fueron estadísticamente significativos en las variables institucionales, es decir, las diferencias de resultados entre Chile y estos países, no se explican por diferencias en la productividad de las variables institucionales. Con todo, es interesante notar que en la comparación con Uruguay, el efecto acumulado del retorno de las variables institucionales es de 10 puntos negativos, sugiriendo que Chile obtiene parte de su ventaja con ese país por la vía de ser más eficiente en el “uso” de estas características institucionales. La variable más relevante en este aspecto está dada por la mayor efectividad comparativa de los establecimientos públicos chilenos respecto de los uruguayos; paradójicamente, el signo negativo de la interacción asociada a este factor, recuerda que en Chile, proporcionalmente menos estudiantes asisten a establecimientos públicos.

En términos generales, las diferencias de retornos asociadas a las variables de los establecimientos,

6 Es importante advertir, sin embargo, que aunque no alcanzan el nivel de significación estadística, varios de los coeficientes asociados a los retornos de variables a nivel del establecimiento, poseen magnitudes relevantes, aunque de variados signos. Por ejemplo, la cantidad de estudiantes por profesor es un factor usado en Chile menos eficientemente que en España y Uruguay, pero más que en Polonia. El tamaño del establecimiento es un factor usado en Chile con menos eficiencia que en Uruguay, pero con más eficiencia que en España. Finalmente, el tamaño del curso es un factor usado en Polonia, con más eficiencia que en Chile. Todas estas asociaciones generan hipótesis interesantes de continuar indagando, quizás con mediciones más precisas. De hecho, en el caso de la comparación con Uruguay, la suma total de estas diferencias de retorno de las variables a nivel del establecimiento, genera un efecto estimado de 58 puntos que Chile “perdería” por ser menos eficiente que Uruguay.

tampoco son una fuente explicativa relevante de las brechas de resultados entre Chile y los países de referencia. En efecto, ninguno de los coeficientes de estas variables es estadísticamente significativo⁶. La única excepción es el efecto asociado al retorno y a la interacción del “ISEC de los compañeros”, pero solo para las comparaciones con España y Polonia. En estos casos, los resultados indican que, si en Chile el nivel socioeconómico y cultural de los compañeros tuviese el retorno que en España o Polonia, la brecha de puntajes con estos países aumentaría⁷. El valor negativo de la interacción asociada a este “efecto par”, indica que estos mayores retornos en Chile se contrarrestan con el hecho de que los estudiantes chilenos poseen en promedio un nivel socioeconómico y cultural más bajo que los españoles y polacos.

Las diferencias estimadas de retorno, de las variables que describen características de los estudiantes, en cambio, son la principal fuente explicativa de las brechas observadas entre los resultados de Chile y Polonia o España, explicando por sí mismas, el total de esas brechas.

La descomposición detallada de la brecha con España, muestra claramente que las diferencias se concentran prácticamente en su totalidad en una sola variable: el retorno asociado a los grados escolares que los estudiantes cursaban al momento de ser evaluados. En términos sustantivos como se muestra en el Gráfico 1 el avance en los resultados de Lectura de España, asociado a un grado adicional de escolaridad,

7 El signo del coeficiente puede inducir a error, puesto que la descomposición detallada de “retornos” es sensible a los signos usados en la construcción de la variable original. En términos intuitivos, Chile tiene un coeficiente de regresión asociado al ISEC de los compañeros, mayor que el de España y Polonia (ver Tabla 2), por lo que este factor debería aportar un valor negativo a la descomposición (es decir, Chile debería aparecer siendo “más eficiente” que los países europeos). El signo positivo, en cambio, se produce porque el valor de Chile en el ISEC es negativo, dada la forma de estandarización de esta variable construida por PISA. Ciertamente, lo mismo es válido para el coeficiente asociado al retorno del ISEC a nivel individual, el cual también debe ser leído con el sentido opuesto a su signo.

es considerablemente mayor que en Chile, al punto que, si en Chile los grados de escolaridad tuviesen la productividad que en España, el puntaje chileno promedio habría aumentado en una magnitud varias veces superior a la brecha total entre ambos países. Esto hace que se diluya completamente la poca ventaja obtenida por Chile, atribuible a que sus estudiantes de 15 años cursan en promedio grados superiores a los españoles.

La comparación detallada con Polonia es más compleja. En primer lugar, al igual que con España, en Chile, el retorno asociado a los grados escolares adicionales, es menor que en Polonia, al punto de que, por sí sola, esta variable da cuenta de la totalidad de la brecha entre ambos países. En este caso, sin embargo, debido a las grandes diferencias de escolaridad que favorecen a Chile, el *trade off* entre mayor escolarización y menor productividad es más notorio (es decir, Chile “recupera” casi la mitad de la desventaja en eficiencia, aumentando los insumos). En segundo lugar, cabe destacar que, en la comparación con Polonia, la diferencia de efectos asociada a los mejores resultados de los estudiantes con mayores expectativas laborales, explica una parte importante de la brecha observada. Así, si las mayores expectativas laborales de Chile tuvieran retornos similares a los observados en Polonia, el promedio nacional en Lectura habría sido 28 puntos superior (esta variable es también relevante y tiene el mismo sentido, en el análisis de la diferencia de productividad con Uruguay). Por último, en Polonia, los estudiantes con mejor nivel socioeconómico y cultural tienen retribuciones, en términos de resultados, muy superiores a sus pares chilenos⁸. Este efecto

⁸ Ciertamente, un mayor retorno de las mejores condiciones socioeconómicas y culturales de los estudiantes puede interpretarse, tanto como un signo de mayor inequidad (mayor “herencia” familiar en educación), como de mayor eficiencia de los establecimientos (mayor capacidad de aprovechar educativamente el *background* de los estudiantes). Este análisis no permite distinguir entre ambas interpretaciones. Recuérdese, sin embargo, que, mientras en Polonia las desigualdades de resultados han disminuido, en Chile han tendido a aumentar. Lo anterior, se suma al hecho de que el “efecto par” (ISEC de los compañeros) es mucho mayor en Chile que en Polonia, lo que da, a nuestro juicio, mayor sustento a la segunda interpretación.

de mayor retorno de las condiciones familiares en Polonia, se suma al hecho de que en ese país, los estudiantes tienen en promedio mejores condiciones socioeconómicas, de forma que ambos se potencian, aumentando la brecha con Chile.

A modo de síntesis, en primer lugar, se puede afirmar que, tomadas en conjunto, las variables referidas a los aspectos institucionales de los establecimientos, no juegan un rol relevante en la explicación de las diferencias entre los resultados de Chile y los tres países estudiados.

En segundo término, en marcado contraste con lo anterior, las variables que describen características de los establecimientos, explican gran parte de la brecha de resultados. En un patrón que es consistente en las comparaciones con los tres países, se puede estimar que, si los establecimientos chilenos tuviesen las características de estos países, los resultados promedio de los estudiantes chilenos habrían sido mayores que los observados. En particular, una parte de los mejores resultados de Polonia y España se explica por el mejor nivel socioeconómico promedio sus estudiantes y sus mayores horas de instrucción en Lenguaje. Así, en la comparación con estos dos países, la dimensión “establecimiento” es la que posee el mayor poder explicativo asociado a diferencias de características (no así, en cuanto a diferencias de retornos). Por su parte, comparado con Uruguay, una fracción de los mejores resultados chilenos, se debe al mayor tiempo dedicado a la instrucción; en cambio, los estudiantes uruguayos contrarrestan parte de su desventaja, con el hecho de tener mejores condiciones socioeconómicas y culturales que sus pares chilenos.

Finalmente, las variables que describen al estudiante, tienen un peso muy importante en la explicación de las diferencias de resultados, pero más en términos de sus retornos, que del nivel de sus características. En efecto, en las comparaciones con España y Polonia, los estudiantes chilenos tienen en promedio mejores

condiciones asociadas al aprendizaje (excepto por el nivel socioeconómico), sin embargo, estos factores tienden a ser menos productivos en términos de logros de competencia lectora. Esta diferencia de productividad es la que explica la mayor parte de la ventaja de los países europeos sobre Chile. En los años de escolaridad, este contraste es particularmente marcado: aunque en promedio, los estudiantes de España y Polonia han cursado menos grados que los chilenos (lo que los perjudica), han obtenido una mayor retribución, en términos de puntajes, por cada grado cursado. Esta diferencia les permite compensar con creces la mencionada desventaja. Esta misma variable es también la más relevante para explicar los mejores resultados chilenos, en comparación con Uruguay, aunque en este caso, ambas, las diferencias de nivel y de retornos, favorecen a Chile. Así, comparados con los uruguayos, los estudiantes chilenos han acumulado más años de escolaridad y han obtenido además mayores avances en Lectura, por cada uno de esos años. De esta forma, ambos efectos combinados son suficientes para explicar el total de la brecha observada a favor de Chile.

3. Conclusiones

3.1 Conclusiones sobre PISA Lectura

El propósito básico de este estudio fue comprender los resultados alcanzados por los estudiantes chilenos en la prueba PISA Lectura en 2006 (un análisis similar para Matemática fue reportado en el artículo anterior). Para lograrlo, se compararon los resultados de Chile, con los de dos países de la OCDE (Polonia y España) que obtuvieron puntajes más altos, y con un país de América Latina (Uruguay) que obtuvo un puntaje inferior. En particular, la investigación indagó acerca de los factores que podrían ser explicativos de las diferencias de resultados entre los estudiantes chilenos y sus pares de estos tres países. Para ello, se aplicó una variante de la metodología de descomposición propuesta por Oaxaca-Blinder. Con este procedimiento,

se pudo determinar qué proporción de la brecha observada de resultados entre Chile y cada uno de los otros tres países, estaba asociada a diferencias en las características de los estudiantes y establecimientos, y qué proporción estaba asociada a diferencias en los efectos que dichas características tienen en los logros de aprendizaje de los estudiantes.

En el estudio se analizó un conjunto amplio de variables, agrupadas en tres dimensiones: los estudiantes, los establecimientos y el nivel institucional, que, en conjunto, permitieron explicar 40% de la variación de los resultados de Chile. Asimismo, se constató que, en términos generales, Chile difería de los demás países, tanto en el nivel de estas variables, como en sus efectos estimados sobre los puntajes obtenidos por los estudiantes.

Vistos en conjunto, los hallazgos indican que la mayor parte de las diferencias de resultados entre Chile y los tres países comparados, no se explican por diferencias en los niveles de recursos o características de los establecimientos y los estudiantes, sino por diferencias en los retornos (logros de aprendizaje) que dichas características o recursos tienen. En términos más específicos, se pudo determinar que la totalidad de la brecha de resultados que separa a Chile de Polonia, se explica por una mayor eficiencia del sistema polaco y no por una mejor dotación. En el caso de España, aproximadamente un tercio de la brecha con Chile puede ser atribuida a desventajas de los estudiantes y de los establecimientos chilenos; los dos tercios restantes, se explicarían por una menor eficiencia de la educación chilena. Finalmente, en la comparación con Uruguay (país con más bajos resultados que Chile), se estima que, alrededor de un tercio de la ventaja chilena, puede ser atribuida a condiciones o características más propicias para el aprendizaje, tanto de los establecimientos, como de los estudiantes. Los dos tercios restantes, se atribuyen a una mayor eficiencia de la educación chilena, comparada con la uruguaya.

Lo anterior no significa que no haya diferencias en ciertas condiciones, que ponen en desventaja a la educación chilena respecto de los países europeos (cuyo mejoramiento podría hacer aumentar los resultados de aprendizaje de los estudiantes). Por ejemplo, los resultados sugieren que, si las familias chilenas tuviesen el nivel socioeconómico y cultural de las familias de los estudiantes españoles o polacos, el promedio nacional de Lectura habría sido 18 puntos mayor. Por otra parte, considerando que los estudiantes polacos tienen más horas de instrucción en Lenguaje que los estudiantes chilenos, se estima que aproximadamente 7 puntos de la brecha con Polonia, se explican por este factor. Lo mismo (aunque esta vez en el sentido opuesto) ocurre en la comparación con Uruguay. En efecto, alrededor de 13 puntos de la ventaja lograda por Chile, pueden ser atribuidos a que los estudiantes chilenos de 15 años han cursado más años de escolaridad que sus pares uruguayos.

En consecuencia, una fracción mayoritaria de la brecha en Lectura de Chile (incluyendo la brecha favorable en la comparación con Uruguay) es susceptible de ser explicada por las diferencias en la eficiencia de la educación entre los países analizados y Chile. Estas diferencias de eficiencia (especialmente en la comparación de Chile con Polonia y España) se explican principalmente por la escolaridad alcanzada por los estudiantes. En efecto, las diferencias en el retorno asociado a cada año de educación adicional, podrían por sí solas dar cuenta de la totalidad de la brecha de resultados observada entre Chile y estos dos países. La comparación con Polonia resulta ilustrativa en este aspecto: mientras los estudiantes chilenos obtienen una ventaja estimada de 33 puntos, producto de que han acumulado más años de escolaridad que sus pares polacos; los estudiantes polacos obtienen una ventaja estimada en 75 puntos, producto de que cada año de educación en Polonia resulta ser más productivo que en Chile.

Por último, aunque sólo en la comparación con Polonia, resulta interesante notar que una parte de la brecha de resultados se explica por las diferencias de retorno asociadas a las expectativas laborales de los estudiantes; es decir, comparativamente con Chile, los estudiantes polacos que manifiestan tener mayores expectativas laborales futuras, obtienen mejores resultados. Aunque difíciles de explicar, este tipo de hallazgos advierte que la menor “eficiencia” de Chile en comparación con estos países, se manifiesta también en los aspectos subjetivos o motivacionales de los estudiantes.

3.2 Análisis conjunto de los hallazgos en Lectura y Matemática

¿Cómo se comparan estos hallazgos de Lectura con los de Matemática? La pregunta no es sólo relevante por tratarse de las dos competencias más valoradas del currículo escolar. Es además interesante, porque los resultados obtenidos por Chile en Lectura fueron comparativamente mejores que los obtenidos en Matemática.

En primer lugar, es importante tener presente que los modelos de regresión utilizados demostraron tener un mayor poder explicativo en Matemática que en Lectura (tendencia más o menos común en este tipo de análisis, a nivel internacional). En efecto, mientras que en Matemática, el conjunto de variables da cuenta del 53% de la variación de resultados; en Lectura, solo explica el 40%. Con todo, se trata, en ambos casos, de proporciones relativamente elevadas para puntajes de pruebas estandarizadas de aprendizaje.

En segundo lugar, en Chile, todas las características de los establecimientos e institucionales que no fueron un factor relevante de los resultados en Matemática, no lo fueron tampoco en Lectura (como el tamaño del curso, el tamaño del establecimiento, el número de estudiantes por profesor, los recursos educativos y las dos medidas de autonomía de los

establecimientos). Sin embargo, algunas variables que explican los resultados de Matemática, no están asociadas sistemáticamente con diferencias en los resultados de Lectura. Esto ocurre, por ejemplo, con la valoración de los estudiantes acerca de la importancia de tener un buen resultado en Matemática y con el nivel de expectativas de los padres por el logro de altos estándares (ambas variables están asociadas positivamente con los logros en Matemática), todo lo cual apunta hacia comunidades escolares más sensibles al buen desempeño en Matemática que en Lectura.

En tercer término, salvo el género, todas las variables relevantes que son comunes para explicar los resultados de ambas asignaturas, poseen el mismo sentido y tienden a tener un efecto estimado similar. Tal es el caso de la asociación positiva con la selección académica de los estudiantes, las horas de instrucción semanal, las altas expectativas laborales y el nivel socioeconómico y cultural de las familias (aunque en este último caso, el “efecto par” parece ser más fuerte para Lectura que para Matemática⁹). Una excepción lo constituye el efecto estimado de las diferencias de grado escolar entre los estudiantes, siendo un poco mayores para Lectura que para Matemática.

Finalmente, aunque en la tendencia internacional, los hombres tienen en promedio mejores resultados en Matemática y las mujeres, en Lectura; en Chile, la brecha a favor de los hombres en Matemática es la mayor entre los países analizados y la de las mujeres en Lectura, es la menor. En otras palabras, las mujeres chilenas se encuentran en una situación menos favorable que sus pares de los demás países, en los resultados de ambas asignaturas.

En cuanto a la explicación de la brecha de resultados entre Chile y los países de referencia, los hallazgos

difieren sensiblemente entre ambas asignaturas, en las comparaciones con España y Uruguay: mientras las diferencias de características no explican nada de la diferencia de resultados en Matemática (más bien apuntan en el sentido opuesto), aproximadamente un tercio de la diferencia en Lectura (negativa con España y positiva con Uruguay) es atribuible a las diferencias de características entre Chile y estos países. En cuanto a Polonia, tanto en Lectura como en Matemática, los resultados sugieren que nada de su ventaja respecto de Chile se explica porque sus establecimientos o estudiantes posean en promedio características más favorables para el aprendizaje que los establecimientos o estudiantes chilenos.

Desagregando estas diferencias de resultados con más detalle, es posible afirmar que, de todas las características analizadas en este estudio, la diferencia en el nivel socioeconómico y cultural promedio de las familias chilenas, respecto de sus pares de los otros tres países, es el factor más relevante para explicar las brechas de resultados. Este hallazgo es común para ambas asignaturas y sistemático en la comparación con los tres países (aunque más relevante para la comparación con España y Polonia). Un hallazgo análogo, aunque de signo opuesto, se obtuvo en el análisis de la incidencia del grado escolar cursado por los estudiantes: en la comparación con los tres países (aunque especialmente relevante con Polonia), tanto en Lectura como en Matemática, los estudiantes chilenos obtienen cierta ventaja en sus resultados, atribuible a su mayor nivel de escolaridad promedio alcanzado a la edad de 15 años, siendo la magnitud de esta ventaja similar en ambas asignaturas.

Por último, el análisis de la fracción de la brecha de resultados de la prueba PISA 2006, asociada a las diferencias de retorno de las variables, mostró que la tendencia general es similar en ambas asignaturas: la mayor parte de la brecha en Lectura y la totalidad de la brecha en Matemática, pueden atribuirse a este componente. En particular, como se destacó,

⁹ Esto es consistente con la literatura sobre el tema, en la que habitualmente se señala que las competencias lectoras son más sensibles a la herencia sociocultural, que las matemáticas.

el factor más relevante para explicar el diferencial de eficiencia en contra de Chile, es la significativa menor productividad de los años de educación cursados por los estudiantes chilenos, en comparación con los polacos y, especialmente, con los españoles.

3.3 Reflexiones de política educativa

El presente estudio se encuentra en el comienzo de una línea de investigación que busca comprender mejor la relación entre procesos, recursos y resultados del sistema escolar chileno, a partir del análisis comparativo de los datos provistos por las pruebas internacionales de aprendizaje. Sorprendentemente, aunque Chile es uno de los países que más ha participado en este tipo de pruebas, el nivel de investigación empírica original en torno suyo en nuestro país, es muy limitado. En particular, no se tiene noticia de otros estudios que hayan indagado sobre factores explicativos de los resultados de Chile, mediante los análisis estadísticos de descomposición de resultados, como los presentados en este artículo y el anterior. Todo esto otorga un valor especial a estos trabajos, pero aconseja también tomar sus resultados con prudencia, en la medida en que se trata de las primeras aproximaciones a una cuestión de suyo compleja.

Adicionalmente, como se explicó en el artículo anterior, tanto los datos disponibles como la técnica de análisis empleada, poseen limitaciones que deben ser consideradas al momento de ponderar estos hallazgos. Por ejemplo, falta información sobre aspectos relevantes como los docentes, y el análisis sólo se basa en los promedios nacionales y no en la distribución total de los datos.

Con todo, es posible sugerir algunas reflexiones en torno a potenciales implicancias de los resultados de este estudio.

En su sentido más global, la investigación entrega poco soporte a aquellos que piensan que los bajos

resultados de Chile en las pruebas internacionales, se deben a una relativa carencia de recursos disponibles en nuestro sistema escolar. Ciertamente, esto no quiere decir que no sea relevante invertir en el equipamiento de los establecimientos (en un sentido amplio); todo lo contrario, quizás precisamente porque en las últimas dos décadas, el país ha hecho un esfuerzo considerable por mejorar las condiciones para la enseñanza y el aprendizaje en los establecimientos (textos, computadores, materiales didácticos, etc.), es que el presente estudio encuentra que en la situación actual, estos no son factores relevantes de desventaja para nuestros estudiantes. Más aún, el único “recurso” que se identificó asociado a las brechas de resultados entre Chile y los demás países de referencia, fue el nivel socioeconómico y cultural de las familias de los estudiantes. Sin embargo, aun siendo el más relevante, éste permite explicar sólo una parte menor de la brecha de resultados, tanto en Matemática como en Lectura. Así, aunque es esperable que los resultados promedio de aprendizaje de los estudiantes chilenos vayan aumentando, conforme el país mejore su nivel de desarrollo; parece del todo infundado confiar en que el sólo mejoramiento del contexto vaya a disminuir la brecha que actualmente separa a Chile de los países con mejor desempeño educacional. Las políticas educacionales tienen aquí un rol insustituible.

En esta materia, los hallazgos del estudio apuntan claramente hacia una menor eficiencia del sistema escolar chileno. En otras palabras, en Chile, la capacidad de la educación de transformar en logros de aprendizaje, los actuales recursos y características de los estudiantes y los establecimientos, es menor que la de los países europeos.

Ciertamente, la información disponible no indaga con la profundidad requerida sobre los procesos de gestión institucional o de enseñanza-aprendizaje, como para poder explicar estas diferencias de eficiencia. Sin embargo, este estudio permitió identificar un factor

enormemente decisivo para comprender esta menor eficiencia chilena en comparación con la polaca y la española: la baja productividad de los años adicionales de educación, alcanzados por los estudiantes de 15 años (en términos de las competencias evaluadas por PISA 2006).

No es evidente la causa de estas diferencias de “valor agregado” entre cursos. Una primera hipótesis, podría ser que en Chile, la promoción de un curso a otro se hace sin rigor ni exigencia académica, como producto de las políticas de combate a la repitencia escolar. Sin embargo, esta hipótesis es inconsistente con el hecho de que en Polonia la repitencia es mucho menor que en Chile, y también es inconsistente en la comparación con Uruguay, donde, a pesar de que la reprobación en secundaria es muy alta, las diferencias de resultados asociadas a los grados cursados, son aun menores que en Chile. Una segunda explicación tentativa, podría atribuir estas diferencias de “valor agregado”, simplemente a las diferencias estimadas de efectividad entre los docentes, pero desafortunadamente, la información recolectada por PISA, no contiene datos referidos a indicadores de calidad de los docentes, mediante los cuales se pudiera testear esta hipótesis. En este contexto, la hipótesis más plausible es la que atribuye la menor eficiencia chilena, a una menor efectividad de la práctica docente, para sacar provecho al tiempo escolar disponible. En efecto, las evaluaciones de impacto sobre la política chilena que aumentó el tiempo escolar (a través de la implementación de la jornada escolar completa), apuntan hacia un efecto positivo de esta medida, aunque de pequeña magnitud¹⁰. Una razón de esta baja efectividad docente podría radicar en el hecho de que Chile considera los grados 7° y 8° como parte de la Educación Primaria, a diferencia de los países europeos, que los consideran parte de la Educación Secundaria. Esto conlleva a que en Chile exista una relativa ausencia de docentes especializados en Lectura y Matemática, que impediría hacer avanzar a los estudiantes hacia

la adquisición de habilidades superiores en estas materias, producto de una tendencia al “repaso” de las materias más bien básicas del currículum. Una razón adicional y complementaria de este hecho, sería que los países europeos han definido con mayor claridad el objetivo de generar en los estudiantes adolescentes y jóvenes, competencias superiores en estos dos campos, mientras el currículum chileno permanece más orientado a la adquisición de conocimientos y a la mayor destreza en operaciones básicas.

En definitiva, sea cual fuere la causa de la falta de eficiencia, lo evidente es que clarificarla, es tremendamente relevante para las políticas educacionales chilenas que se propongan a futuro. Por ejemplo, si la hipótesis que aquí se ha considerado como más plausible, fuera confirmada, ello constituiría un respaldo adicional a la decisión reciente de modificar la estructura de división entre la educación primaria y secundaria en Chile, disminuyendo el primer ciclo de ocho a seis años y aumentando el segundo, de cuatro a seis.

Igualmente interesante, resulta la constatación que se realizó en este estudio, sobre la poca o nula relevancia explicativa de los factores más comúnmente considerados en los debates sobre políticas educacionales, sobre algunos de los cuales Chile ha estado apostando fuertemente en los últimos años. Tal es el caso de las diferencias en disponibilidad de recursos educativos, el grado de autonomía de los establecimientos para decidir sobre el currículum o para gestionar los recursos, el tamaño de los cursos, la dotación docente, o el nivel de presión por los resultados que los padres ponen sobre los establecimientos. Lo anterior, no implica, obviamente, que alguno de estos factores no sea eventualmente importante para una buena educación, pero sí sugiere que su modificación no es la clave para acercarnos a los países con mejores resultados en Lectura y Matemática, al menos en lo que respecta a la evaluación PISA. Esto nos recuerda que la agenda de mejoramiento de la calidad y

10 Ver Valenzuela, 2005 y Bellei, 2009.

la equidad educativas, en un marco de creciente eficiencia en el uso de los recursos disponibles, es enormemente más compleja que una de expansión escolar y equipamiento del sistema. En este sentido, quienes diseñan las políticas deberían estar cada vez más abiertos a analizar la evidencia científica, pues ésta muchas veces contradice las recomendaciones del sentido común o las demandas de los grupos de interés.

Por último, cabe destacar que, comparado con los tres países de referencia, Chile es el país en donde la selección de los estudiantes por criterios académicos en los procesos de admisión, no solo es una práctica más generalizada, sino que incide positivamente en los resultados: cuanto más selectivo el establecimiento, mejores resultados. A este respecto, cabe preguntarse ¿es que en Chile la presencia de mayor segregación académica de los estudiantes ha hecho que los establecimientos especialicen sus procesos de enseñanza-aprendizaje y sean por ello más efectivos?, ¿o se trata sólo de un efecto de redistribución de estudiantes según sus capacidades, sin efectos sobre la calidad, pero que aumenta la inequidad de la educación chilena? El debate sobre la relación entre establecimientos públicos y privados en Chile ha estado muy marcado por este punto, al que, en este caso, se agregan los fenómenos de segregación socioeconómica de la población escolar. En este sentido, cabe mencionar que los efectos sobre calidad y equidad educativas que puede tener un sistema escolar más o menos inclinado a la segregación académica y social son enormes. Como se sabe, en Chile, las condiciones institucionales (incluso recientemente ratificadas por la Ley General de Educación) dejan un amplio margen para que dichos procesos de segregación se produzcan. Aunque se trata de un asunto controversial en la literatura especializada, nuestra interpretación tiende a ser pesimista, en términos de que hay pocas ventajas y muchos riesgos en un sistema escolar así organizado.

Referencias

- Ammermüller, A. (2004). "PISA: What makes the difference? Explaining the gap in PISA test scores between Finland and Germany". *Discussion paper*, 04-04. Mannheim, Alemania: Centre for European Economic Research.
- Bellei, C. (2009). "Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile", *Economics of Education Review*, en prensa. St. Louis, Missouri: Elsevier B.V.
- Blinder, A. (1973). Wage discrimination: reduced form and structural estimates. *The Journal of Human Resources*, 4(8). Wisconsin, Estados Unidos: The University of Wisconsin Press.
- McEwan, P. y Marshall, J. (2004). "Why does academic achievement vary across countries? Evidence from Cuba and Mexico". *Education Economics*, 12 (3). Oxford, Reino Unido: Taylor & Francis.
- Oaxaca, R. (1973). "Male – Female wage differentials in urban labor markets". *International Economic Review*, 14 (3). Philadelphia: University of Pennsylvania.
- Sakellariou, C. (2008). "Peer effects and the indigenous / non indigenous early test-score gap in Peru". *Education Economics*, 16, 4. Oxford, Reino Unido: Taylor & Francis.
- Valenzuela, J. P. (2005). "Partial evaluation of a big reform in the Chilean education system: From a half day to a full day schooling". Tesis doctoral. Ann Arbor, Michigan: University of Michigan.

Características de la prueba

3



- Percepciones de los estudiantes de pedagogía sobre PISA
- Resultados de los estudiantes chilenos en la prueba PISA Ciencias 2006: una mirada a sus competencias
- Análisis comparativo de los resultados chilenos en las pruebas de Matemática SIMCE y PISA
- Análisis comparativo de la estructura interna y objetivos de evaluación de las pruebas de Lectura SIMCE y PISA

Percepciones de los estudiantes de pedagogía sobre PISA

Por
Gabriela Cares¹

Resumen

Este artículo presenta los resultados de una investigación cualitativa sobre las percepciones que tienen los estudiantes de pedagogía sobre la prueba PISA. En particular, el estudio indagó en cuáles podrían ser los usos del material de PISA para el trabajo docente en aula y qué elementos podrían facilitar o limitar estos usos. Para obtener la información, se realizaron focus groups con estudiantes de pedagogía próximos a egresar, con el propósito de llegar a conclusiones y recomendaciones de política, relacionadas con evaluación externa de aprendizajes, relevantes para el proceso de formación inicial docente.

El conocimiento demostrado por los estudiantes de pedagogía sobre la evaluación estandarizada de aprendizajes, es bastante vago. En particular, PISA es muy poco conocida, pero a medida que se va conociendo, gana aceptación entre los estudiantes, quienes ven en el material de esta evaluación una posible fuente para retroalimentar las prácticas pedagógicas. Sin embargo, se vio que hay un conjunto de limitantes para el uso de la información de PISA, en las prácticas de los docentes chilenos, entre las que destacan el desconocimiento sobre evaluación estandarizada y la falta de motivación para modificar las prácticas docentes que hoy se observan.

¹ Gabriela Cares fue miembro del Equipo de Estudios Internacionales del SIMCE.

La autora agradece a Leonor Cariola y al Equipo de Estudios Internacionales.

Introducción

La evaluación externa de aprendizajes tendrá un impacto relevante en el mejoramiento de la calidad de la educación, sólo en la medida en que sus resultados sean comprendidos y utilizados por los distintos actores que conforman el sistema educacional.

A nivel de establecimientos, la batería de instrumentos empleados por PISA y la información que entrega sobre resultados de aprendizajes, podría ser útil para retroalimentar las prácticas docentes. Actualmente, se cuenta con mucha información proveniente de las evaluaciones PISA, que podría servir a este propósito. Sin embargo, a diferencia de otros países, no existen en Chile iniciativas orientadas específicamente, a la comprensión y uso de esta información por parte de los docentes. En México por ejemplo, se ha diseñado material en el que se desglosan ítemes publicados por PISA 2000 y 2003 (INEE, 2005), describiendo qué se evalúa y cómo esto podría ser utilizado por los profesores para obtener información sobre el aprendizaje logrado por sus estudiantes, teniendo como parámetro las definiciones de PISA.

También en Italia se ha diseñado material para docentes en base a las preguntas de PISA (INVALSI, 2008), con el objetivo de difundir la información que PISA entrega y promover su uso pedagógico. Esfuerzos como estos apuntan a la difusión del material que PISA entrega, y a promover su comprensión desde una perspectiva útil para los docentes.

Para que la evaluación externa pueda ser utilizada como apoyo al proceso de enseñanza-aprendizaje, la información debe ser conocida, comprendida y percibida como útil. Bajo el supuesto de que el material de PISA puede ser un aporte para retroalimentar positivamente las prácticas docentes en Chile, se desarrolló la investigación que sustenta este artículo.

En el informe de McKinsey & Company, se ha hecho una revisión de sistemas educacionales en el mundo y se concluye de manera categórica, que “la calidad de la educación de un país no puede exceder la calidad de sus docentes” (McKinsey & Company, 2007). En sintonía con esta afirmación, la presente investigación tuvo como actores principales a futuros docentes, con el propósito de poder llegar a conclusiones que permitan, dentro de los alcances de este emprendimiento, retroalimentar su proceso de formación universitaria.

La pregunta central que guió la investigación surgió a partir de reflexiones al interior del Equipo de Estudios Internacionales del SIMCE, donde se observó que, pese a la gran cantidad de información que se genera en las evaluaciones externas de aprendizaje, se hace un limitado uso de ésta entre los distintos actores del sistema educacional que podrían aprovecharla.

Considerando que en Chile, en el último tiempo, se han intensificado los esfuerzos por el mejoramiento de la calidad de los docentes, y que la formación inicial se ha convertido en importante foco de atención, se planteó como objetivo de esta investigación describir las percepciones que tienen los estudiantes de pedagogía próximos a egresar, sobre la evaluación de aprendizajes llevada a cabo por PISA. No se trabajó directamente con profesores en ejercicio, ya que se privilegió la posibilidad de hacer propuestas enfocadas en la formación inicial docente.

Los objetivos de investigación específicos, fueron los siguientes:

1. Describir la imagen que los estudiantes de pedagogía tienen sobre PISA y explorar cómo ésta se ve afectada según la fuente de información.
2. Identificar las valoraciones (positivas o negativas) que los estudiantes de pedagogía manifiestan sobre PISA.
3. Identificar posibles usos pedagógicos de la información que PISA entrega.

4. Explorar, desde la perspectiva de los estudiantes de pedagogía, posibles facilitadores y limitantes para que el uso de la información entregada por PISA, sirva para retroalimentar las prácticas pedagógicas.

Mediante estos objetivos se buscó, en primer lugar, avanzar hacia un diagnóstico sobre el conocimiento y valoración de PISA entre los futuros docentes, para así saber cuál es el escenario que se enfrentaría si se quiere inducir el uso de la información que esta evaluación reporta. En segundo lugar, asumiendo las valoraciones que hoy se observan sobre PISA, se quiso explorar en los usos que los propios estudiantes darían al material y cuáles podrían ser los facilitadores y limitantes de estos usos.

A continuación se describe el método utilizado para la recolección de la información. Luego, se presentan los principales resultados de este estudio, divididos en los cuatro ámbitos explorados: la imagen que tienen sobre PISA los estudiantes de pedagogía; las valoraciones que expresan los mismos estudiantes, una vez que se les presenta material de PISA; los posibles usos pedagógicos que los entrevistados identifican en el material y las limitantes para el uso pedagógico de la información (aquellas identificadas por los propios estudiantes y las que se pueden deducir de sus opiniones). Finalmente, se presentan conclusiones y recomendaciones de política pública.

1. Metodología

El presente artículo se basa en los resultados de una investigación cualitativa de carácter exploratorio, sin pretensiones de representatividad.

Para la recolección de información se utilizó la técnica de focus group, la cual permite recoger opiniones sobre determinados temas incluidos en una pauta, además de indagar en opiniones sobre temas no considerados *a priori*, sino surgidos de los propios participantes. Considerando que uno de los objetivos del estudio

era conocer los posibles usos pedagógicos que los estudiantes de pedagogía identifican en el material de PISA, resultaba especialmente importante obtener las opiniones que pudieran surgir desde ellos mismos y de la interacción con sus compañeros de carrera.

La estructura general de los focus groups, fue la siguiente²:

1. Imagen de PISA.
2. Nivelación de información: entrega de conceptos claves sobre PISA.
3. Entrega de material (y tiempo para analizarlo):
 - 3.a Niveles de desempeño: resultados PISA 2006 y descripciones de cada nivel (por áreas).
 - 3.b Preguntas liberadas (por áreas).
 - 3.c Marco de evaluación (extracto por áreas).
4. Opiniones sobre el material entregado.
 - 4.a Opiniones generales.
 - 4.b Valoraciones positivas o negativas sobre pertinencia, utilidad y claridad.
 - 4.c Posibles usos pedagógicos de la información.
 - 4.d Facilitadores y limitantes para el uso de la información.
5. Resultados en PISA según variables de contexto (género y nivel socioeconómico):
 - 5.a Opiniones generales.
 - 5.b ¿Qué se puede hacer? (visión país).
 - 5.c ¿Qué están preparados para hacer ellos mismos?
6. Cierre con más información sobre PISA.

Para llegar a conclusiones sobre cada uno de los temas, se empleó un criterio de saturación de información, que consiste en indagar reiteradamente sobre lo mismo, hasta llegar a un punto en que el investigador nota que los distintos participantes de los grupos de conversación ya no revelan nuevas perspectivas.

² En el Anexo se encuentra la pauta general empleada para la realización de los focus groups.

Para el análisis de las conversaciones, éstas fueron grabadas y luego transcritas.

Para analizar el material se establecieron *a priori* las variables relevantes y sus respectivos indicadores, cada uno de los cuales recibió un código y con ellos se clasificó la información entregada por los estudiantes. Las variables y sus indicadores fueron los siguientes:

1. Imagen de PISA.
2. Percepciones sobre PISA:
 - 2.1 Preguntas de PISA:
 - a. Positivo.
 - b. Negativo.
 - c. Posible uso pedagógico.
 - 2.2 Niveles de desempeño:
 - a. Positivo.
 - b. Negativo.
 - c. Posible uso pedagógico.
 - 2.3 Marco de evaluación:
 - a. Positivo.
 - b. Negativo.
 - c. Posible uso pedagógico.
3. Posibles facilitadores y limitantes para el uso pedagógico de la información de PISA:
 - 3.1 Facilitadores y/o limitantes identificadas por los propios estudiantes.
 - 3.2 Facilitadores y/o limitantes identificadas externamente.

Ahora bien, dados los objetivos del estudio, se convocó a participar, sólo a estudiantes de pedagogía que:

- estuvieran próximos a egresar de sus carreras
- cursaran una especialidad evaluada por las pruebas PISA: Lectura, Matemática y Ciencias (aunque no fue posible conseguir estudiantes de pedagogía en Ciencias) y
- correspondieran a futuros profesores de estudiantes evaluados por PISA (de 15 años).

Considerando estas tres características, se logró convocar, a través de académicos de diversas universidades, a un total de 26 participantes³. Todos los participantes mostraron motivación y buena disposición para colaborar, lo que probablemente se relacione con la posibilidad de autoselección para participar en el estudio. En la Tabla 1 se presenta un resumen de las características de la muestra⁴.

3 En el caso de la UMCE, se retomaron en el segundo focus group, algunos temas que no fueron cerrados durante la primera sesión. Tres estudiantes participaron en ambas sesiones.

4 Para más información sobre la muestra, véase el Anexo.

Tabla 1.
Universidad y área de formación de los estudiantes de la muestra

	Universidad Bolivariana	Universidad de Santiago de Chile	Universidad de Chile	Univ. Alberto Hurtado	Univ. Metropolitana de Ciencias de la Educación
Matemática	1 sesión 4 estudiantes	2 sesiones 10 estudiantes			2 sesiones 8 estudiantes
Lenguaje			1 sesión 4 estudiantes	1 sesión 3 estudiantes	

Cada una de las sesiones de conversación fue guiada por integrantes del Equipo de Estudios Internacionales del SIMCE, y se estructuró de manera tal que los participantes tuvieran tiempo para revisar el material de PISA, antes de comentarlo. Estas sesiones se realizaron en las sedes de las universidades de pertenencia de los participantes, salvo en un caso, que se realizó en una sala de reuniones del Ministerio de Educación.

Durante las sesiones, de aproximadamente 1 hora de duración, los moderadores indujeron deliberadamente a hablar del SIMCE, con el objeto de gatillar una posible comparación con PISA. En general, este tema se planteó al principio de las conversaciones, inmediatamente después de indagar en la imagen de PISA. Luego, se entregó a cada participante diversos materiales informativos sobre PISA, para posteriormente comentarlos en grupo. El material con el que se trabajó fueron preguntas publicadas, descripciones de los niveles de desempeño, distribución de los estudiantes chilenos en los niveles de desempeño de la prueba PISA 2006, y extractos del marco de evaluación de dicho instrumento⁵.

2. Resultados

A continuación, se presentan los principales resultados obtenidos de los focus group. Para cada variable, se señala una síntesis de la información recabada y algunas citas de lo señalado por los participantes, a modo ilustrativo⁶.

2.1 Imagen de PISA

Teniendo en cuenta toda la información recogida con relación a la imagen que tienen los estudiantes de pedagogía sobre PISA, se podría decir que el perfil más recurrente es un estudiante que desconoce

PISA, pero tiende a tener una intuición positiva hacia esta evaluación. Según señalan, aunque con cierta inseguridad, PISA evalúa apropiación (y no memorización) de conocimientos, lo cual se refleja al contestar preguntas relacionadas con su aplicación en la vida diaria.

Otros participantes señalan desconocer esta evaluación, por lo que no se atreven a expresar una opinión sobre la misma. En estos casos, sólo se identifica con claridad que se trata de una evaluación internacional.

Lo hemos visto en clases, creo que hemos hablado, pero no así de fondo.

Lenguaje

Quienes demuestran tener cierto conocimiento sobre PISA, reconocen que la principal fuente de información es la propia universidad. Allí PISA es estudiada por su enfoque evaluativo —que pone en el centro la comprensión y uso del conocimiento (en contraste con la memorización)—, por el tipo de preguntas que realiza, y en menor medida, por el tipo de resultados que arroja.

Lo que nosotros sabemos de esta prueba es que en realidad lo que busca... (es) medir capacidad de razonamiento de los estudiantes, de ver cómo ellos se enfocan, cómo se enfrentan al problema, si saben leer los datos, si saben realmente lo que les están preguntando.

Matemática

He oído súper poco acerca de la prueba, pero sí hemos analizado algunas preguntas, de repente en clases.

Matemática

Un grupo reducido señala que, además de haber oído acerca de PISA en sus universidades, han leído por su cuenta sobre PISA en la prensa chilena. Ellos recuerdan que aquí se destacan los bajos resultados de Chile en esta evaluación internacional.

5 Véanse los materiales entregados a los participantes en el Anexo.

6 Los participantes que son citados, sólo se identifican con la especialidad de la carrera que se encontraban cursando.

Entre algunos estudiantes de pedagogía existe una diferenciación espontánea entre SIMCE y PISA, asociando la primera a una evaluación más cercana a “las fórmulas” o “conocimientos puros” y a PISA, a la posibilidad de usarlos.

Respecto de la validez de sus resultados, algunos señalan que, tratándose de una evaluación impulsada desde el Ministerio de Educación, les genera desconfianza. Sugieren que los datos podrían ser manipulados, principalmente en el muestreo, de modo que los resultados de Chile aparezcan más altos de lo que en realidad son.

Yo creo que la selección de escuelas se puede prestar a un índice engañoso.

Matemática

También sobre la validez de los resultados, algunos participantes señalan sus sospechas sobre la posible preparación de los estudiantes para rendir una evaluación externa, cuya fecha de aplicación es conocida de antemano por los establecimientos.

Pero los resultados de repente igual son engañosos, porque si se sabe que el 2009 va a ser la prueba y la orientación es hacia tal cosa, se trabaja con los alumnos que posiblemente van a dar la prueba. Y va a estar engañoso, no se va a medir la realidad.

Matemática

Pese a estas impresiones, al indagar si los logros que demuestran los estudiantes que han sido preparados para rendir una prueba, reflejan o no lo que realmente saben, se observa que no existe consenso entre los participantes.

2.2 Percepciones sobre PISA

Al presentarse material de PISA (preguntas, niveles de desempeño y extractos del marco de evaluación), se observa por parte de los estudiantes, una valoración predominantemente positiva hacia esta evaluación.

Respecto del marco de evaluación, lo más valorado es la relevancia que PISA asigna al efecto que tiene en los resultados, una actitud positiva de los estudiantes hacia el aprendizaje.

Esa parte que dice acá de los buenos sentimientos, yo creo que le da mayor valor... es fundamental lo que ellos sienten de repente con las matemáticas en sí, o con lo que están haciendo.

Matemática

Lo primero que llama la atención de los estudiantes de pedagogía es la forma como PISA evalúa, es decir, el tipo de preguntas más que los contenidos en sí.

Analizando las preguntas, en la mayoría de los casos, los entrevistados consideran que no se trata de contenidos demasiado difíciles o inapropiados para los alumnos. También en las preguntas, los entrevistados reconocen la pertinencia para evaluar competencias relevantes para cualquier ciudadano.

Yo creo que es una evaluación bastante buena. De hecho, yo creo que todos los niños deberían saber todos estos ejercicios de la prueba PISA, sin necesidad que estudien algo relacionado con matemáticas.

Matemática

Los participantes consideran que las preguntas se podrían responder mediante contenidos curricularmente apropiados para estudiantes chilenos de 15 años, en ocasiones teniendo que seleccionar una estrategia conocida, entre varias posibles, para llegar a la respuesta correcta. En general, los objetivos de evaluación identificados por los estudiantes de pedagogía tras las preguntas, son considerados apropiados y relevantes

para jóvenes de 15 años, tanto desde su perspectiva personal, como desde lo que ellos consideran que se valora en el currículum nacional.

Respecto del formato de las preguntas, se observan principalmente tres tipos de comentarios, relacionados con:

a. Los estímulos empleados en las preguntas.

Estos en general, son bien valorados por su creatividad y atractivo, pero también se menciona, en más de una oportunidad, que las situaciones planteadas por PISA, quizás por romper el esquema de preguntas al cual estarían más acostumbrados los estudiantes chilenos, podrían resultar más complejas.

b. La extensión de las preguntas.

Entre sus preguntas, PISA incluye algunas que demandan bastante lectura (en las distintas áreas evaluadas), lo cual se considera un obstáculo para ser contestadas por los alumnos chilenos, a quienes, a juicio de los entrevistados, no les gusta leer.

c. La integración de habilidades de distintas áreas.

Desde la perspectiva de los estudiantes de pedagogía, las preguntas de PISA demandan una comprensión más amplia que la estrictamente matemática o lectora. Esto se ve como positivo, ya que a su juicio, la forma de evaluación de PISA es integral y tiene relación con cómo se presentan los desafíos en la vida real.

Los estudiantes de pedagogía creen que la propuesta de PISA resulta muy novedosa y distinta respecto de las prácticas que hoy hay en Chile, tanto en las aulas escolares, como en las propias universidades. Si bien existe en la “academia”, una valoración teórica hacia la internalización y posible aplicación de los conocimientos y habilidades propios de cada área, las prácticas docentes no se condicen con ello, ni en los establecimientos escolares ni en la universidad.

Si hasta a nosotros nos cuesta, leemos las preguntas y (decimos) ¿qué es esto? Tenemos mucho conocimiento, mucho concepto, pero cuando hay que pensar... de analizar hay cero énfasis.

Matemática

De la información que PISA entrega, los niveles de desempeño parecieran ser lo más difícil de comprender por los estudiantes. Se observa en algunos casos, comprensión errónea en la lectura de los gráficos de los niveles de desempeño (que muestran una progresión y cómo se distribuyen los alumnos en ella). Pese a recibir una explicación previa (la misma para todos), hubo estudiantes que no comprendieron que los niveles de desempeño más bajos agrupan a los estudiantes que demuestran más débilmente las competencias evaluadas por PISA.

Sobre las expectativas de aprendizaje descritas en los niveles de desempeño de PISA, existe consenso en que las correspondientes a los niveles más altos, superan ampliamente lo que actualmente son capaces de demostrar los estudiantes chilenos de 15 años. Ya sea porque estas competencias van más allá de lo planteado en el currículum o porque sería ingenuo pensar lograrlas, dada la realidad nacional. Las expectativas de PISA, según los estudiantes de pedagogía, no parecen muy alineadas al contexto chileno⁷. Aunque las preguntas de PISA no se perciben tan difíciles, se señala que los niveles de desempeño más altos parecen poco alcanzables (esta divergencia no se aclaró durante la investigación).

Obviamente nos va a ir mal siempre, por el nivel que nosotros esperamos y exigimos de nuestros estudiantes. Este nivel (el más alto de PISA en Lectura) sería ya para los 18 años.

Lenguaje

⁷ Los niveles de desempeño describen los aprendizajes que según la OCDE, requieren ser capaces de demostrar los jóvenes a los 15 años. Cuánto se aproximan los estudiantes a estos logros, se mide a través de las preguntas de PISA.

Pese a las altas exigencias de los niveles de desempeño, los participantes creen que se trata de descripciones muy deseables de lograr y pertinentes de esperar.

Una vez que los estudiantes conocen y analizan el material de PISA y una pequeña porción del tipo de resultados que reporta, plantean a grandes rasgos tres tipos de comentarios respecto de la relevancia que tiene para Chile participar en esta evaluación:

- a. Participar en una evaluación internacional de aprendizajes es relevante para Chile y representa una inversión, ya que permite establecer comparaciones con referentes internacionales e identificar prácticas que podrían imitarse (adaptarse) en nuestro país.

Yo creo que está bien, porque uno no puede estar siempre comparándose con uno mismo, necesito saber cómo voy yo en comparación con el resto y, entre comillas, qué puedo copiarle yo al resto para avanzar. Si no, estoy todo el rato en lo mismo y no crecemos.

Matemática

- b. Está bien participar, pero la comunicación de resultados debería ser más transparente, más franca. Los estudiantes de pedagogía manifiestan que los resultados tienden a presentarse de tal manera que encubren el bajo rendimiento de los alumnos chilenos.

Decir que estamos bien en la Región... ¿por qué no decir que somos malos y la Región está mal?

Matemática

Aquí la Ministra de Educación dijo que salimos primeros...o terceros, pero son como tres los países que participan en PISA a nivel Latinoamericano⁸.

Matemática

- c. La relevancia y posibles beneficios de participar en una evaluación internacional se pierden en Chile, ya que no se toman las medidas necesarias para mejorar los magros resultados del país.

Evalúan y evalúan, pero no hay ningún cambio, siempre es igual.

Matemática

Siempre se sabe la diferencia que hay, pero nunca hay como...ya, apliquemos soluciones para que eso no siga ocurriendo o la brecha sea más chiquitita. Siempre hay puros resultados y todos los años le dan con lo mismo, pero no hay ningún tipo de estrategia para solucionar eso.

Matemática

En resumen, se puede observar que, tras haber conocido el material de PISA, los participantes tienen una valoración positiva hacia esta evaluación: a su enfoque, al tipo de preguntas que realiza y a que Chile participe, aunque esto último con ciertos reparos a la forma en que se utilizan los resultados.

2.3 Posibles usos pedagógicos de PISA

Tras revisar el material presentado en los focus groups, los futuros docentes son capaces de identificar potencial de uso pedagógico en las preguntas de PISA y en menor medida, en los niveles de desempeño.

En las preguntas publicadas, se reconoce un material de fácil acceso y uso, alineado con los énfasis que

⁸ Esta cita expresa el desacuerdo de un participante respecto del manejo público de los resultados obtenidos por Chile en PISA. Sin embargo, también refleja falta de información, ya que en PISA 2006 participaron seis países latinoamericanos.

quisieran dar los futuros docentes a sus clases, y coherente con los contenidos que se deben evaluar en 2° Año de Enseñanza Media en Chile, de acuerdo con lo establecido en el currículum.

Es increíble cómo se aplica a lo cotidiano. Un campo, una casa, un entretecho, aquí hay muchas cosas... sirven mucho para enseñar de otras formas la matemática.

Matemática

En contraste con las preguntas que se elaboran en PISA, los estudiantes critican las guías de ejercicios que hoy en día suelen emplear los docentes, las cuales son réplicas de lo que se viene haciendo desde hace muchos años y no promueven la real adquisición de competencias. En este sentido, creen que el material de PISA podría ser un buen aporte para retroalimentar la evaluación de aula, ya que está fácilmente disponible para todos y plantea problemas contextualizados, más estimulantes para los alumnos.

En cuanto a los niveles de desempeño, algunos estudiantes de pedagogía los comparan con los mapas de progreso que entrega el Ministerio de Educación, ya que ambas herramientas servirían para tener un referente de hacia dónde se quiere movilizar a los alumnos o bien qué conocimientos y habilidades se podría esperar de ellos en cada área (Ciencias, Lectura y Matemática)⁹.

Es como claramente una escala de evolución del nivel que deberían tener los estudiantes, desde un nivel más básico hasta un nivel superior. Se supone que tiene que ir adquiriendo o demostrando el manejo de determinadas competencias, como

⁹ Los niveles de desempeño de PISA describen lo que los alumnos saben y pueden hacer, en forma similar a los niveles de logro reportados por SIMCE, es decir, asociando una descripción de desempeño a un determinado rango de puntajes de la prueba. Por otra parte, los mapas de progreso son descripciones de aprendizaje, ordenadas en una secuencia progresiva "típica", que permiten a los profesores observar en qué momento del desarrollo de su aprendizaje se puede situar a un estudiante.

realizar inferencias, establecer relaciones, no sé. Nosotros comentábamos ahora que tiene mucha relación con los mapas de progreso que hay de Lenguaje que también son de Lectura, y que un poco apuntan a lo mismo.

Lenguaje

Aunque las descripciones empleadas en PISA, son consideradas un referente muy exigente para el contexto nacional, los estudiantes igualmente creen que podrían ser útiles en su trabajo como docentes. Por ejemplo, la descripción de distintos niveles permitiría identificar, contrastando mediante evaluación de aula, cuál es la distribución de los estudiantes, y con ello trabajar la diversidad.

Puedo identificar quiénes son los estudiantes que están en determinados porcentajes... de ahí empiezo a trabajar personalmente.

Lenguaje



2.4 Limitantes para usar pedagógicamente la información de PISA

Pese a los comentarios positivos de los estudiantes de pedagogía hacia los instrumentos de PISA, su marco de evaluación y los resultados que reporta, se identifican principalmente, obstaculizadores para la incorporación del material de PISA en las prácticas de los docentes chilenos. Si bien hay circunstancias que pueden considerarse como positivas, por ejemplo, la buena disposición para participar en una investigación relacionada con PISA, no se encuentran variables con potencial de intervención desde las políticas educativas.

A continuación se revisarán las principales limitantes para el uso pedagógico de la información de PISA, en primer lugar, aquellas explícitamente reconocidas por los estudiantes de pedagogía y, en seguida, aquellas limitantes que se deducen del análisis de las conversaciones.

2.4.1 Limitantes desde la propia perspectiva de los estudiantes

Las dificultades para trabajar pedagógicamente con la información de PISA, tienen que ver principalmente, con ocuparse de implementar en el aula una propuesta que, en el actual contexto del sistema de educación, sería muy innovadora. Además, demandaría de mayores esfuerzos de los que estrictamente se solicita a los profesores.

El desarrollo de clases donde se promueva el gusto por el aprendizaje y conocimientos significativos, algo valorado por los estudiantes de pedagogía en el enfoque de PISA, requeriría de una utilización de recursos para el aprendizaje distinta de la actual; de un rol más activo de los alumnos, y de mayor preparación de los docentes. A juicio de los entrevistados, las clases que hoy se desarrollan en Chile son predominantemente, orientadas al conocimiento enciclopédico, más bien abstractas y centradas en los contenidos.

Hay acuerdo entre los participantes, en que la introducción del material de PISA en el trabajo en aula, requiere esfuerzo de equipo y apoyo de los directivos para lograr un efecto real y positivo en los aprendizajes de los estudiantes. Se cree que, en la actualidad, los directivos de los establecimientos están preocupados únicamente de los resultados, en desmedro de los procesos, lo cual desalentaría a los profesores para mejorar sus prácticas.

Por otra parte, no sólo se atribuye responsabilidad a los directivos. Los profesores reticentes al cambio, representan un obstáculo para la introducción de la perspectiva y el material de PISA, en el trabajo de aula. Con relación a esas falencias, los participantes marcan una clara distinción entre las competencias con que ellos cuentan como futuros docentes y el bajo nivel que observan en las prácticas de varios de los actuales profesores.

Nosotros llegamos al colegio y somos pero más que fantásticos, conocemos cosas increíbles, y son cosas mínimas que uno debería enseñar.

Lenguaje

Ellos no quieren aprender nada, ellos tienen sus clases y las repiten todos los años, en todos los colegios, a toda hora.

Matemática

Los propios estudiantes reconocen tener una visión derrotista sobre cómo será su desempeño a lo largo de su trayectoria profesional. Hoy, ellos reconocen en sí mismos un alto nivel de motivación, consideran que su formación les permitiría desarrollar clases de mayor calidad a las que les ha tocado observar, pero asumen que deberán enfrentarse a un sistema restrictivo, con falta de recursos materiales en los establecimientos, colegas desmotivados y bajos salarios, lo que muy probablemente los llevará al desánimo.

Yo tengo el miedo de que la mayoría de nosotros vamos a caer en el sistema. Porque yo hago

práctica en un colegio y uno ve que el inspector general no está capacitado y el director menos... eso produce una desmotivación.

Matemática

Tal vez cuando tenga 45, ya no voy a saber a quién le está yendo mal ni mucho menos.

Matemática

Una crítica señalada reiteradamente por los estudiantes, es que no existe claridad sobre cuáles son los resultados de aprendizaje esperados por las autoridades, específicamente por el Ministerio de Educación. Ellos critican un sistema donde los énfasis no son claros ni se perciben como congruentes. Por otra parte, el enfoque de PISA, tal como los estudiantes lo comprenden, es altamente valorado por ellos, pero lo ven poco compatible con las orientaciones que actualmente tendría el sistema educacional chileno.

Es mucha materia, o sea estamos hablando de materia pasada, no de contenidos, no de conocimientos adquiridos, no de habilidades, entonces todavía yo creo que aquí en Chile se habla en esos términos, no se habla en términos de competencias, se está hablando en términos de materia, y yo creo que eso es lo que cuesta cambiar.

Matemática

2.4.2 Limitantes inferidas a partir del discurso de los estudiantes

a) Desconocimiento sobre el potencial de la información entregada por PISA y SIMCE

Ya se ha dicho que el conocimiento que muestran los estudiantes de pedagogía sobre PISA es muy escaso y difuso. Incluso en los casos en que se recuerda haber oído de PISA en la universidad es sólo en relación al tipo de preguntas, pero no se sabe sobre el tipo de resultados que entrega, ni hay una reflexión acerca de sus fortalezas o debilidades.

Por otra parte, los estudiantes muestran tener una opinión categórica sobre el sistema nacional de evaluación, al cual se acusa de haberse convertido en el principal referente de calidad del sistema educacional, en desmedro de los procesos que se llevan a cabo en los establecimientos.

Estamos acostumbrados a que la evaluación es lo fundamental, no estamos acostumbrados a que el proceso sea lo fundamental... en todas partes pasa lo mismo... o sea, el colegio particular quiere ganarle al otro particular, porque así van a llegar más alumnos y con eso se dedican al SIMCE.

Lenguaje

Pese a identificar una gran relevancia en el SIMCE, algunos estudiantes de pedagogía afirman que existe amplio desconocimiento acerca de la metodología de evaluación y del análisis de resultados que se puede derivar a partir de esta evaluación (entre docentes y futuros docentes). Esto lleva a desaprovechar el potencial de la información para retroalimentar su trabajo.

El sentido que yo siento que le deberíamos dar nosotros como pedagogos (al SIMCE), es ver efectivamente cuál es el nivel de ellos, y en base a eso poder tomar medidas y ver qué estás haciendo bien o qué estás haciendo mal. En la realidad la cuestión no es así y la importancia de subir el puntaje es: me va a llegar plata o no me va a llegar plata, y que le vaya bien aquí no más y no importa, después lo dejamos botado..., debería ser todo lo contrario debería haber un análisis, ya, en qué nos equivocamos, qué estamos fallando y seguir trabajando en eso, que son habilidades que uno tiene que trabajar.

Lenguaje

En términos generales se puede observar que la mayoría de los entrevistados no conoce la metodología de evaluación ni de análisis de los resultados, lo que

conduce a un escaso aprovechamiento del potencial de la información para retroalimentar su trabajo.

Un ejemplo del desconocimiento de los propósitos y forma de evaluar del SIMCE, se refleja en la utilidad asignada por un participante a poder contar con las preguntas de PISA, para “preparar el SIMCE”.

Por otra parte, respecto de la posibilidad de usar la información que el SIMCE entrega, algunos estudiantes de pedagogía criticaron a sus futuros colegas que reclaman tener poco acceso a los datos o características de las pruebas.

Comentó que él no encontraba material SIMCE, no lo encontraba tampoco en la red y, como que nos contó y a cada rato nos reíamos, porque está lleno... entonces, partiendo por eso... Yo creo que es una flojera acá, que dicen que no hay y está lleno ... en internet está lleno también de cómo trabajar la habilidad de los contenidos, tipo de preguntas, está todo.

Lenguaje

b) Desconocimiento de lo que implica una evaluación estandarizada

De las conversaciones con los estudiantes de pedagogía, se observa en reiteradas oportunidades, la percepción de la estandarización como una limitante de la validez de los resultados. Ellos consideran “injusto” esperar que jóvenes inmersos en contextos disímiles, logren el desarrollo de las mismas competencias.

No va a ser lo mismo aquí resolver problemas que en China o que en Estados Unidos. Porque a lo mejor para un niño pobre, va a ser necesario resolver problemas de la vida cotidiana con matemática, y para otra persona no va a ser lo mismo... para otra persona, que maneje todo por la tecnología y que tenga todas las cosas, (saber contar) va a ser un pelo de la cola.

Matemática

También referidos a las diferencias de contexto, pero a nivel nacional, algunos participantes consideran injustas las comparaciones hechas entre establecimientos o estudiantes que se encuentran en condiciones muy distintas para enfrentar un mismo desafío de evaluación. Algunos estudiantes de pedagogía consideran que las pruebas en sí son inequitativas.

Esa manera de estandarizar, yo creo que está bien, pero creo que es un poquito cuantitativa... es súper difícil manejar todas las variables como para decir: ya, este niño está en la misma situación que este otro niño... de repente medirlos a todos por igual, cuando las circunstancias no son las mismas, es un poco perjudicial.

Matemática

Yo veo que este tipo de pruebas debieran estar haciéndose con énfasis al contexto donde se van a aplicar. No es lo mismo aplicar este tipo de pruebas en un colegio municipal, con déficit, con problemas, donde el alumno tiene problemas serios, a un colegio particular.

Matemática

Si apenas se está luchando con que el niño pueda asistir y pueda hacer las cosas, pueda lograr estar bien en ciertas áreas, no puede rendir una prueba así.

Matemática

c) Falta de estrategias apropiadas para enfrentar desafíos intrínsecos al aula

Algunos estudiantes declaran explícitamente tener falencias en cuanto a competencias pedagógicas y otros estudiantes lo dejan de manifiesto mediante sus discursos. Por ejemplo, al pedir a los participantes imaginar un aula donde se replicara la distribución obtenida por los alumnos chilenos en los niveles de desempeño de PISA 2006, algunos señalan ideas sobre cómo reaccionarían para enfrentar la diversidad

y aprovechar la información con que se cuenta. Sin embargo, a otros estudiantes no se le ocurre qué hacer con una distribución tan disímil y algunos declaran que definitivamente, no sabrían cómo reaccionar ante un grupo de alumnos con tantas diferencias de rendimiento.

No tengo una estrategia yo, por ejemplo, para saber cómo agarro a los de más atrás y seguir subiéndolos a los que están arriba.

Matemática

Otra falencia que queda al descubierto, es la falta de reflexión sobre temas de género y su posible impacto en el rendimiento de los alumnos. En efecto, al comentar sobre las diferencias en los resultados de hombres y mujeres en PISA 2006, muchos de los participantes demuestran sorpresa y, al intentar explicar estas diferencias, muestran nulo conocimiento sobre las variables explicativas que se plantean comúnmente en la investigación educacional.

Yo también he escuchado que hacen como responsable de esos resultados a que las mujeres a esa edad están más preocupadas de gustarle al niño.

Lenguaje

Yo creo como ellas maduran primero, ya a esa edad (15 años) ya están pensando en el pololo.

Matemática

Las preguntas de PISA son más intuitivas, no es que: ya, hoy vamos a aplicar tal fórmula... entonces ahí es donde ganan los hombres, en la intuición.

Matemática

3. Conclusiones

Como se señaló al comienzo, este estudio de carácter exploratorio, busca avanzar en un diagnóstico sobre la imagen y valoraciones (positivas o negativas) que los estudiantes de pedagogía tienen sobre PISA. En segunda instancia se quiere identificar, desde la perspectiva de los futuros docentes, posibles usos pedagógicos del material de PISA y factores que podrían actuar como facilitadores o limitantes de ello.

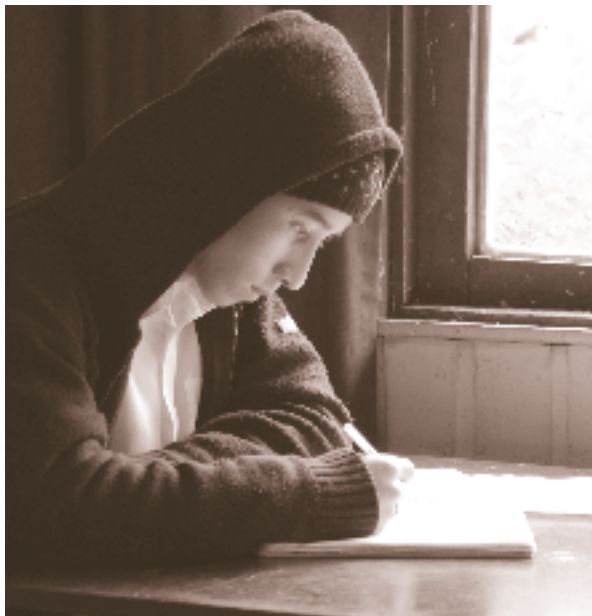
Respecto de la imagen de PISA, se observa que esta evaluación es poco conocida por los estudiantes de pedagogía, quienes, sin embargo, muestran una inclinación positiva hacia los pocos antecedentes que manejan. Una vez que los participantes conocen el material de PISA, se forman una opinión positiva de esta evaluación, destacando especialmente, el tipo de preguntas (contextualizadas en situaciones cotidianas) y la importancia que se asigna a la actitud hacia el aprendizaje.

El potencial de uso pedagógico del material que PISA provee es reconocido por los futuros docentes, pero se perciben múltiples limitantes para poder implementarlo y obtener resultados. En un comienzo, considerando la información recabada, se podría promover un mayor aprovechamiento del material publicado por PISA mediante un uso sencillo, como retroalimentar la evaluación de aula, utilizando para ello las preguntas publicadas. En este contexto, se podría aprovechar el material ya publicado en México (disponible en la web). Por su parte, las descripciones de niveles de logro, se perciben por los futuros docentes, como una herramienta que permitiría clarificar las metas de aprendizaje de sus alumnos.

En cuanto a los facilitadores para el uso de la información que reporta PISA, cabe destacar el interés que despierta el material, dado lo innovador de su enfoque. Sin embargo, es esta misma innovación, la que detona las principales limitantes encontradas. En efecto,

dados múltiples factores del sistema escolar actual, los futuros docentes perciben que entre los actuales profesores, hay cierta reticencia frente a “lo nuevo”. En este contexto, aparecen como principales elementos, la falta de motivación por actualizarse (derivada del escaso tiempo y los bajos sueldos); la carencia de elementos de juicio para abordar el contenido de los documentos (solo en escasas oportunidades y de manera parcial, se abordan en la formación docente, las particularidades de las evaluaciones estandarizadas) y la débil capacidad de comprender una información como la de PISA: compleja y multivariable.

Asimismo, aun en el contexto de que se comprendieran a cabalidad los resultados PISA, pareciera haber deficiencias en el ámbito de la formación docente, relativas a poder abordar las dificultades o diferencias en el aprendizaje demostrado por los alumnos en este tipo de evaluaciones, por ejemplo, estrategias para abordar las diferencias de género y la diversidad de niveles de logro de los distintos aprendizajes. Ello, sin duda constituye una limitante del uso de la información, pues, sin las herramientas para revertir los resultados, el mero conocimiento de los mismos, no podría sino conducir a frustración.



Considerando lo anteriormente expuesto, resulta crítico que los estudiantes de pedagogía, que prontamente serán profesores, reciban una formación que no solo les provea herramientas para abordar e interpretar adecuadamente los resultados de evaluaciones como PISA (y SIMCE), sino que especialmente, les entregue métodos de enseñanza y técnicas pedagógicas que les permitan enfrentar de manera eficiente la diversidad de niveles de aprendizaje con la que se encontrarán en las salas de clases.

Por último, queda al descubierto la necesidad de dar a conocer y tematizar formalmente, en el proceso de formación inicial docente, las metodologías de trabajo e implicancias que tiene en Chile la evaluación externa, de modo que los estudiantes sean capaces de tener una opinión informada. Para ello se tiene a favor la buena disposición de docentes universitarios y autoridades académicas, por lo que el desafío queda abierto, tanto para ellos como para el Ministerio de Educación, entre quienes se debería establecer un puente que facilite y promueva el flujo de información a este respecto.

Referencias

- INEE (Instituto Nacional para la Evaluación de la Educación) (2005). *PISA para docentes: La evaluación como oportunidad de aprendizaje*. México, D.F.
- INVALSI (Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione) (2008). *Compendio Prove PISA Insegnanti*. Italia: Centro Europeo dell'Educazione (CEDE). Extraído desde: <http://www.invalsi.it/>
- McKinsey & Company (2007). *How the World's Best Performing School Systems Come out on Top*. Extraído desde <http://www.mckinsey.com/>
- MINEDUC (Ministerio de Educación de Chile) (2008). *PISA 2006: Rendimientos de estudiantes de 15 años en Ciencias. Lectura y Matemática*. Santiago, Chile. Extraído desde: <http://www.simce.cl/>
- OCDE (Organización para la Cooperación y el Desarrollo Económico) (2006). *PISA 2006, Marco de la evaluación. Conocimientos y habilidades en Ciencias, Matemáticas y Lectura*. París, Francia. Extraído desde: <http://www.oecd.org/>

Anexo

1. Características de la muestra

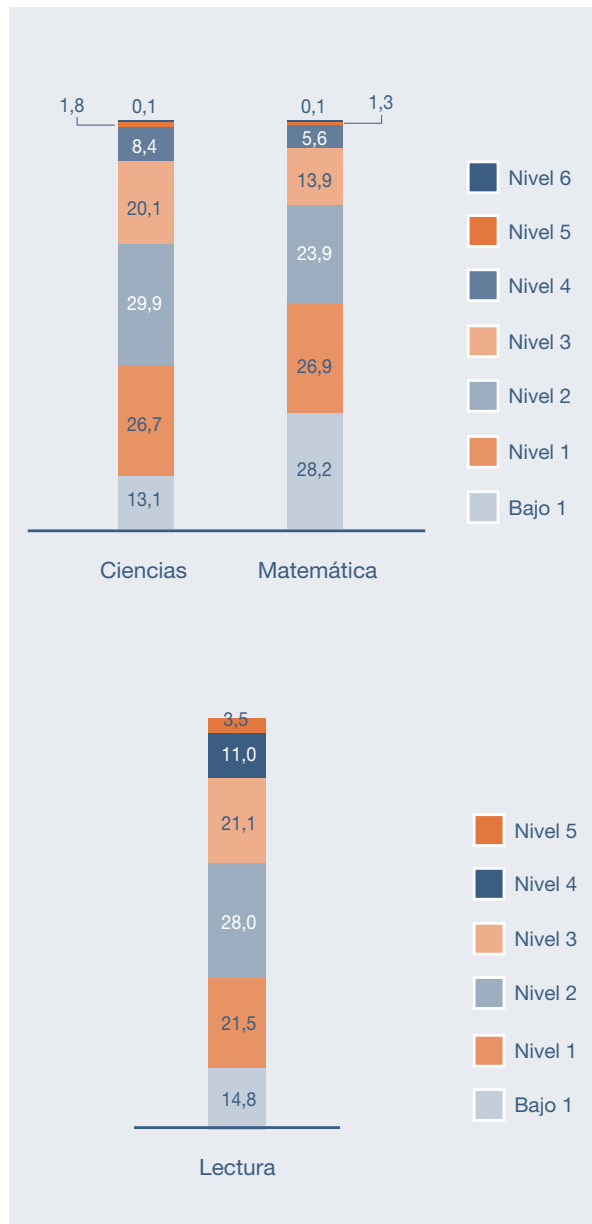
Tabla A.1
Características de la formación docente de los participantes

Carrera	Universidad	Título profesional	Grados	Duración (en semestres)
Licenciatura en Educación en Matemática y Computación	U. de Santiago de Chile (pública)	Profesor de Estado en Matemática y Computación	Licenciado en Educación Matemática y Computación	8
Pedagogía en Lengua Castellana y Comunicación	U. Alberto Hurtado (privada)	Profesor/a de Lengua Castellana y Comunicación	Bachiller en Humanidades Licenciado/a en Educación	10
Licenciatura en Lengua y Literatura Hispánicas, Mención Literatura o Lingüística ¹⁰	U. de Chile (pública)	-	Licenciado en Lengua y Literatura Hispánicas, con mención en Literatura o en Lingüística	8
Licenciatura en Educación Matemática y Pedagogía en Matemática	U. Metropolitana de Ciencias de la Educación (pública)	Profesor de Matemática	Licenciado en Educación Matemática	10
Pedagogía en Educación Básica	U. Bolivariana (privada)	Profesor/a de Educación General Básica con Mención en Matemática	Licenciado/a en Educación	8

¹⁰ Tras aprobar este programa, los estudiantes continúan sus estudios en el Centro de Estudios Pedagógicos de la Universidad de Chile. Los participantes de este estudio, ya se encontraban cursando los ramos de este segundo programa.

Figura A.1

Distribución (porcentaje) de los estudiantes chilenos en los niveles de desempeño descritos en PISA 2006



A los estudiantes se les hizo notar que Ciencias y Matemática tienen seis niveles de desempeño, mientras que Lectura sólo cinco.

2. Material entregado a los participantes en los focus groups

A los participantes se les entregó un cuadernillo con las preguntas publicadas por PISA 2000, Lectura¹¹ y otro con las preguntas publicadas por PISA 2000, Matemática¹².

Además se les entregó la distribución de los estudiantes chilenos en los niveles de desempeño en Ciencias, Matemática y Lectura.

¹¹ Se trabajó con las preguntas de PISA 2000, ya que Chile no participó en PISA 2003 y, de PISA 2006, sólo se publicaron preguntas de Ciencias. No se incluye aquí el material con las preguntas para economizar espacio. Disponible en: http://www.simce.cl/fileadmin/Documentos_y_archivos_SIMCE/evaluaciones_inter/pisa_2000/Itemes_Lectura_Liberados_PISA_2000.PDF

¹² Disponible en: http://www.simce.cl/fileadmin/Documentos_y_archivos_SIMCE/evaluaciones_inter/pisa_2000/Itemes_Matematica_liberados_PISA_2000.PDF

Figura A.2
Descripción de los niveles de desempeño en la escala de Lectura

Nivel	Puntaje en límite inferior	¿Qué pueden hacer los estudiantes?
5	625.6	<p>Los estudiantes pueden ubicar y posiblemente ordenar en secuencias o combinar múltiples fragmentos de información que no se visualiza fácilmente en el texto, parte de la cual puede estar fuera del cuerpo principal del mismo. Inferir qué información en el texto es relevante para desarrollar una tarea. Pueden discriminar satisfactoriamente entre mucha información principal que compete entre sí y construir el significado de sutilezas del lenguaje o demostrar una comprensión completa y detallada de un texto. Son capaces de evaluar críticamente y formular hipótesis, basándose en conocimiento especializado, manejar conceptos contrarios a las expectativas y lograr una profunda comprensión de textos largos o complejos. Pueden analizar textos continuos, cuya estructura discursiva no es obvia o no está claramente delineada, con el objetivo de encontrar la relación entre partes específicas del texto con su tema o propósito implícitos. En textos discontinuos, los estudiantes pueden identificar patrones entre muchos fragmentos de información presentados de manera extensa y detallada, refiriéndose algunas veces a información externa. El lector puede necesitar darse cuenta de que la total comprensión de un fragmento del texto requiere referirse a una parte separada del mismo documento, como un pie de página.</p>
4	552.9	<p>Son capaces de ubicar y posiblemente ordenar en secuencias o combinar múltiples fragmentos de información que no se visualiza fácilmente, cada uno de los cuales puede necesitar satisfacer múltiples criterios, en un texto de contenido o forma familiar. Inferir qué información en el texto es relevante para desarrollar una tarea. Usar una inferencia de alto nivel, basada en el texto, para comprender y aplicar categorías en un contexto no familiar, y para construir el significado de una sección del texto, considerando el texto como un todo. Manejar ambigüedades, ideas que son contrarias a lo esperado o formuladas negativamente. Usar conocimiento formal o general para formular hipótesis o evaluar críticamente un texto. Mostrar una comprensión precisa de textos complejos o largos. En textos continuos, establecer relaciones lingüísticas o temáticas entre varios párrafos, frecuentemente en ausencia de marcas discursivas claras, con el objetivo de localizar, interpretar o evaluar información no destacada, o para inferir sentido psicológico o metafísico. En textos discontinuos, explorar un texto largo y detallado con el propósito de encontrar información relevante, frecuentemente con poca o ninguna ayuda de organizadores como rótulos o formatos especiales, para localizar diversos fragmentos de información que deban ser comparados o combinados.</p>

(Continuación en página siguiente)

(Continuación Figura A.2)

Nivel	Puntaje en límite inferior	¿Qué pueden hacer los estudiantes?
3	480.2	<p>Los estudiantes pueden ubicar y en ocasiones reconocer la relación entre fragmentos de información, cada uno de los cuales puede requerir satisfacer múltiples criterios. Discriminar satisfactoriamente entre información relevante que compite entre sí. Integrar varias partes de un texto para identificar la idea principal, comprender una relación o construir el significado de una palabra u oración. Comparar, contrastar o categorizar tomando en cuenta diversos criterios. Son capaces de establecer conexiones o comparaciones, dar explicaciones o evaluar una característica de un texto; también, demostrar una comprensión detallada del texto en relación con conocimiento familiar, cotidiano o basándose en conocimiento menos común. En textos continuos, pueden usar convenciones sobre la organización de textos, cuando están presentes, y establecer relaciones lógicas —implícitas o explícitas— como relaciones causa-efecto entre oraciones o párrafos, para localizar, interpretar o evaluar información. En textos discontinuos, considerar una forma de presentación a la luz de una segunda, separar documentos o formas de presentación, posiblemente en un formato distinto, o combinar diversos fragmentos de información espacial, verbal y numérica en un gráfico o mapa para extraer conclusiones acerca de la información representada.</p>
2	407.5	<p>Los estudiantes son capaces de ubicar uno o más fragmentos de información, cada uno de los cuales puede requerir satisfacer múltiples criterios. Discriminar satisfactoriamente entre información que compite entre sí. Identificar la idea principal en un texto, comprender relaciones, establecer o aplicar categorías simples, o construir significado dentro de una parte limitada del texto, cuando la información no está destacada y se requieren inferencias simples. Hacer comparaciones o conexiones entre el texto y conocimiento extratextual, o explicar una característica del texto basándose en experiencias y actitudes personales. En textos continuos, establecer relaciones lógicas y lingüísticas dentro de un párrafo, para localizar o interpretar información; o sintetizar información entre textos, o de partes de un texto, para inferir el propósito del autor. En textos discontinuos, demostrar una noción de la estructura subyacente a una presentación visual, por ejemplo, un diagrama de árbol o un cuadro, o combinar dos fragmentos de información de un gráfico o un cuadro.</p>
1	334.8	<p>Pueden ubicar uno o más fragmentos independientes de información explícita que comúnmente requieren satisfacer un solo criterio, con poca o ninguna información que compite entre sí en un texto. Reconocer el tema principal o el propósito del autor en un texto sobre un tema que les resulta familiar, cuando la información en el texto está destacada. Establecer una relación simple entre la información del texto y el conocimiento de la vida cotidiana. En textos continuos, pueden usar redundancia, encabezados de párrafos o convenciones comunes de formato, para formarse una impresión de la idea principal del texto, o para localizar información explícita dentro de una sección breve del mismo. En textos discontinuos, son capaces de enfocar fragmentos discretos de información, usualmente en una sola forma de presentación, por ejemplo, un mapa simple, un gráfico de líneas o uno de barras con sólo una pequeña cantidad de información de manera directa, y en el que la mayor parte del texto verbal está limitado a una pequeña cantidad de palabras o frases.</p>

Fuente: MINEDUC, 2008

Figura A.3
Descripción de los niveles de desempeño en la escala de Matemática

Nivel	Puntaje en límite inferior	¿Qué pueden hacer los estudiantes?
6	669.3	Los estudiantes conceptualizan, generalizan y usan información basada en sus investigaciones y en el modelamiento de situaciones problemáticas complejas. Pueden relacionar diferentes fuentes de información y representaciones y hacer traducciones entre ellas de manera flexible. Logran un razonamiento y pensamiento matemático avanzado, y pueden aplicarlo, junto con el dominio de las operaciones y relaciones matemáticas simbólicas y formales, en el desarrollo de nuevas aproximaciones y estrategias para enfrentar situaciones novedosas. Pueden formular y comunicar con precisión sus acciones y reflexiones relacionadas con sus descubrimientos, interpretaciones, argumentos, y la adecuación de éstas respecto de las situaciones originales.
5	607.0	Pueden desarrollar y trabajar con modelos para situaciones complejas, identificando sus limitaciones y especificando sus supuestos; seleccionar, comparar y evaluar estrategias de resolución de problemas adecuadas para abordar problemas complejos referidos a estos modelos. Pueden trabajar de manera estratégica, usando habilidades de pensamiento y razonamiento amplias y bien desarrolladas, representaciones vinculadas adecuadamente, caracterizaciones simbólicas y formales, y comprensión profunda de estas situaciones. Son capaces de reflexionar sobre sus acciones y formular y comunicar sus interpretaciones y razonamiento.
4	544.7	Los estudiantes pueden trabajar de manera eficiente con modelos explícitos de situaciones concretas complejas que involucren condicionantes o la necesidad de reconocer supuestos. Pueden seleccionar e integrar diferentes representaciones, incluyendo las simbólicas, relacionándolas directamente con situaciones del mundo real. Son capaces de usar habilidades bien desarrolladas y razonar con flexibilidad y cierta profundización en estos contextos. Ellos pueden elaborar y comunicar explicaciones y razonamientos, basándose en sus propias interpretaciones, argumentos y acciones. Los estudiantes pueden trabajar de manera eficiente con modelos explícitos de situaciones concretas complejas que involucren condicionantes o la necesidad de reconocer supuestos. Pueden seleccionar e integrar diferentes representaciones, incluyendo las simbólicas, relacionándolas directamente con situaciones del mundo real. Son capaces de usar habilidades bien desarrolladas y razonar con flexibilidad y cierta profundización en estos contextos. Ellos pueden elaborar y comunicar explicaciones y razonamientos, basándose en sus propias interpretaciones, argumentos y acciones.
3	482.4	Pueden ejecutar procedimientos claramente descritos, incluyendo los que requieren decisiones secuenciales. Ellos pueden seleccionar y aplicar estrategias simples de resolución de problemas, interpretar y usar representaciones basadas en diferentes fuentes de información, y razonar directamente a partir de ellas. Son capaces de elaborar comunicaciones breves para reportar sus interpretaciones, resultados y razonamiento.
2	420.1	Son aptos para interpretar y reconocer situaciones en contextos que requieren sólo inferencia directa. Ellos pueden extraer información relevante de una fuente de información a la vez y hacer uso de una manera de representación. Pueden utilizar algoritmos, fórmulas, procedimientos o convenciones básicas. Son capaces de razonar directamente y hacer interpretaciones literales de los resultados.
1	357.8	Los estudiantes pueden responder preguntas que involucren contextos familiares, en los cuales toda la información relevante está presente y las preguntas se encuentran claramente definidas. Son capaces de identificar información y llevar a cabo procedimientos rutinarios, siguiendo instrucciones directas en situaciones explícitas. Pueden realizar acciones que son obvias y se desprenden directamente de los estímulos presentados.

Fuente: MINEDUC, 2008

2.1 Extracto del marco de evaluación de PISA 2006 Lectura

Definición del área de evaluación

Las definiciones de la lectura y de la *competencia lectora* han evolucionado paralelamente a los cambios sociales, económicos y culturales. El concepto de aprendizaje, y en especial el de aprendizaje a lo largo de toda la vida, han transformado las percepciones de la *competencia lectora* y de las necesidades a las que ha de hacer frente. Ya no se considera que la capacidad de lectura sea algo que se adquiere en la infancia durante los primeros años de escolarización. Más bien se ve como un conjunto en evolución que incluye una serie de conocimientos, habilidades y estrategias que las personas van construyendo con los años, según las diversas situaciones que viven y mediante la interacción con sus compañeros y con las comunidades más extensas en las que participan.

Como resultado de un proceso de consenso en el que participaron los expertos en lectura designados por los países participantes y los grupos de asesoramiento de PISA, se adoptó para el estudio la siguiente definición de *competencia lectora*:

La *competencia lectora* es la capacidad de comprender, utilizar y analizar textos escritos para alcanzar los objetivos del lector, desarrollar sus conocimientos y posibilidades y participar en la sociedad.

Esa definición supera la idea tradicional de *competencia lectora* como proceso de descodificación y comprensión literal. En lugar de ello, parte de la base de que la *competencia lectora* comporta comprender informaciones escritas, utilizarlas y reflexionar sobre ellas para cumplir una gran variedad de fines. La definición tiene, pues, en cuenta el papel activo e interactivo del lector que adquiere información a partir de textos escritos. La definición también está

abierta a la enorme variedad de situaciones en las que la *competencia lectora* puede desempeñar un papel para los adultos jóvenes, situaciones que van desde lo público a lo privado, desde el entorno escolar al laboral, desde el ejercicio activo de la ciudadanía hasta el aprendizaje continuo. Asimismo, hace explícita la idea de que la capacidad de lectura permite al individuo dar satisfacción a una serie de aspiraciones personales, que abarcan desde la consecución de metas específicas, como la cualificación educativa o el éxito profesional, hasta objetivos menos inmediatos destinados a enriquecer y mejorar la vida personal. La *competencia lectora* también proporciona a las personas unos instrumentos lingüísticos que resultan cada vez más necesarios para poder hacer frente a las exigencias de las sociedades modernas, con su extenso aparato burocrático, sus instituciones formales y sus complejos sistemas legales.

Mientras tratan de comprender y utilizar aquello que están leyendo, los lectores reaccionan ante un texto determinado de muy distintas maneras. Ese proceso dinámico incluye muchos factores, algunos de los cuales pueden ponerse en juego en un estudio a gran escala como PISA. Tres que se pueden mencionar son la situación de la lectura, la estructura del propio texto y las características de las preguntas que se suscitan sobre el texto (la rúbrica del texto). Todos estos factores constituyen elementos importantes del proceso de lectura y fueron tenidos en cuenta a la hora de elaborar los ejercicios para la evaluación.

2.2 Extracto del marco de evaluación de PISA 2006 Matemática

Definición del área de evaluación

El área de la *competencia matemática* definido por PISA hace referencia a la capacidad de los alumnos para analizar, razonar y comunicarse eficazmente cuando plantean, formulan, resuelven e interpretan problemas matemáticos en diversas situaciones. En

lugar de limitarse al tipo de situaciones y problemas que suelen encontrarse en las aulas, la evaluación PISA se centra en los problemas del mundo real. En un entorno real, los ciudadanos han de hacer frente a una serie de situaciones al ir de compras, viajar, cocinar, ocuparse de su economía doméstica, valorar cuestiones políticas, etc., en las que el empleo de un razonamiento cuantitativo o espacial, u otras capacidades matemáticas, contribuirá a aclarar, formular o resolver los problemas que se les planteen. Estos usos de las matemáticas se basan en las habilidades que se han aprendido y practicado mediante el tipo de problemas que suelen presentarse en los libros de texto y en las aulas. Sin embargo, exigen asimismo la capacidad de aplicar esas habilidades a unos contextos menos estructurados, que carecen de instrucciones precisas y en los que el alumno debe decidir cuál será el conocimiento más adecuado al caso y cuál será la forma más útil de aplicarlo.

El concepto de *competencia matemática* de PISA pretende evaluar en qué medida los alumnos de 15 años pueden ser considerados unos ciudadanos reflexivos e informados y unos consumidores inteligentes. Cada vez es más normal que los ciudadanos de cualquier país se vean enfrentados a una multiplicidad de tareas que entrañan conceptos matemáticos de carácter cuantitativo, espacial, probabilístico o de algún otro tipo. Los medios de comunicación (periódicos, revistas, televisión e internet) están repletos de información en forma de tablas, diagramas o gráficos, donde se tratan temas como el clima, la economía, la medicina o los deportes, por mencionar tan solo unos pocos ejemplos. Los ciudadanos se ven sometidos a un bombardeo informativo sobre temas como «el calentamiento global y el efecto invernadero», «el crecimiento demográfico», «los vertidos de petróleo en los mares», «la desaparición de espacios naturales». Por último, sin ser por ello menos importante, los ciudadanos se ven en la necesidad de leer formularios, interpretar horarios de trenes y autobuses, llevar a cabo transacciones monetarias de forma satisfactoria, decidir cuál es la mejor compra en el mercado, etc.

La *competencia matemática* de PISA se centra en la capacidad de los alumnos de 15 años (una edad en la que muchos de ellos están a punto de completar el ciclo de formación obligatoria en matemáticas) para dotar de sentido estas cuestiones y llevar a cabo las tareas que requieren, recurriendo a sus conocimientos y su comprensión de las matemáticas.

PISA define así la *competencia matemática*:

Competencia matemática es una capacidad del individuo para identificar y entender la función que desempeñan las matemáticas en el mundo, emitir juicios fundados y utilizar y relacionarse con las matemáticas de forma que se puedan satisfacer las necesidades de la vida de los individuos como ciudadanos constructivos, comprometidos y reflexivos.

3. Pauta Focus Groups¹³

I. Conocimiento e imagen sobre PISA

¿Han oído hablar de PISA? ¿Qué saben de ella?

¿Conocen el SIMCE? ¿Podrían comparar SIMCE y PISA?

- Evaluación estandarizada (*¿Conocen? ¿Qué opinan?*)
- Evaluación internacional (*¿Conocen? ¿Qué opinan?*)
- Evaluación de tendencias (*¿Conocen? ¿Qué opinan?*)

II. Presentación de PISA

Evaluación internacional que permite obtener puntajes sobre el desempeño de los estudiantes de 15 años, en Ciencias, Matemática y Lectura. Permite

¹³ Esta fue la pauta general empleada en la realización de los focus groups, cuyo orden en la práctica, varió de acuerdo a las opiniones que se fueron recogiendo. Además, se incorporaron en algunas oportunidades, preguntas que fueron surgiendo a partir de las opiniones expuestas por los participantes.

obtener puntajes para cada uno de los países y compararlos entre ellos. Además de las pruebas sobre rendimiento académico, se aplican cuestionarios que recogen información de contexto de los estudiantes, de sus familias y de sus establecimientos. Esto permite relacionar el rendimiento académico con otras variables, como nivel socioeconómico de los estudiantes, uso de computadores o características de los establecimientos.

Ahora voy a entregarles un material sobre el cual me gustaría que opinen.

Entrega de material:

Información sobre niveles de desempeño:

Los niveles de desempeño se asocian a tramos de puntajes en la escala de PISA. Para cada nivel de desempeño, se describen las competencias demostradas por los estudiantes al responder las preguntas de PISA. Los niveles más altos describen las competencias que son capaces de demostrar aquellos estudiantes que logran responder correctamente las preguntas más difíciles de la prueba.

Información sobre preguntas: entrega de dossier con preguntas liberadas de PISA en Matemática y Lectura. Se solicita revisar y resolver un par de preguntas a cada estudiante y se da algo de tiempo para hojear las demás preguntas.

Información sobre marco de evaluación: se entrega a cada estudiante el fragmento con el marco de evaluación.

Les voy a entregar un material para que después conversemos sobre eso. Es una breve referencia sobre cuál es el marco de evaluación en el que PISA evalúa Matemática y Lectura. Porque hay una definición de qué es lo que PISA quiere evaluar, qué es lo que se

entiende por competencia Matemática y Lectora. Se los voy a pasar para que lo revisemos.

III. Percepción sobre las preguntas, los niveles de desempeño y el marco de evaluación de PISA

Objetivo de evaluación de la pregunta

- ¿Qué creen que evalúa esta pregunta?
- ¿Qué tendrían que saber los niños para contestar bien?

Dificultad de la pregunta

- ¿Cuán difícil les parece la pregunta que revisaron?

Pertinencia de las preguntas

- ¿Cuán adecuadas les parecen estas preguntas para estudiantes chilenos de 15 años? (considerando el nivel de conocimiento que pueden tener, considerando las características socioculturales y geográficas de Chile; para hombres y mujeres; jóvenes de distintos grupos socioeconómicos).
- ¿Qué les parece el formato de las preguntas?
- ¿Qué les parece la pauta de corrección?

Relevancia de las preguntas

¿Cuán relevantes les parecen estas preguntas?:

- De acuerdo a los conocimientos y habilidades que el curriculum prescribe.
- De acuerdo a las competencias que un estudiante de 15 años debería haber logrado.

Pertinencia y utilidad de los niveles de desempeño

- ¿Qué les parecieron los niveles de desempeño? ¿Comprensibles? ¿Útiles? ¿Apropiados para estudiantes de chilenos de 15 años?

Marco de evaluación

- ¿Qué les pareció? ¿Se entiende? ¿Qué les llama la atención?

IV. Percepción sobre las propias competencias para trabajar con las preguntas de PISA

- *¿Han oído hablar sobre PISA en su formación como docentes? (algo sobre evaluación estandarizada, internacional (muestreo), tendencias).*
- *¿Cuán preparados se sienten para enseñar a los estudiantes a contestar preguntas como las de PISA?*
- *¿Creen que el material de PISA, la información que se genera con esta prueba, podría utilizarse para mejorar sus prácticas como docentes de aula?
¿Cómo?*
- *Si les han informado de PISA en su carrera, ¿desde qué perspectiva se abordó?*
- *¿Creen que una prueba como PISA es un reflejo válido de las competencias que los estudiantes han desarrollado?*

¿Qué nos dice PISA sobre la educación de los jóvenes en Chile?

Resultados de los estudiantes chilenos en la prueba PISA Ciencias 2006: una mirada a sus competencias

Por
Johanna Gubler
Alexis Williamson¹

Resumen

En la prueba PISA Ciencias 2006 los estudiantes chilenos tuvieron un rendimiento significativamente más bajo que el promedio de países de la OCDE. El presente estudio busca identificar qué aspectos de los evaluados por dicha prueba constituyen debilidades o fortalezas en el dominio científico de los estudiantes chilenos respecto de sus pares de la OCDE. Para esto, primero se indagó si había diferencias significativas en el porcentaje de respuestas correctas (PRC) promedio de cada una de las preguntas, entre los estudiantes chilenos y los de los países OCDE, y luego, se realizó un análisis de conglomerados para cada subescala, que permitió, en base al PRC promedio, agrupar las preguntas de la prueba en tres grupos de dificultad (baja, media y alta). Este mismo análisis se hizo diferenciando por género. En una segunda etapa, se analizaron las preguntas de cada grupo de dificultad, considerando las variables formato de la pregunta, contexto de aplicación, área de aplicación, área de contenido y habilidad evaluada. Se observó que en Chile, al comparar con la OCDE, la distribución de las preguntas tiende a desplazarse de manera relativamente uniforme hacia los grupos de dificultad superior. Del análisis de las preguntas, se pudo establecer que, para ambos grupos de estudiantes, las preguntas de formato abierto y las preguntas planteadas en contexto global tienden a resultar más difíciles, y que para los estudiantes chilenos resultan particularmente difíciles las habilidades de interpretar pruebas científicas y formular conclusiones y de reconocer cuestiones susceptibles de ser investigadas científicamente. Respecto de los resultados según género, la competencia de explicar fenómenos científicamente resultó ser la que presenta mayores diferencias entre hombres y mujeres en Chile, a favor de los hombres y, al comparar los resultados de Chile con el grupo OCDE, las estudiantes chilenas se distancian en mayor medida de sus pares de la OCDE que los estudiantes hombres.

¹ Johanna Gubler es Coordinadora Nacional del Estudio Internacional de Tendencias en Matemáticas y Ciencias TIMSS, de SIMCE. Alexis Williamson es Coordinador del Equipo Disciplinario de Ciencias Naturales de SIMCE.

Los autores agradecen al comité técnico de esta publicación, por su valiosa retroalimentación a este artículo y, especialmente, a Claudia Matus por su disposición y apoyo en todo lo referido al análisis estadístico.

Introducción

Los resultados de la prueba PISA de Ciencias aplicada el año 2006 (a la que, en adelante, referiremos como “PISA Ciencias”) dan cuenta de que los estudiantes chilenos presentan falencias respecto de su competencia científica. En el presente estudio se examinaron estos resultados con el objetivo de determinar en qué aspectos de los evaluados, se observan las principales debilidades o fortalezas de los estudiantes chilenos, en relación al dominio que muestra el promedio de estudiantes de los países pertenecientes a la OCDE. Esto se inserta en el interés de utilizar de manera exhaustiva la información que entregan las pruebas internacionales, con el objeto de proveer al sistema una perspectiva detallada de los resultados en áreas específicas de aprendizaje, aportando de este modo a la tarea de fortalecer el desarrollo de las competencias de los estudiantes, mediante la entrega de información útil, tanto para la enseñanza en aula, como para la toma de decisiones de política pública.

Para dar cumplimiento al propósito de este estudio, se realizó una comparación de los resultados de los estudiantes de Chile y de la OCDE en la escala general de ciencias y en las subescalas correspondientes a las competencias evaluadas por la prueba. Esta comparación permitió determinar las particularidades de los estudiantes de Chile en relación con las de los estudiantes de países que, en su mayoría, han alcanzado altos niveles de desarrollo económico y social ².

A continuación, se presentan los principales antecedentes sobre la prueba PISA Ciencias, que sirven de base a

esta investigación, luego se explicitan los objetivos del estudio y la metodología que se ha seguido y, finalmente, se exponen sus principales resultados y conclusiones.

1. Antecedentes de la Prueba PISA Ciencias

La prueba PISA adopta un enfoque amplio para evaluar el conocimiento y las destrezas que reflejan los cambios actuales de los currículos, superando el enfoque basado en la escuela, es decir, teniendo en cuenta la utilización del conocimiento, en las tareas y desafíos de cada día, en situaciones nuevas para los estudiantes (OCDE, 2006).

El proyecto PISA, se basa en un modelo dinámico de aprendizaje y se centra en aquello que los estudiantes de 15 años necesitarán en el futuro, por lo que pretende valorar lo que son capaces de hacer con lo que han aprendido. Así, al tiempo que evalúa los conocimientos de los estudiantes, PISA también evalúa su capacidad para aplicar su conocimiento y experiencia a diversos asuntos del mundo real. Se utiliza el término “competencia” para expresar esta concepción amplia de los conocimientos y destrezas.

El énfasis en evaluar el dominio funcional (grado de aptitud adquirido) de conceptos amplios, cobra especial significación, a la luz del interés que muestran las naciones por desarrollar su capital humano. Este es definido por la OCDE como: “Los conocimientos, destrezas y otros atributos de los individuos que influyen de manera significativa en el bienestar personal, social y económico” (OCDE, 2006).

En este contexto, la prueba PISA Ciencias, entiende la competencia científica, como la capacidad para usar los conocimientos y los procesos científicos no solo para comprender el mundo natural, sino también para tomar decisiones informadas sobre el ambiente natural y social, y sobre la propia salud y bienestar.

² En el año 2007 Chile fue invitado a formar parte de este conjunto de países y se espera que para fines del 2009, sea miembro pleno de la OCDE. Ser parte de esta organización proveerá de un marco para identificar, comparar, coordinar y evaluar políticas nacionales e internacionales en educación. Esto representa parte de los desafíos que nuestro país se ha propuesto en cuanto a desarrollo económico y social, e implica aspirar al logro de un alto nivel educacional para toda la población. De esta manera, los niveles educativos alcanzados por los países OCDE constituyen tanto un punto de referencia como un importante desafío para nuestro país.

1.1 Competencias evaluadas

La prueba PISA Ciencias, evaluó el grado de dominio de tres competencias científicas:

- *Explicar fenómenos científicamente.*
- *Identificar cuestiones científicas.*
- *Usar evidencia científica.*

A continuación se presentan las definiciones de cada una de estas competencias, contenidas en el Marco de Evaluación de dicha prueba (OCDE, 2006).

La competencia de *explicar fenómenos científicamente*, supone aplicar el conocimiento adecuado de la ciencia a una determinada situación. Esta capacidad implica describir, analizar e interpretar fenómenos, predecir cambios, y reconocer las descripciones, explicaciones y predicciones apropiadas a cada caso, aplicando conocimientos y comprendiendo los principales conceptos y principios científicos.

La competencia de *identificar cuestiones científicas*, por su parte, supone distinguir temas y contenidos científicos de otros tipos de asuntos. El aspecto más importante es que los asuntos científicos deben poder resolverse mediante respuestas basadas en pruebas de carácter científico. Esta competencia implica reconocer interrogantes que pueden ser investigadas científicamente e identificar términos clave para buscar información científica sobre un determinado tema. Además, incluye la capacidad de reconocer los rasgos característicos de una investigación de corte científico, por ejemplo, qué elementos deben ser comparados, qué variables deben modificarse o someterse a control, qué información complementaria se requiere o qué medidas han de adoptarse para recoger los datos apropiados al caso.

Finalmente, la competencia de *usar evidencia científica* supone que los estudiantes sean capaces de captar el sentido de los hallazgos científicos, con el fin de usarlos

como pruebas para realizar afirmaciones o conclusiones. Implica la capacidad de acceder a información científica, así como la de elaborar argumentos y conclusiones basados en pruebas científicas. Esta competencia puede englobar aspectos tales como: seleccionar conclusiones alternativas, en función de las pruebas de que se dispone; dar razones a favor y en contra de una determinada conclusión, según los procesos empleados para llegar a ella, e identificar los supuestos que se han asumido para llegar a una conclusión. La reflexión sobre las implicancias sociales de los avances científicos o tecnológicos constituye otro aspecto de esta competencia. Para dar cuenta de esta competencia, a los estudiantes se les puede pedir que comuniquen sus pruebas o evidencias y decisiones ante un público determinado, bien con sus propias palabras, bien mediante el uso de diagramas u otros sistemas de representación apropiados. En suma, los estudiantes deben poder presentar en forma lógica y clara las conexiones entre las pruebas científicas y sus conclusiones o decisiones.

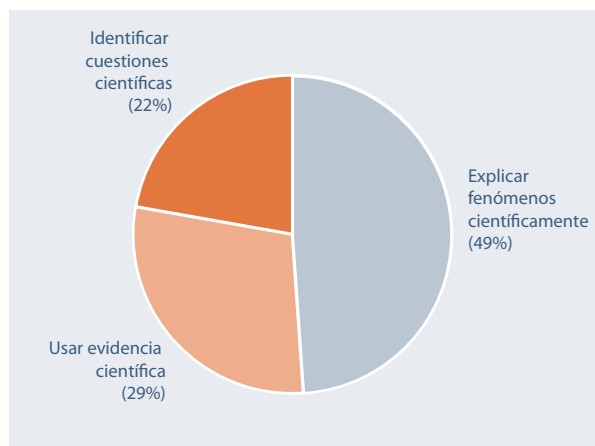
Estas tres competencias constituyen las subescalas de la prueba PISA Ciencias 2006 y, de acuerdo con lo que se señala en el Marco de Evaluación de la prueba (OCDE, 2006), fueron seleccionadas por su importancia en la práctica de las ciencias y por su relación con habilidades cognitivas clave, tales como, el razonamiento inductivo/deductivo, el pensamiento basado en sistemas y modelos, la toma de decisiones con sentido crítico, la transformación de la información, la elaboración y comunicación de argumentos y explicaciones basadas en datos, y la utilización de las ciencias. En la sección 4 de este artículo, se presentan ejemplos de preguntas pertenecientes a cada una de estas subescalas.

1.2 Características de las preguntas

La prueba PISA Ciencias se estructuró en grupos de preguntas asociadas a un mismo estímulo o tema. En total, la prueba incluyó 108 preguntas, de las cuales 33% son de *respuestas abiertas* (36 preguntas), es decir, preguntas en las que el estudiante debía construir su propia respuesta (breve o extensa, dependiendo de la pregunta) y que en este artículo llamaremos *preguntas abiertas*. El 67% restante es de selección múltiple (72 preguntas), es decir, preguntas en las que el alumno debe elegir una respuesta dada, y que en este artículo llamaremos *preguntas cerradas*. En el presente estudio se consideraron 103 de las 108 preguntas totales dado que internacionalmente 5 preguntas presentaron problemas métricos. En la sección 4 de este artículo, se presentan ejemplos de preguntas de distintos formatos.

La distribución de las preguntas en términos de las competencias evaluadas, muestra que aproximadamente la mitad de la prueba se refirió a la competencia de *explicar fenómenos científicamente*. En el Gráfico 1 se muestra la distribución porcentual de preguntas en la prueba, en cada una de las competencias evaluadas.

Gráfico 1: Distribución de las preguntas de PISA Ciencias, según competencia evaluada



Fuente: OCDE, 2006

Cada una de las preguntas de la prueba se elaboró para evaluar el desempeño en al menos una de estas competencias y sobre la base de variables comunes y de variables propias de algunas competencias. Entre las variables comunes, se cuentan el *formato* de la pregunta, el *contexto* en que se sitúa y el *área de aplicación*. Entre las variables propias de algunas competencias, se distinguen, el *área de contenido*³ que aborda la pregunta (en la competencia *explicar fenómenos científicamente*) y la *habilidad* que evalúa la pregunta (en las competencias *identificar cuestiones científicas* y *usar evidencia científica*). En la Tabla 1, se muestra cada una de estas variables con sus respectivas categorías.

En este estudio, a partir de la base de datos de PISA Ciencias 2006, se pudo clasificar cada una de las preguntas según la competencia que evalúa y las categorías que adopta para las variables señaladas. En la sección 4 de este artículo, se presentan ejemplos de preguntas que ilustran estas distintas categorías.



³ La Prueba PISA 2006 considera como áreas de contenido “conocimiento de la ciencia” y “conocimiento sobre la ciencia”, el primero fue el que se consideró en el presente estudio.

Tabla 1.
Variables y sus categorías para cada subescala o competencia

Subescala	Variables	Categorías
Explicar fenómenos científicamente	Formato	Abierta - Cerrada
	Contexto	Personal - Social - Global
	Área de contenido	Sistemas Vivos - Sistemas de la Tierra y el Espacio - Sistemas Físicos- Sistemas Tecnológicos
	Área de aplicación	Salud - Recursos Naturales - Medio Ambiente - Fronteras - Riesgos de la Ciencia y la Tecnología
Usar evidencia científica	Formato	Abierta - Cerrada
	Contexto	Personal - Social - Global
	Área de aplicación	Salud - Recursos Naturales - Medio Ambiente - Fronteras - Riesgos de la Ciencia y la Tecnología
	Habilidad	Interpretar pruebas científicas y elaborar y comunicar conclusiones Identificar los supuestos, las pruebas y los razonamientos que subyacen a las conclusiones Reflexionar sobre las implicancias sociales de los avances científicos y tecnológicos
Identificar cuestiones científicas	Formato	Abierta - Cerrada
	Contexto	Personal - Social - Global
	Área de aplicación	Salud - Recursos Naturales - Medio Ambiente - Fronteras - Riesgos de la Ciencia y la Tecnología
	Habilidad	Reconocer cuestiones susceptibles de ser investigadas científicamente Identificar términos clave para la búsqueda de información científica Reconocer los rasgos clave de la investigación científica

Fuente: OCDE, 2006

Tabla 2.
Ejemplos de contextos y áreas de aplicación de PISA Ciencias

Áreas de aplicación	Contextos		
	Personal (individuo, familia, compañeros)	Social (la comunidad)	Global (la vida, en todo el mundo)
Salud	Conservación de la salud personal, nutrición	Transmisión social de enfermedades, salud comunitaria	Epidemias
Recursos Naturales	Consumo personal de materiales y energía	Manutención de poblaciones humanas	Crecimiento demográfico, uso sostenible de las especies
Medio Ambiente	Uso y desecho de materiales	Impacto medioambiental, climas locales	Biodiversidad, sostenibilidad ecológica
Riesgos	Decisiones sobre la vivienda	Cambios rápidos (terremotos)	Cambio climático
Fronteras de la Ciencia y la Tecnología	Deporte y ocio, música, tecnología personal	Nuevos materiales, aparatos y procesos	Extinción de especies, exploración del espacio

Fuente: OCDE, 2006

1.3 Niveles de desempeño

PISA describe seis niveles de desempeño que muestran cómo progresan los conocimientos y habilidades de los estudiantes desde los más elementales, en el nivel 1, hasta los más complejos, en el nivel 6. Estos niveles son inclusivos, es decir, si un estudiante se ubica en cierto nivel, es capaz de realizar todas o la mayoría de las tareas definidas para los niveles inferiores. Cada estudiante es clasificado en un determinado nivel de desempeño, de acuerdo con el puntaje obtenido en la prueba⁴.

1.4 Resultados de los estudiantes chilenos

1.4.1 Resultados generales

En la escala general de ciencias (escala que combina el puntaje de las tres competencias o subescalas evaluadas), los estudiantes chilenos obtuvieron un puntaje promedio de 438 puntos. Este puntaje fue significativamente inferior al puntaje promedio de los países miembros de la OCDE, fijado en 500 puntos⁵.

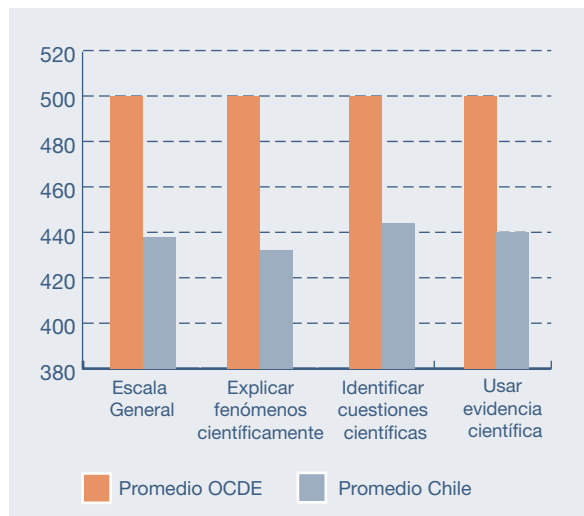
En cuanto a las competencias evaluadas, el promedio obtenido en la subescala *explicar fenómenos científicamente* fue de 432 puntos, en la subescala *identificar cuestiones científicas* fue de 444 puntos y en *usar evidencia científica* fue de 440 puntos. Todos estos puntajes también fueron significativamente inferiores a los 500 puntos promedio obtenidos por el conjunto de países OCDE, en cada una de estas subescalas.

4 En el primer capítulo de esta publicación se encuentra la descripción de cada uno de los niveles de logro de PISA Ciencias 2006.

5 Cabe destacar que este puntaje es significativamente superior al de todos los países latinoamericanos que participaron en PISA 2006 (MINEDUC-UCE, 2007).

En el Gráfico 2, se muestran los puntajes promedio obtenidos en la escala general de ciencias y en las subescalas de cada competencia evaluada, tanto por los estudiantes chilenos, como por los de los países OCDE. Como se observa, el mejor rendimiento se registró en la competencia *identificar cuestiones científicas* y el más bajo, en la competencia *explicar fenómenos científicamente*, por lo que se puede afirmar que esta es la competencia que presenta mayores dificultades para los estudiantes de Chile.

Gráfico 2: Puntaje promedio en la escala general y en las subescalas, para Chile y el grupo OCDE



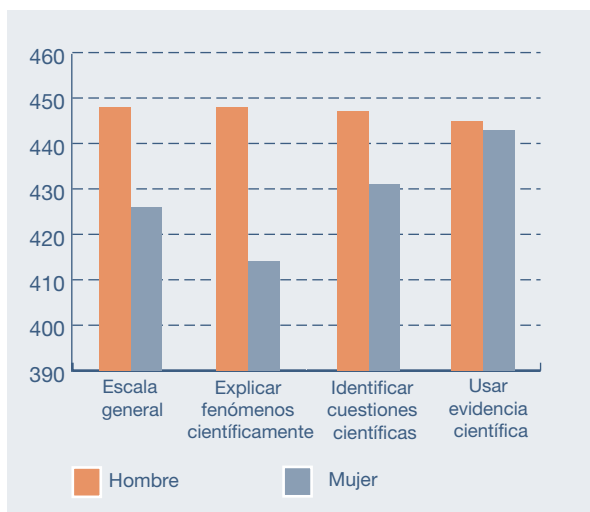
Fuente: MINEDUC-UCE, 2007

1.4.2 Resultados según género

La mayoría de los países que rindió la prueba PISA Ciencias en 2006 (37, de los 57 evaluados), no presentó diferencias significativas de puntaje entre hombres y mujeres, pero entre los 20 países donde hubo diferencias, en 12 ésta fue a favor de las mujeres y solo en 8, a favor de los hombres. En este último grupo se ubicó Chile, país que presentó diferencias

a favor de los hombres en todas las subescalas y, como consecuencia, también en la escala general de ciencias. Adicionalmente, de entre todos los países participantes, Chile presentó una de las mayores diferencias según género en favor de los hombres, siendo la subescala *explicar fenómenos científicamente* la que registró la mayor diferencia. En el Gráfico 3, se muestran las diferencias en los resultados según género, para Chile y la OCDE, en la escala general y en las subescalas de ciencias.

Gráfico 3: Diferencias de resultados por género, en la escala general de ciencias y en las subescalas de competencias, en los estudiantes chilenos



Fuente: MINEDUC-UCE, 2007

Los resultados de los estudiantes chilenos muestran que en 43,6% de las preguntas de la prueba, los hombres tuvieron un PRC promedio significativamente superior al de las mujeres. De estas preguntas, 30 evaluaban la competencia *explicar fenómenos científicamente*, lo que equivale al 70% de las 49 preguntas que evaluaban dicha competencia. Por su parte, solo hubo 3 preguntas de la prueba que presentaron diferencias significativas a favor de las mujeres, lo que equivale a un 2,9% del total de preguntas.

Para el grupo OCDE, aunque en la mayoría de los casos no hubo diferencias significativas por género en los puntajes totales de la prueba, se pudo observar la existencia de ciertas fortalezas diferenciadas, al interior de las competencias. Al igual que en Chile, se observó que los hombres tienden a rendir mejor en la competencia *explicar fenómenos científicamente*, aunque la diferencia es bastante menor en el grupo OCDE que en Chile. Sin embargo, en el promedio de países OCDE las mujeres tienden a rendir mejor en *identificar cuestiones científicas*, mientras que en los resultados para Chile, esta subescala no presenta diferencias por género.

1.4.3 Resultados según niveles de desempeño

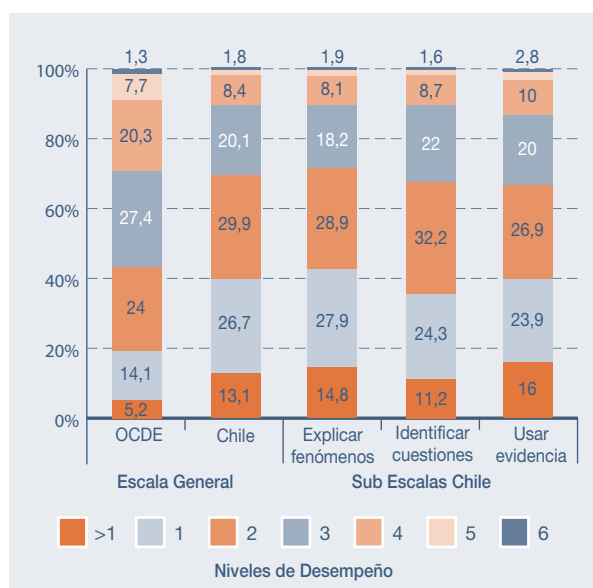
Como se observa en el Gráfico 4, en todas las subescalas, cerca de 40% de los estudiantes chilenos no alcanzó el nivel 2 de la escala general, que es el nivel que PISA considera de “aptitud básica”, pues a partir de él, los estudiantes comienzan a demostrar las competencias en ciencias que les permitirán participar de manera eficaz en situaciones de la vida real, relacionadas con la ciencia y la tecnología (OCDE, 2008).

Ahora bien, si se consideran los resultados por subescala, se observa que, en la competencia *explicar fenómenos científicamente*, mientras 57,2% de los estudiantes chilenos, alcanzó o superó el nivel 2, el 42,7% no lo alcanzó. Por su parte, en la competencia *identificar cuestiones científicas* mientras 64,5% alcanzó o superó el nivel 2, 35,5% de los estudiantes chilenos no lo alcanzó. Finalmente, en la competencia *usar evidencia científica*, 60,1% de los estudiantes chilenos alcanzó o superó el nivel 2, mientras 39,9% no lo logró. En cuanto a las tres competencias evaluadas, la gran mayoría de estudiantes chilenos que alcanzó o superó

el nivel 2, se concentró entre los niveles 2 y 4 (96,5%) y solo un porcentaje cercano al 2%, alcanzó los dos más altos niveles de logro.

A diferencia de lo que ocurre con los estudiantes de Chile, en promedio, solo alrededor de 20% de los estudiantes de los países de la OCDE, no alcanzó el nivel 2 y un porcentaje cercano al 10% alcanzó los dos niveles más altos de desempeño⁶.

Gráfico 4: Porcentaje de estudiantes chilenos y OCDE en cada nivel de desempeño



Fuente: OCDE, 2006

⁶ En el anexo del texto "Rendimiento de estudiantes de 15 años en Ciencias, Lectura y Matemática" (MINEDUC-UCE, 2007), se encuentra la descripción de cada uno de los niveles de logro de PISA Ciencias 2006.

2. Objetivos del estudio

El objetivo general de este estudio, consistió en identificar, a partir de los resultados obtenidos por los estudiantes chilenos en la escala general y en las subescalas de la prueba PISA Ciencias, los aspectos en los que los estudiantes chilenos presentan las mayores dificultades y/o fortalezas respecto de su competencia científica, tomando como referencia los resultados de los estudiantes del conjunto de países OCDE.

Sobre la base de este objetivo general, se plantearon los siguientes objetivos específicos:

1. Determinar si la dificultad de las preguntas se asocia, en el caso de los estudiantes chilenos, con el formato de las preguntas, es decir, si corresponden a preguntas abiertas o preguntas cerradas, tomando como referencia los resultados de los estudiantes del conjunto de países OCDE.
2. Determinar si la dificultad de las preguntas se asocia, en el caso de los estudiantes chilenos, con alguno de los contextos, áreas de aplicación, o áreas de contenido en que se plantean las preguntas, o bien, con alguna de las habilidades evaluadas, tomando como referencia los resultados de los estudiantes del conjunto de países OCDE.
3. Identificar fortalezas de los estudiantes chilenos respecto de su competencia científica, considerando los distintos aspectos evaluados, tomando como referencia el conjunto de países OCDE.
4. Identificar algunos de los aspectos de las preguntas que resultan particularmente difíciles para las estudiantes mujeres en Chile, tomando como referencia los resultados de las mujeres del conjunto de países OCDE.

3. Metodología

La primera etapa de este estudio, consistió en extraer de la base de datos de los resultados PISA Ciencias 2006, los PRC promedio de los estudiantes chilenos, los de los países OCDE, y los PRC promedio de hombres y mujeres por separado, en cada una de las preguntas de la prueba. Luego, se aplicó el test de comparación de proporciones para muestras independientes de Wald (Agresti, 2007), para detectar diferencias significativas entre los PRC promedio de los diferentes grupos de estudiantes.

A continuación, para cada una de las tres subescalas, se realizó un análisis estadístico de clasificación de conglomerados⁷, intencionando la formación de tres grupos de distinta dificultad (baja, media y alta), en base al PRC promedio de cada pregunta que conforma cada subescala. Como resultado, se formaron, para Chile y para la OCDE por separado, tres grupos de dificultad en cada subescala y tres grupos de dificultad en la subescala *explicar fenómenos científicamente* para cada género, indicando para cada grupo, la cantidad de preguntas que lo conforman, su PRC promedio⁸ y la identificación de las preguntas que lo integran.

Luego de tener cada una de las preguntas clasificadas en alguno de los grupos de dificultad, se observó la distribución de las mismas, en las distintas categorías de las variables⁹ y se comparó esta distribución en los grupos de dificultad formados para Chile y para la OCDE. Por ejemplo, al considerar la variable *formato de pregunta*, con sus categorías *abierto-cerrado*, se buscó identificar si había mayor cantidad relativa de preguntas abiertas en alguno de los grupos de dificultad

para Chile y, luego, si ocurría de la misma manera en los grupos de dificultad de los países de la OCDE.

Finalmente, se realizó el mismo análisis de conglomerados, esta vez, diferenciando por género, de manera de identificar grupos de preguntas de distinta dificultad para las mujeres y para los hombres por separado, tanto para Chile como para la OCDE. Esto permitió caracterizar con más detalle las diferencias de resultados entre hombres y mujeres, utilizando las mismas variables estudiadas para la población general. Dado que se reportó que la subescala que presentó mayores diferencias entre hombres y mujeres en Chile fue *explicar fenómenos científicamente* (MINEDUC, 2007), este análisis se realizó solo para esta subescala.

4. Resultados

4.1 Diferencias en los PRC promedio en la escala general y en las subescalas de ciencias, entre Chile y el grupo OCDE

En la escala general de ciencias, alrededor del 86% de las preguntas presentan un PRC promedio significativamente superior, a favor de la OCDE.

Respecto de las subescalas, se observa que, en alrededor del 90% de las preguntas de la subescala *explicar fenómenos científicamente*, Chile presenta PRC promedio significativamente inferiores a los de la OCDE. Es decir, proporcionalmente, hay más preguntas con diferencia significativa en esta subescala que en la escala general. En alrededor del 94% de las preguntas de la subescala *usar evidencia científica*, Chile presenta PRC promedio significativamente inferiores a los de la OCDE. Proporcionalmente, esta subescala presenta la mayor cantidad de preguntas con diferencia significativa en sus PRC entre Chile y la OCDE. Por último, en la subescala *identificar cuestiones científicas*, se observa que, en alrededor del 70% de las preguntas, Chile presenta PRC promedio significativamente inferiores a los de la OCDE.

7 El análisis de conglomerados permite clasificar unidades (en este caso, preguntas) respecto de una o más variables (en este caso, Porcentaje de Respuesta Correcta promedio), de manera que las preguntas que pertenezcan a un mismo grupo sean lo más parecidas posibles dentro del grupo y lo más distintas entre los grupos.

8 Estos PRC promedio se usaron para confirmar las diferencias entre los grupos de distinta dificultad de Chile y la OCDE.

9 El detalle de las categorías que adopta cada variable estudiada se encuentra en la Tabla 1.

4.2 Grupos de dificultad en la escala general y en las subescalas de ciencias para Chile y el grupo OCDE

El análisis de conglomerados para la escala general y las subescalas, arrojó las siguientes conformaciones de grupos, tanto para los estudiantes chilenos, como para los de la OCDE:

Tabla 3.
Grupos de dificultad en la escala general de ciencias

Grupo de dificultad		Chile	OCDE
Baja	Cantidad de preguntas	20	26
	PRC promedio	67,9	74,5
Media	Cantidad de preguntas	46	41
	PRC promedio	45,8	56
Alta	Cantidad de preguntas	37	36
	PRC promedio	24,9	36,2
Cantidad total de preguntas		103	103

Tabla 4.
Grupos de dificultad en la subescala explicar fenómenos científicamente

Grupo de dificultad		Chile	OCDE
Baja	Cantidad de preguntas	11	14
	PRC promedio	70,1	76,2
Media	Cantidad de preguntas	22	21
	PRC promedio	45,2	55,6
Alta	Cantidad de preguntas	16	14
	PRC promedio	23,2	34,5
Cantidad total de preguntas		49	49

Tabla 5.
Grupos de dificultad en la subescala usar evidencia científica

Grupo de dificultad		Chile	OCDE
Baja	Cantidad de preguntas	1	7
	PRC promedio	84,06	68,35
Media	Cantidad de preguntas	14	14
	PRC promedio	49,28	52,51
Alta	Cantidad de preguntas	16	10
	PRC promedio	26,89	34,07
Cantidad total de preguntas		31	31

Tabla 6.
Grupos de dificultad en la subescala identificar cuestiones científicas

Grupo de dificultad		Chile	OCDE
Baja	Cantidad de preguntas	1	1
	PRC promedio	80,26	80,26
Media	Cantidad de preguntas	13	13
	PRC promedio	53,22	53,22
Alta	Cantidad de preguntas	9	9
	PRC promedio	30,88	30,88
Cantidad total de preguntas		23	23

En las Tablas 5 y 6 se observa que en las subescalas *usar evidencia científica e identificar cuestiones científicas*, en el caso de Chile, prácticamente no hay preguntas de baja dificultad, situación que no ocurre en la OCDE.

4.3 Distribución de las preguntas, según PRC promedio, por grupos de dificultad, en Chile y la OCDE

En los Gráficos 5, 6 y 7, se ilustra el PRC promedio de los estudiantes de Chile y la OCDE en cada pregunta, en los distintos grupos de dificultad para cada subescala, respectivamente. De esta manera, se han ordenado en cada gráfico todas las preguntas que evalúan la competencia correspondiente, de mayor a menor PRC promedio, identificando con distintos símbolos los tres grupos de dificultad formados. En el eje X se representan las distintas preguntas de cada subescala, ordenadas de mayor a menor PRC promedio.

Gráfico 5: PRC promedio de los estudiantes de Chile y la OCDE, en cada una de las preguntas de cada grupo de dificultad, de la subescala *explicar fenómenos científicamente*

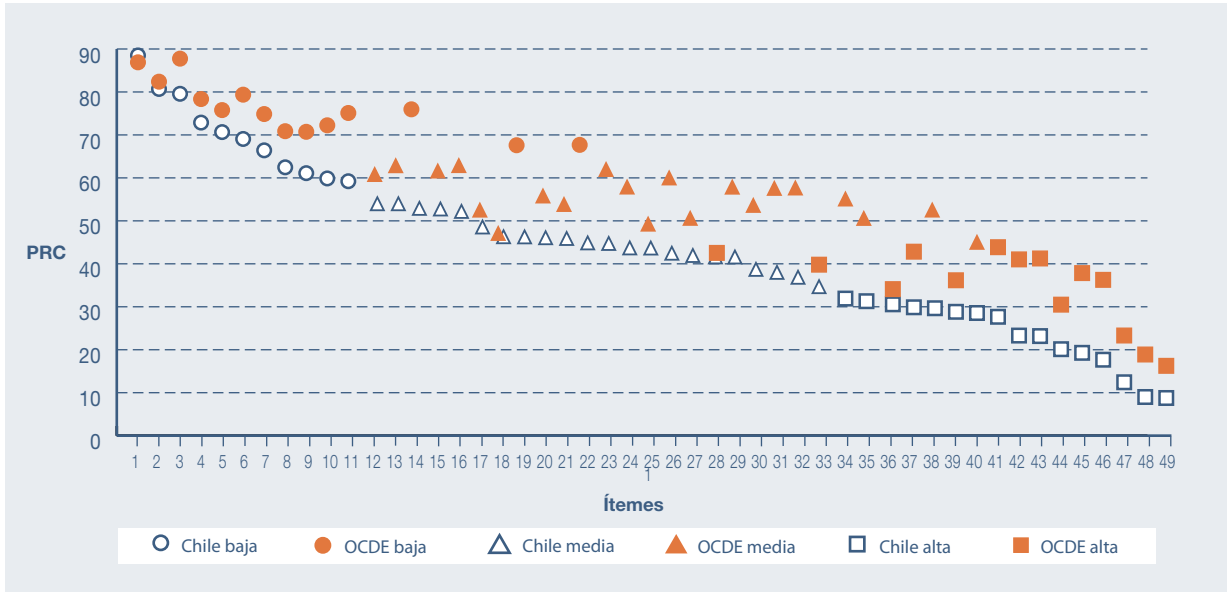


Gráfico 6: PRC promedio de los estudiantes de Chile y la OCDE, en cada una de las preguntas de cada grupo de dificultad, de la subescala *usar evidencia científica*

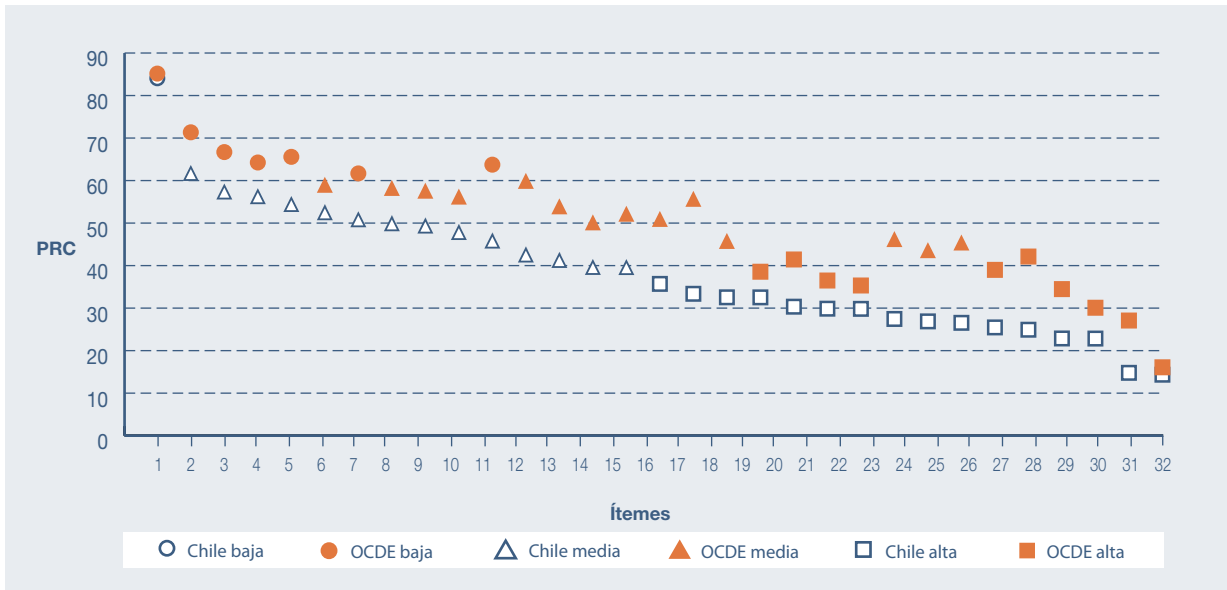
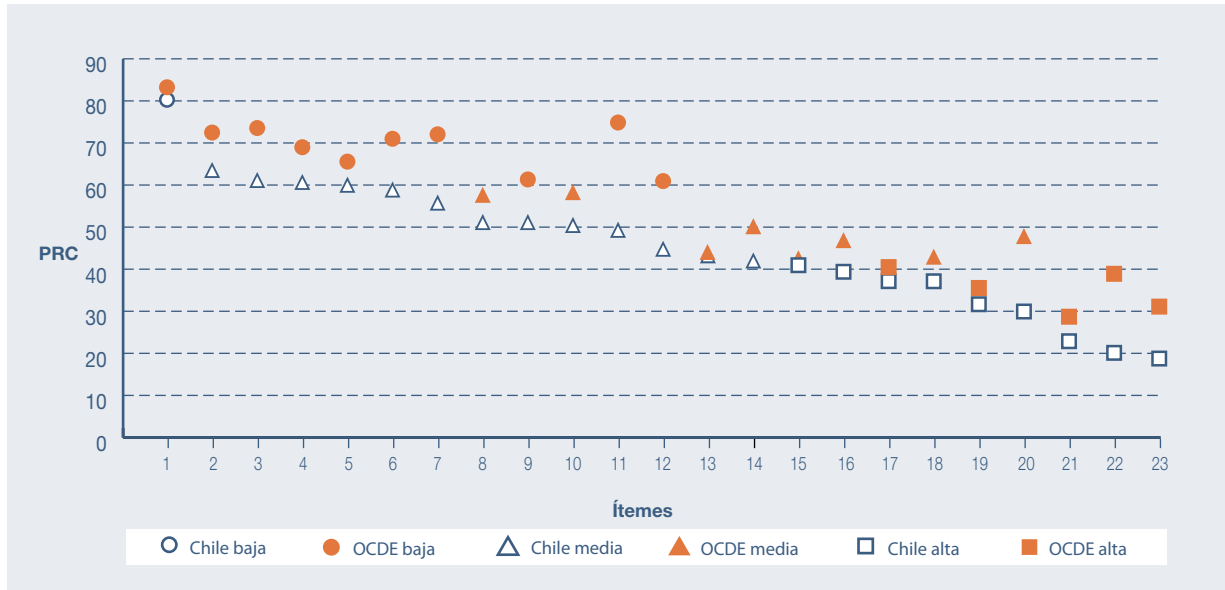


Gráfico 7: PRC promedio de los estudiantes de Chile y la OCDE, en cada una de las preguntas de cada grupo de dificultad, de la subescala *identificar cuestiones científicas*



En los gráficos anteriores, se puede observar que la mayoría de las preguntas tiene un mayor PRC promedio para los estudiantes OCDE. Además, se observa cómo una misma pregunta puede clasificarse en distintos grupos de dificultad, dependiendo de si observamos los grupos de los estudiantes chilenos o los de la OCDE. Es decir, algunas de las preguntas que para la OCDE se ubicaron en los grupos de dificultad baja, para Chile se ubicaron en los grupos de dificultad media, y algunas de las preguntas que para la OCDE se ubicaron en los grupos de dificultad media, para Chile se ubicaron en los grupos de dificultad alta.

Este desplazamiento de las preguntas en los grupos de dificultad, se observa como un patrón en todas las

subescalas, aunque no en la misma magnitud. Así, en la subescala *explicar fenómenos científicamente* (Gráfico 5) hay menor distancia con la OCDE para las preguntas de dificultad baja y alta; en cambio, las preguntas de los grupos de dificultad media se encuentran más distantes. Es decir, hay comparativamente más brecha en los resultados, en las preguntas de mediana dificultad. En el Gráfico 6, correspondiente a la subescala *usar evidencia científica*, se observa un comportamiento similar al anterior, salvo para las preguntas de la 18 a la 22. Por último en la subescala *identificar cuestiones científicas* (Gráfico 7) se puede observar que, salvo para algunas preguntas, los PRC promedio en Chile se mantienen a una distancia similar de los PRC promedio de la OCDE.

4.4 Distribución de las preguntas, según formato, por grupos de dificultad, en Chile y la OCDE

En el Anexo, se presentan las tablas con la información que sirvió de base para el análisis presentado en las secciones 4.4 a 4.8.

Desde el punto de vista del *formato de las preguntas* que conforman los distintos grupos de dificultad, tanto en la escala general como en las tres subescalas de competencias, la distribución de las preguntas es similar entre Chile y los países OCDE. En efecto, tanto para los estudiantes chilenos como para los de la OCDE, la mayor parte de las *preguntas abiertas* se

ubican en el grupo de preguntas de alta dificultad y en ambos casos el grupo de dificultad baja concentra un mínimo de *preguntas abiertas*.

Cabe destacar que en las subescalas *usar evidencia científica e identificar cuestiones científicas*, en Chile prácticamente no hay *preguntas cerradas* que se ubiquen en el grupo de baja dificultad: en su mayoría pertenecen al grupo de dificultad alta. En cambio, en la OCDE, las *preguntas cerradas* se encuentran distribuidas en los tres grupos de dificultad.

A continuación se muestran algunos ejemplos de *preguntas abiertas* con sus respectivas pautas de corrección.

EFECTO INVERNADERO

Lee los siguiente textos y responde las preguntas que aparecen a continuación.

EL EFECTO INVERNADERO: ¿REALIDAD O FICCIÓN?

Los seres vivos necesitan energía para sobrevivir. La energía que mantiene la vida sobre la Tierra proviene del Sol, el cual irradia energía al espacio porque es muy caliente. Una pequeña proporción de esta energía llega a la Tierra.

La atmósfera terrestre actúa como una capa protectora de la superficie de nuestro planeta, evitando las variaciones de temperatura que existirían en un mundo sin aire.

La mayor parte de la energía irradiada proveniente del Sol pasa a través de la atmósfera terrestre. La Tierra absorbe una parte de esta energía y otra parte es reflejada por la superficie de la Tierra. Parte de esta energía reflejada es absorbida por la atmósfera.

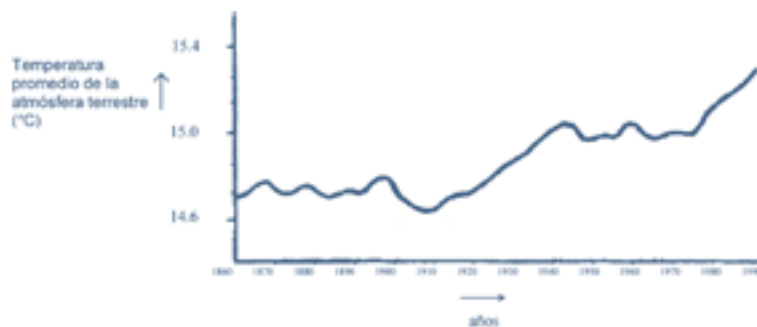
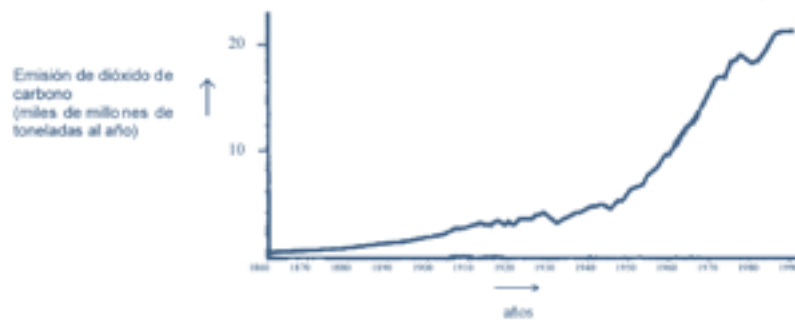
Como resultado de esto, la temperatura promedio sobre la superficie de la Tierra es más alta de lo que sería si no existiera atmósfera. La atmósfera terrestre funciona como un invernadero, de ahí la expresión *efecto invernadero*.

Se dice que el efecto invernadero se ha acentuado durante el siglo veinte.

Es un hecho que la temperatura promedio de la atmósfera ha aumentado. En diarios y revistas se afirma con frecuencia que el aumento de la emisión de dióxido de carbono es la principal causa del calentamiento ocurrido durante el siglo veinte.

Un alumno llamado Andrés, se interesa por la posible relación entre el promedio de la temperatura de la atmósfera terrestre y la emisión de dióxido de carbono en la Tierra.

En una biblioteca encuentra los dos gráficos siguientes:



A partir de estos dos gráficos, Andrés concluye que es cierto que el aumento de la temperatura promedio de la atmósfera terrestre se debe al aumento de la emisión de dióxido de carbono.

Pregunta 1: EFECTO INVERNADERO

¿Qué se observa en los gráficos que respaldaría la conclusión de Andrés?

.....

.....

Subescala	Usar evidencia científica			
Formato de pregunta	Abierta			
Contexto	Global			
Área de aplicación	Medio ambiente			
Habilidad	Identificar supuestos, pruebas razonamientos que subyacen a las conclusiones			
Resultados (PRC promedio)	OCDE: 54%		Chile: 41,3%	
	Mujeres: 55,08%	Hombres: 52,8%	Mujeres: 39,97%	Hombres: 42,47%

Pauta de Corrección	
Puntaje Completo	<p>Se refiere al incremento de ambos (promedio): temperatura y emisión de dióxido de carbono.</p> <ul style="list-style-type: none"> - <i>A medida que aumentaron las emisiones aumentó la temperatura.</i> - <i>Ambos gráficos muestran un aumento.</i> <p>Menciona (en términos generales) una relación positiva entre la temperatura y la emisión de dióxido de carbono.</p> <p>Nota: Se acepta terminología usada por los estudiantes como “relación positiva”, “forma similar” o “directamente proporcional”. Aunque los siguientes ejemplos de respuestas no son estrictamente correctos, muestran una comprensión suficiente para dar este puntaje.</p> <ul style="list-style-type: none"> - <i>La cantidad de CO2 y la temperatura promedio de la Tierra es directamente proporcional.</i> - <i>Tienen una forma similar, indicando una relación.</i>
Sin Puntaje	<p>Se refiere al aumento de la temperatura (promedio) o a la emisión de dióxido de carbono.</p> <ul style="list-style-type: none"> - <i>El CO2 está aumentando.</i> - <i>Muestra un cambio dramático en las temperaturas.</i> <p>Se refiere a la temperatura y las emisiones de dióxido de carbono sin estar claro sobre la naturaleza de la relación.</p> <ul style="list-style-type: none"> - <i>Las emisiones de CO2 tienen un efecto en el alza de la temperatura de la tierra.</i> <p>Otras Respuestas:</p> <ul style="list-style-type: none"> - <i>La emisión de CO2 está aumentando considerablemente más que el promedio de la temperatura de la Tierra.</i> Nota: Esta respuesta es incorrecta porque considera que el grado en que la emisión de CO2 y la temperatura están aumentando es la respuesta y no el hecho de que ambas están aumentando. - <i>El aumento de CO2 a través de los años se debe al aumento de la temperatura de la atmósfera terrestre.</i> - <i>La manera en que el gráfico sube.</i> - <i>Hay un aumento.</i>

En el ejemplo anterior, para contestar correctamente, los estudiantes tienen que identificar una parte de los gráficos que les sirva de evidencia para una determinada conclusión y señalar con claridad a qué parte se refieren. Específicamente, requiere que los estudiantes se den cuenta de que ambos gráficos, en promedio, en el

mismo periodo de tiempo, muestran una tendencia creciente. De esta manera, esta pregunta de formato abierto evalúa la competencia usar *evidencia científica* en la medida en que requiere que el alumno identifique evidencia que le sirve para apoyar una determinada conclusión, y la presente (OCDE, 2009).

Pregunta 2: EFECTO INVERNADERO

Otra alumna, Rosa, no está de acuerdo con la conclusión de Andrés. Ella compara los dos gráficos y afirma que algunas partes de los gráficos no respaldan su conclusión.

Da como ejemplo una parte de los gráficos que no respalde la conclusión de Andrés. Explica tu respuesta.

.....

.....

Subescala	Usar evidencia científica			
Formato de pregunta	Abierta			
Contexto	Global			
Área de aplicación	Medio ambiente			
Habilidad	Identificar supuestos, pruebas razonamientos que subyacen a las conclusiones			
Resultados (PRC promedio)	OCDE: 34,5%		Chile: 22,9%	
	Mujeres: 34,31%	Hombres: 34,63%	Mujeres: 20,95%	Hombres: 24,45%

Pauta de Corrección	
Puntaje Completo	<p>Menciona una parte en particular de los gráficos donde ambas curvas no están descendiendo o ascendiendo y proporciona la explicación correspondiente.</p> <ul style="list-style-type: none"> - De 1900-1910 (aproximadamente) el CO2 estaba aumentando, mientras la temperatura estaba bajando. - De 1980-1983 el dióxido de carbono bajó y la temperatura subió. - La temperatura en los 1800 es casi la misma pero el primer gráfico mantiene un alza. - Entre 1950 y 1980 la temperatura no aumentó pero el CO2 si lo hizo. - Desde 1940 hasta 1975 la temperatura permanece casi igual pero la emisión de dióxido de carbono muestra un drástico aumento.
Puntaje Parcial	<p>Menciona un periodo correcto, sin ninguna explicación.</p> <ul style="list-style-type: none"> - 1930-1933. - Antes de 1910. <p>Menciona sólo un año en particular (no un período de tiempo), con una explicación aceptable.</p> <ul style="list-style-type: none"> -En 1980 las emisiones bajaron pero la temperatura siguió subiendo. <p>Da un ejemplo que no apoya la conclusión de Andrés pero comete un error al mencionar el período. (Nota: Debe haber evidencia de este error- por ejemplo, en el gráfico se marca un área que claramente indica una respuesta correcta y posteriormente se comete un error al transferir esta información al texto).</p> <ul style="list-style-type: none"> -Entre 1950 y 1960 la temperatura disminuyó y la emisión de dióxido de carbono aumentó. <p>Menciona las diferencias entre las dos curvas, sin mencionar un periodo específico.</p> <ul style="list-style-type: none"> -En algunos lugares la temperatura se eleva aun si la emisión disminuye. <p>Se refiere a una irregularidad en uno de los gráficos.</p> <ul style="list-style-type: none"> -Es aproximadamente en 1910, cuando la temperatura había descendido y esto duró un cierto periodo de tiempo. <p>Señala diferencias claras entre los gráficos, pero su explicación es insuficiente.</p> <ul style="list-style-type: none"> -En la década de 1940 el calor fue muy alto pero el dióxido de carbono fue muy bajo.
Sin Puntaje	<p>Menciona una irregularidad en una curva sin referirse específicamente a los gráficos.</p> <ul style="list-style-type: none"> - Sube y baja un poco. - Bajó en 1930. <p>Se refiere a un periodo o año no muy claro ninguna explicación.</p> <ul style="list-style-type: none"> - La parte media. - 1910. <p>Otras Respuestas</p> <ul style="list-style-type: none"> - En 1940 la temperatura promedio aumentó, pero no lo hizo la emisión de dióxido de carbono. - Alrededor de 1910 la temperatura ha aumentado pero no la emisión.

En este segundo ejemplo, para contestar correctamente, los estudiantes tienen que identificar y presentar una parte de los gráficos que no provea de evidencia para una determinada conclusión. Específicamente, requiere que los estudiantes busquen diferencias entre los gráficos, que tienen en general una tendencia parecida al ascenso. Así, tienen que reconocer que deben identificar una parte de los gráficos donde ambas curvas no están ascendiendo o descendiendo, y presentar este hallazgo como parte de una justificación para la conclusión. En otras palabras, deben explicar la diferencia que han encontrado. Aquellos estudiantes que solo reconocen que hay una diferencia, pero no proveen explicación de por qué sirve como evidencia contraria a la conclusión, se clasifican en el nivel 4 de desempeño. Así, esta pregunta de formato abierto evalúa la competencia *usar evidencia científica* en la medida en que requiere que el alumno identifique evidencia contraria a una determinada conclusión (OCDE, 2009).

4.5 Distribución de las preguntas, según contexto, por grupos de dificultad, en Chile y la OCDE

En la escala general y en las subescalas *explicar fenómenos científicos* y *usar evidencia científica*, tanto para los estudiantes chilenos, como para los de la OCDE, la mayor parte de las preguntas planteadas en un *contexto global* se concentra en los grupos de dificultad media y alta y prácticamente no hay preguntas de *contexto global* en el grupo de dificultad baja. Dado que en la subescala *identificar cuestiones científicas* hay solo una pregunta referida a *contexto global*, no es posible establecer un patrón en esta subescala.

Ahora bien, en la subescala *usar evidencia científica*, a diferencia de la OCDE, en Chile, las preguntas

planteadas en *contexto social* se ubican en su mayoría en el grupo de dificultad alta, mientras que para la OCDE, estas preguntas se distribuyen de manera similar en por los tres grupos de dificultad. En la subescala *identificar cuestiones científicas* también se observa una diferencia entre Chile y la OCDE, en las preguntas planteadas en *contexto social*: mientras Chile tiene solo una de estas preguntas en el grupo de dificultad baja, la OCDE tiene más de la mitad de estas preguntas en ese grupo. En la sección 4.4 se muestran ejemplos de preguntas planteadas en un *contexto global*.

4.6 Distribución de las preguntas, según área de aplicación, por grupos de dificultad, en Chile y la OCDE

En la *escala general* y en la subescala *explicar fenómenos científicamente* existe una distribución similar de dificultad entre Chile y la OCDE, de las preguntas según *área de aplicación*. En ambos casos, para cada una de las áreas de aplicación las preguntas se distribuyen en los tres grupos de dificultad. Sin embargo, en cada área de aplicación, independientemente del grupo de dificultad, los estudiantes de la OCDE logran un mayor PRC promedio.

En la subescala *usar evidencia científica*, las áreas de aplicación *salud* y *medio ambiente* presentan distribuciones de las preguntas por grupo de dificultad similares para Chile y la OCDE y en ambos casos prácticamente no hay preguntas de *salud* y *medioambiente* que se ubiquen en los grupos de dificultad baja (solo una en el caso de la OCDE). Ahora bien, en las áreas de aplicación *recursos naturales* y *fronteras de la ciencia y la tecnología*, se observa una diferencia entre Chile y la OCDE: en Chile, en ninguna de estas áreas de aplicación, hay preguntas en el grupo de dificultad baja, en cambio en la OCDE sí las hay.

En la subescala *identificar cuestiones científicas*, al igual que en la subescala anterior, tanto para Chile como para la OCDE no hay preguntas del área de aplicación *salud* que se ubiquen en los grupos de dificultad baja. Ahora bien, en las áreas de aplicación *medio ambiente* y *fronteras de la ciencia y la tecnología* se observa una diferencia entre Chile y la OCDE: mientras que para Chile prácticamente no hay preguntas de estas

áreas de aplicación en el grupo de dificultad baja; la mayoría de las preguntas de estas áreas de aplicación se ubican en el grupo fácil de la OCDE.

A continuación se muestran dos preguntas del área de aplicación *fronteras de la ciencia y la tecnología* y tres preguntas del área de aplicación *salud*.

CULTIVOS GENÉTICAMENTE MODIFICADOS

DEBERÍA PROHIBIRSE EL MAÍZ GM

Los grupos de conservación de la vida silvestre están exigiendo que se prohíba un nuevo maíz genéticamente modificado (GM).

Este maíz GM está diseñado para resistir un nuevo y potente herbicida que mata las plantas de maíz tradicional. Este nuevo herbicida matará la mayor parte de las malezas que crecen en los campos de cultivo de maíz.

Los conservacionistas dicen que, dado que estas malezas son el alimento de pequeños animales, especialmente insectos, el uso de este nuevo herbicida en el cultivo del maíz GM será dañino para el medio ambiente. Los que están a favor del uso del maíz GM afirman que un estudio científico ha demostrado que esto no ocurrirá.

Aquí están los detalles del estudio científico mencionado en el artículo anterior.

- Se sembró maíz en 200 campos a lo largo del país.
- Cada campo se dividió en dos. En una mitad se sembró el maíz genéticamente modificado (GM) tratado con el potente herbicida nuevo, y en la otra mitad se sembró el maíz convencional tratado con un herbicida convencional.
- El número de insectos encontrado en el maíz GM, tratado con el nuevo herbicida, fue más o menos el mismo que el número de insectos en el maíz convencional, tratado con el herbicida convencional.

Pregunta 1: CULTIVOS GENÉTICAMENTE MODIFICADOS

En el estudio científico mencionado en el artículo, ¿qué factores fueron variados intencionalmente? Encierra en un círculo "Sí" o "No" para cada uno de los siguientes factores.

¿Fue variado intencionalmente este factor en el estudio?	¿Sí o No?
El número de insectos en el ambiente.	Sí / No
Los tipos de herbicida usados.	Sí / No

Subescala	Identificar cuestiones científicas			
Formato de pregunta	Cerrada			
Contexto	Social			
Área de aplicación	Fronteras de la ciencia y la tecnología			
Habilidad	Reconocer los rasgos clave de una investigación científica			
Resultados (PRC promedio)	OCDE: 60,95%		Chile: 44,86%	
	Mujeres: 63,10%	Hombres: 58,81%	Mujeres: 45,59%	Hombres: 44,25%

Pauta de Corrección	
Puntaje completo	Las dos correctas: No, Sí, en este orden.

Pregunta 2: CULTIVOS GENÉTICAMENTE MODIFICADOS

El maíz se sembró en 200 campos a lo largo del país. ¿Por qué los científicos usaron más de un lugar?

- A Para que muchos agricultores pudieran probar el nuevo maíz GM.
- B Para saber cuánto maíz GM podían cultivar.
- C Para cubrir tanta tierra como fuera posible con el cultivo GM.
- D Para incluir varias condiciones de crecimiento para el maíz.

Subescala	Identificar cuestiones científicas			
Formato de pregunta	Cerrada			
Contexto	Social			
Área de aplicación	Fronteras de la ciencia y la tecnología			
Habilidad	Reconocer los rasgos clave de una investigación científica			
Clave	D. Para incluir varias condiciones de crecimiento para el maíz			
Resultados (PRC promedio)	OCDE: 73,6%		Chile: 61,2%	
	Mujeres: 74,42%	Hombres: 72,71%	Mujeres: 59,75%	Hombres: 62,42%

Los estudiantes que contestan correctamente estas preguntas, demuestran, en el caso de la primera pregunta, haber comprendido el diseño del experimento, lo que permite reconocer un factor que en el estudio fue controlado. En el caso de la segunda pregunta, demuestran haber comprendido la necesidad de estudiar el comportamiento del cultivo en variadas condiciones ambientales, lo que se da al cultivarlo en 200 sitios a lo largo del país, ya que de lo contrario, las conclusiones sobre el comportamiento del cultivo tendrían un alcance más limitado.

Estas preguntas se insertan en el área *fronteras de la ciencia y la tecnología* y evalúan la competencia *identificar cuestiones científicas*, en la medida en que demandan del estudiante la habilidad de reconocer rasgos clave de la investigación científica, a partir de una situación experimental concreta relacionada con la modificación genética de determinados cultivos y sus posibles consecuencias.

MARY MONTAGU

Lee el siguiente artículo de un diario y contesta las preguntas que se presentan a continuación.

LA HISTORIA DE LAS VACUNAS

Mary Montagu era una bella mujer. En 1715, sobrevivió a un ataque de viruela, pero quedó cubierta de cicatrices.

Mientras vivía en Turquía, en 1717, observó un método llamado inoculación que se practicaba allí comúnmente.

Este tratamiento consistía en rasguñar la piel, para transmitir una forma atenuada del virus de la viruela a personas saludables y jóvenes, que luego se enfermaban, pero en la mayoría de los casos solo adquirían una forma suave de la enfermedad.

Mary Montagu estaba tan convencida de que estas inoculaciones eran seguras, que permitió que se inoculara a su hijo y a su hija. En 1796, Edward Jenner usó inoculaciones con una enfermedad similar, la viruela de las vacas, para producir anticuerpos contra la viruela. Comparado con la inoculación de la viruela, este tratamiento tenía menos efectos secundarios y la persona tratada no podía infectar a otros. El tratamiento se hizo conocido como vacunación.

Pregunta 1: MARY MONTAGU

¿Contra qué tipo de enfermedades se puede vacunar a la gente?

- A Enfermedades hereditarias como la hemofilia.
- B Enfermedades causadas por virus, como la polio.
- C Enfermedades producto del mal funcionamiento del cuerpo, como la diabetes.
- D Cualquier enfermedad que no tenga cura.

Subescala	Explicar fenómenos científicamente			
Formato de pregunta	Cerrada			
Contexto	Social			
Área de aplicación	Salud			
Área de contenido	Sistemas Vivos			
Clave:	B. Enfermedades causadas por virus, como la polio			
Resultados (PRC promedio)	OCDE: 74,9%		Chile: 66,4%	
	Mujeres: 77,24%	Hombres: 72,56%	Mujeres: 63,99%	Hombres: 68,57%

Pregunta 2: MARY MONTAGU

Si los animales o los seres humanos contraen alguna enfermedad bacteriana infecciosa y luego se recuperan, generalmente, el tipo de bacteria que les causó la enfermedad no vuelve a enfermarlos.

¿Por qué sucede esto?

- A El cuerpo mata a todas las bacterias que pueden causar el mismo tipo de enfermedad.
- B El cuerpo genera anticuerpos que matan a este tipo de bacterias, antes de que se multipliquen.
- C Los glóbulos rojos matan a todas las bacterias que puedan causar el mismo tipo de enfermedad.
- D Los glóbulos rojos capturan y eliminan este tipo de bacterias del cuerpo.

Subescala	Explicar fenómenos científicamente			
Formato de pregunta	Cerrada			
Contexto	Social			
Área de aplicación	Salud			
Área de contenido	Sistemas vivos			
Clave	B. El cuerpo genera anticuerpos que matan a este tipo de bacterias, antes de que se multipliquen			
Resultados (PRC promedio)	OCDE: 75,1%		Chile: 59,2%	
	Mujeres: 74.93%	Hombres: 75.28%	Mujeres: 55.31%	Hombres: 62.75%

Pregunta 3: MARY MONTAGU

Da una razón por la que se recomienda que, en particular los niños pequeños y los ancianos, se vacunen contra la influenza (gripe).

.....

.....

.....

Subescala	Explicar fenómenos científicamente			
Formato de pregunta	Abierta			
Contexto	Social			
Área de aplicación	Salud			
Área de contenido	Sistemas vivos			
Resultados (PRC promedio)	OCDE: 58,76%		Chile: 52,89%	
	Mujeres: 61,53%	Hombres: 55,99%	Mujeres: 55,03%	Hombres: 50,94%

Pauta de Corrección	
Puntaje Completo	<p>Respuestas que mencionan que la gente muy joven y/o los ancianos tienen sistemas inmunes más débiles que otras personas, o algo similar. La razón (razones) dadas deben referirse a gente joven o anciana en particular, no a todo el mundo en general. Además, la respuesta debe indicar, directa o indirectamente, que estas personas tienen sistemas inmunes más débiles que otras personas, no solo que son, en términos generales, “más débiles”.</p> <ul style="list-style-type: none"> - <i>Ellos tienen menos resistencia a las enfermedades.</i> - <i>Los jóvenes y viejos no pueden combatir la enfermedad tan bien como los demás.</i> - <i>Es más probable que les dé gripe.</i> - <i>Si les da gripe, los efectos son peores para ellos.</i>
Sin Puntaje	<p>Otras respuestas.</p> <ul style="list-style-type: none"> - <i>Para que no les dé gripe.</i> - <i>Son más débiles.</i> - <i>Necesitan ayuda para combatir la gripe.</i>

Las tres preguntas pertenecen al área de aplicación salud y a la competencia explicar fenómenos científicamente en la medida en que las tres requieren que el estudiante aplique conocimientos científicos sobre el sistema inmunológico para, en la primera pregunta, poder establecer sobre qué tipo de enfermedades es posible desarrollar vacunas, en la segunda, poder reconocer la explicación al hecho de que, luego de haber contraído cierta enfermedad infecciosa bacteriana y recuperarse, el cuerpo no se vuelve a enfermar y, por último, en la tercera pregunta explicar el motivo por el que los ancianos y los niños requieren de vacunas contra la influenza.

4.7 Distribución de las preguntas, según área de contenido, por grupos de dificultad, en Chile y la OCDE

En los grupos de dificultad según *área de contenido*, se observa una distribución similar entre Chile y la OCDE, es decir, hay un subconjunto similar de preguntas en cada grupo de dificultad de cada una de las áreas de contenido. Ahora bien, cabe destacar que tanto para Chile como para la OCDE, cerca de la mitad de las preguntas del área de contenido *sistemas vivos*, se ubican en los grupos de dificultad alta.

4.8 Distribución de las preguntas, según habilidad, por grupos de dificultad, en Chile y la OCDE

En la subescala *usar evidencia científica* en Chile, a diferencia de la OCDE, no hay preguntas de la habilidad *interpretar pruebas científicas y formular conclusiones*, que resulten fáciles, sino que la mayor parte de las preguntas se concentran en el grupo de dificultad alta. En cambio, para los estudiantes de la OCDE, si bien hay preguntas en el grupo de dificultad alta, la mayor parte se concentra en los grupos de dificultad media y baja. No se observan mayores diferencias con la OCDE en la distribución de las preguntas pertenecientes a las otras dos habilidades.

En la subescala *identificar cuestiones científicas* se observa una diferencia entre Chile y la OCDE. En Chile no hay preguntas de baja dificultad en las habilidades *reconocer cuestiones susceptibles de ser investigadas científicamente* y *reconocer rasgos de la investigación científica*. En la OCDE, en cambio, se pueden encontrar preguntas de ambas habilidades, en los tres grupos de dificultad.

A continuación se muestra una pregunta de la prueba que pertenece a la subescala *identificar cuestiones científicas*.

LLUVIA ÁCIDA

Es posible simular el efecto de la lluvia ácida sobre el mármol colocando trozos de mármol en vinagre durante una noche. El vinagre y la lluvia ácida tienen prácticamente el mismo nivel de acidez. Cuando se pone un trozo de mármol en vinagre, se forman burbujas de gas. Uno puede determinar la masa del trozo de mármol seco antes y después del experimento.

Pregunta 1: LLUVIA ÁCIDA

Los alumnos que llevaron a cabo este experimento también pusieron los trozos de mármol en agua pura (destilada) por toda una noche.

Explica por qué los alumnos incluyeron este paso en su experimento.

.....

Subescala	Identificar cuestiones científicas			
Formato de pregunta	Abierta			
Contexto	Personal			
Área de aplicación	Riesgos			
Área de contenido	Reconocer los rasgos clave de una investigación científica			
Resultados (PRC promedio)	OCDE: 35,54%		Chile: 31,67%	
	Mujeres: 37,62%	Hombres: 33,52%	Mujeres: 29,86%	Hombres: 33,17%

Pauta de Corrección

Puntaje Completo	<p>Para mostrar que el ácido (vinagre) es necesario para la reacción.</p> <ul style="list-style-type: none"> - Para asegurarse de que el agua de la lluvia debe ser ácida, como la lluvia ácida, para causar esta reacción. - Para ver si hay otras razones para los agujeros en los trozos de mármol. - Porque muestra que los trozos de mármol no reaccionan con cualquier fluido ya que el agua es neutra.
Puntaje Parcial	<p>Para comparar con la prueba de vinagre y mármol, pero no deja claro que se realiza para mostrar que el ácido (vinagre) es necesario para la reacción.</p> <ul style="list-style-type: none"> - Para comparar con el otro tubo de ensayo. - Para ver si el trozo de mármol cambia en agua pura. - Los estudiantes incluyeron este paso para mostrar qué pasa cuando llueve normalmente sobre el mármol. - Para actuar como control. - Para ver la diferencia entre agua normal y agua ácida (vinagre).
Sin Puntaje	<p>Otras respuestas.</p> <ul style="list-style-type: none"> - Para mostrar que el agua destilada no era un ácido.

Los estudiantes que obtienen puntaje completo en esta pregunta, comprenden que es necesario mostrar que la reacción no ocurrirá en agua y que el vinagre es un reactante necesario. Poner mármol en agua, muestra comprensión del concepto de control en un experimento de carácter científico. Los estudiantes tienen que poder dar una explicación de manera clara. Aquellos que obtienen puntaje parcial, muestran cierto conocimiento de que el experimento implica una comparación, pero no logran demostrar que comprenden que el objetivo es mostrar que el vinagre es un reactante necesario, sin establecer que este es el objetivo final de la comparación. Esta pregunta requiere que los estudiantes muestren comprensión sobre la estructura de un experimento científico y por ello corresponde a la competencia *identificar cuestiones científicas* (OCDE, 2009).

4.9 Resultados del análisis por género¹⁰

El análisis por conglomerados, para la subescala *explicar fenómenos científicamente* arrojó las conformaciones de grupos, por género, para Chile y la OCDE que se muestran en la Tabla 7.

Tabla 7.
Grupos de dificultad en estudiantes mujeres en la subescala *explicar fenómenos científicamente*, para Chile y la OCDE

Grupo de dificultad		Chile	OCDE
Baja	Cantidad de preguntas	6	15
	PRC promedio	75,07	75,03
Media	Cantidad de preguntas	22	21
	PRC promedio	46,79	52,77
Alta	Cantidad de preguntas	21	13
	PRC promedio	23,12	31,84
Cantidad total de preguntas		49	49

En la Tabla 7, se observa que las preguntas se distribuyen de manera distinta en los grupos de dificultad de

¹⁰ En el Anexo se incluye el resto de las tablas con la información a partir de la cual se realizó el análisis presentado en esta sección.

Chile y la OCDE. Mientras para las mujeres de la OCDE 30,6% de las preguntas de esta subescala se ubican en el grupo de dificultad baja, para las mujeres chilenas solo 12,2% de las preguntas se ubican en ese grupo.

Tabla 8.
Grupos de dificultad en estudiantes hombres en la subescala *explicar fenómenos científicamente*, para Chile y la OCDE

Grupo de dificultad		Chile	OCDE
Baja	Cantidad de preguntas	13	17
	PRC promedio	71,36	74,78
Media	Cantidad de preguntas	21	26
	PRC promedio	46,42	52,02
Alta	Cantidad de preguntas	15	6
	PRC promedio	24,91	27,31
Cantidad total de preguntas		49	49

En la Tabla 8, se observa que, aunque la distribución de las preguntas por grupo de dificultad es diferente entre los estudiantes hombres de Chile y de la OCDE, esta diferencia es de menor magnitud que la observada entre las mujeres. Para los hombres de la OCDE, 34,7% de las preguntas de esta subescala se ubican en el grupo de dificultad baja y para los de Chile, 26,5% de las preguntas se ubica en ese grupo, lo que representa más del doble de la cantidad de preguntas de dificultad baja para las estudiantes mujeres de Chile (12,2%).

4.10 Distribución de las preguntas según género, por los grupos de dificultad, en Chile y la OCDE

En el Gráfico 8 se puede apreciar que existe, entre el desempeño de mujeres y hombres de Chile, un desplazamiento del grado de dificultad de las preguntas. Una gran cantidad de las preguntas que para los hombres se ubicaron en el grupo de dificultad baja, para las mujeres se ubicaron en el grupo de dificultad media, y algunas de las preguntas que para los hombres

Gráfico 8: PRC promedio de preguntas de dificultad baja, media y alta en hombres y mujeres de Chile, en la subescala *explicar fenómenos científicamente*

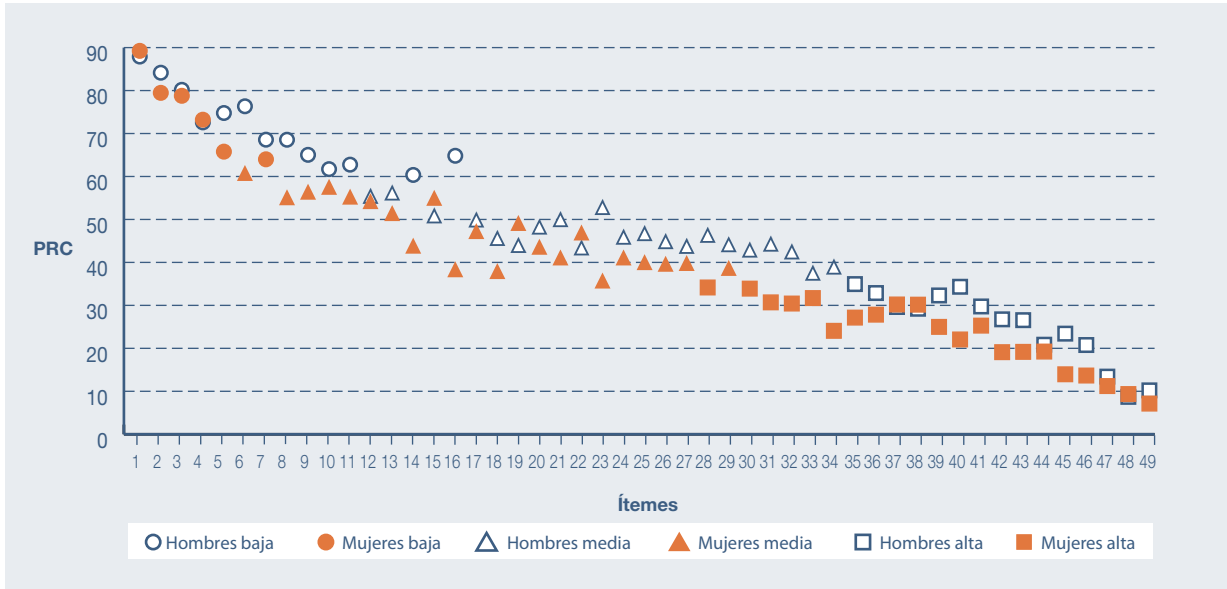
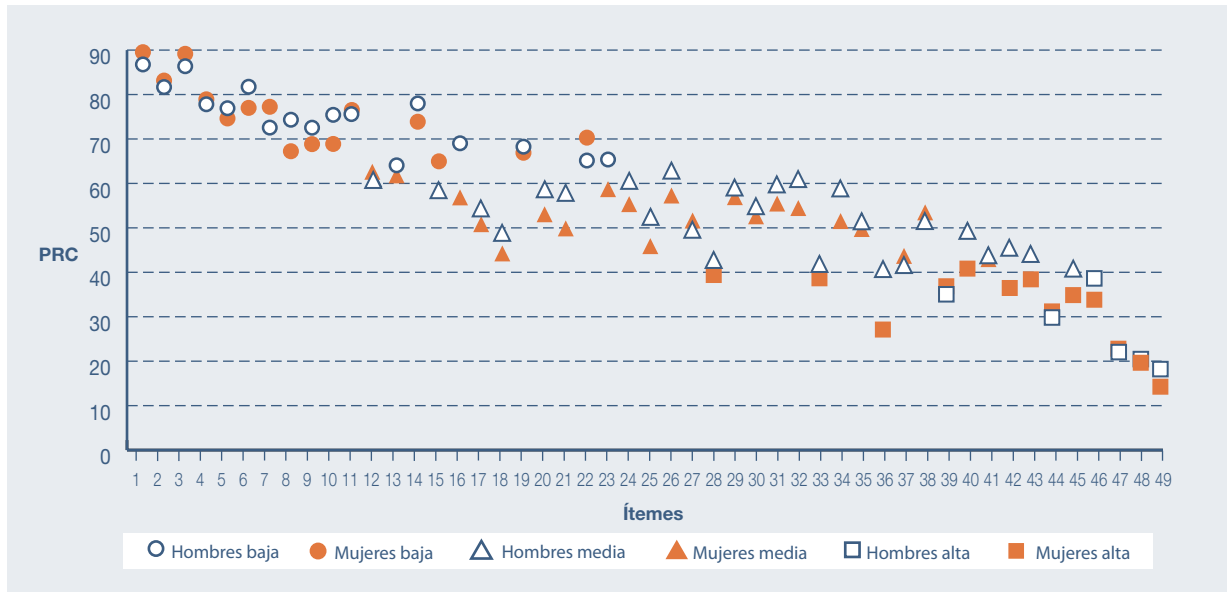


Gráfico 9: PRC promedio de preguntas de dificultad baja, media y alta en hombres y mujeres del grupo OCDE, en la subescala *explicar fenómenos científicamente*



se ubicaron en el grupo de dificultad media, para las mujeres se ubicaron en el grupo de dificultad alta.

También se puede observar que en la mayoría de las preguntas, los estudiantes hombres chilenos obtuvieron mayores PRC promedio que las mujeres.

En el Gráfico 9, en cambio, se puede observar que en la mayoría de las preguntas, los PRC promedio de hombres y mujeres tienden a ser más cercanos, es decir, se observa que estudiantes hombres y mujeres de la OCDE presentan desempeños similares.

A partir de los gráficos 10 y 11 se observa que, al comparar los resultados de los estudiantes hombres de Chile con sus pares OCDE y los resultados de las estudiantes mujeres de Chile con sus pares OCDE, hay

más similitudes en los resultados entre los dos grupos de hombres que entre los dos grupos de mujeres. Es decir, las mujeres chilenas se distancian más de sus pares OCDE que los hombres.

Gráfico 10: PRC promedio de preguntas de dificultad baja, media y alta en hombres de Chile y el grupo OCDE, en la subescala *explicar fenómenos científicamente*

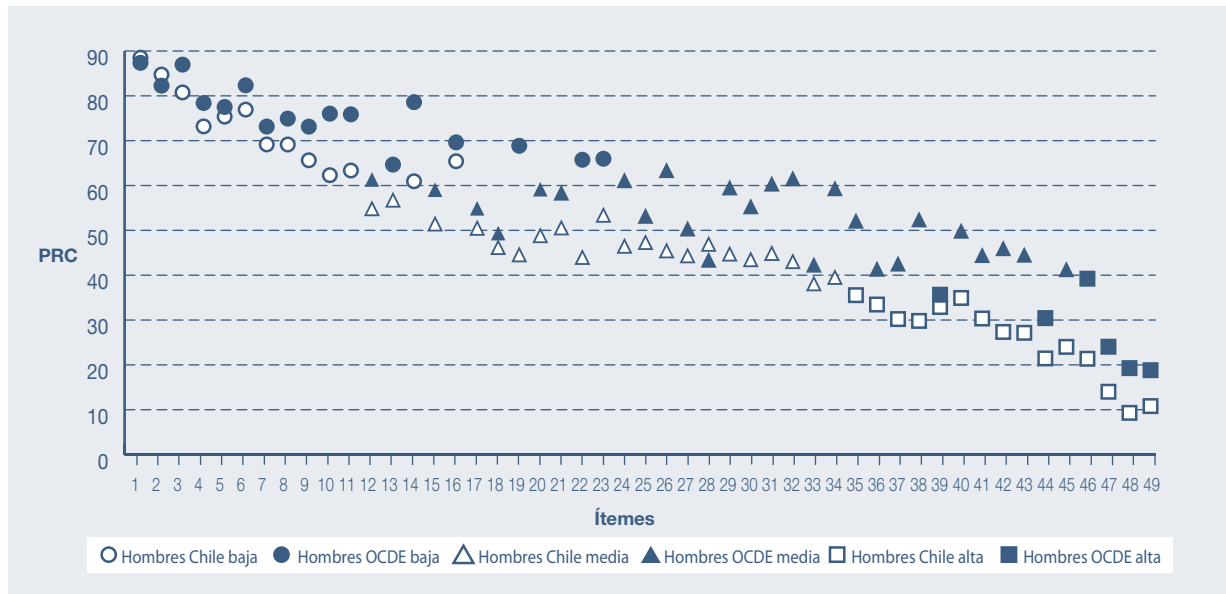
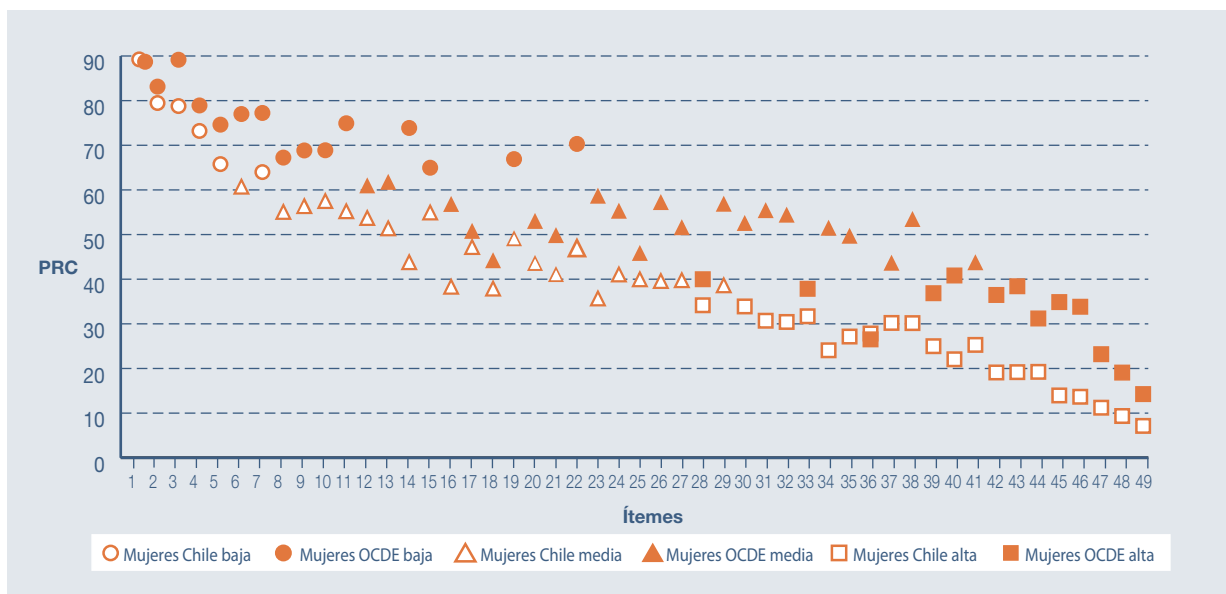


Gráfico 11: PRC promedio de preguntas de dificultad baja, media y alta en mujeres de Chile y del grupo OCDE, en la subescala *explicar fenómenos científicamente*



4.11 Análisis comparado de la distribución de las preguntas por grupo de dificultad, de acuerdo a formato, contexto, área de aplicación y área de contenido, en Chile y la OCDE

Respecto de las preguntas de *formato abierto*, se observa una distribución similar a la observada en los análisis anteriores, sin distinción por género. Es decir, prácticamente no se ubican preguntas de *formato abierto* en los grupos de baja dificultad, tanto para hombres como para mujeres de Chile y la OCDE. En cambio, al observar las preguntas de *formato cerrado*, se aprecia que en el caso de las mujeres de Chile, a diferencia de sus pares OCDE y de los hombres chilenos y de la OCDE, solo 17,6% de estas preguntas se ubican en el grupo de baja dificultad, mientras que para los otros grupos de estudiantes, sobre 35,2% de estas preguntas se ubican en el grupo de dificultad baja.

Respecto de las preguntas planteadas en *contexto global*, se observa una distribución similar a la observada en los análisis anteriores, sin distinción por género. Es decir, prácticamente no se ubican preguntas de *contexto global* en los grupos de baja dificultad ni para hombres ni para mujeres de Chile y la OCDE.

En relación con las *áreas de aplicación* en que se plantean las preguntas, ocurre lo mismo, es decir, se observa una distribución similar a la observada en los análisis anteriores, sin distinción por género.

Donde sí se encontraron diferencias, fue en las preguntas que tienen contenidos referidos a *sistemas vivos*: mientras para las mujeres chilenas, la mayoría de estas preguntas se ubican en el grupo de dificultad alta, para las mujeres de la OCDE estas preguntas se distribuyen en la misma proporción, en los tres grupos de dificultad.

5. Conclusiones

El análisis estadístico de conformación de conglomerados, a partir del cual se formaron tres grupos de preguntas de distinta dificultad para la escala general y para las subescalas de la prueba PISA Ciencias 2006, usando los resultados de Chile y de los países OCDE, permitió identificar aspectos comunes en los resultados de los estudiantes chilenos y en los de la OCDE. Al mismo tiempo, permitió identificar diferencias y distinguir algunas debilidades presentes en el aprendizaje de los estudiantes chilenos.

En términos generales, en cada una de las categorías de las variables estudiadas, hay una distribución similar de las preguntas por grupo de dificultad, es decir, en general, hay preguntas de dificultad baja, media y alta, tanto para los estudiantes chilenos como para los de la OCDE. Ahora bien, en el caso de Chile, la mayoría de las preguntas se encuentran desplazadas a un grupo superior de dificultad, es decir, preguntas que para la OCDE resultan fáciles, para Chile resultan de dificultad media; las que para la OCDE resultan de dificultad media, tienden a resultar difíciles para Chile, y las que resultan difíciles para la OCDE, también lo son para Chile. En otras palabras, se observa un patrón similar en la distribución de las preguntas por categorías, pero un desplazamiento a lo largo de los grupos de dificultad, donde los PRC promedio de cada uno de estos grupos, resultan inferiores para el caso de Chile.

En relación con la variable *formato de la pregunta*, aquellas preguntas que tienen *formato abierto* tienden a resultar más difíciles, tanto para los estudiantes chilenos, como para los de la OCDE, con independencia de la competencia que se está evaluando. Las preguntas abiertas demandan al estudiante construir y comunicar su respuesta y en muchos casos requieren que el estudiante desarrolle una explicación, argumento o justificación con claridad. De esta manera, más

allá del contenido o habilidad particular que se esté evaluando en una determinada pregunta, el *formato abierto* introduce la complejidad de poner en juego la capacidad de producir y comunicar de manera eficaz una respuesta, aspectos que están ausentes en las preguntas de *formato cerrado*. En la prueba PISA, más de un tercio de las preguntas son de formato abierto, dando cuenta de la relevancia de desarrollar en los estudiantes estas capacidades en el área de Ciencias, como parte de la competencia científica.

Por su parte, tanto para los estudiantes chilenos como para los de la OCDE, las preguntas planteadas en un *contexto global* son las que resultan más difíciles. Ahora bien, a los estudiantes chilenos, en las subescalas *usar pruebas científicas* e *identificar cuestiones científicas* también les resulta difícil resolver las preguntas que están situadas en *contextos sociales*. En este sentido, sería interesante indagar, por un lado, cuán aplicados a situaciones que van más allá de situaciones cercanas a los estudiantes, están aprendiendo los contenidos, los estudiantes chilenos y, por otro, si al momento de evaluar sus aprendizajes, se les está proveyendo de situaciones realmente nuevas que les permitan no solo reproducir su conocimiento, sino verdaderamente aplicarlo. Esto implica la importancia de exponer a los estudiantes a ejemplos de situaciones que vayan más allá de los explícitamente presentados en las clases.

Al analizar las preguntas, considerando la variable *área de aplicación*, tanto para los estudiantes chilenos como para sus pares de la OCDE, las preguntas referidas al área de *salud* tienden a resultar más difíciles. Una posible línea para indagar los motivos de que salud resulte más difícil tiene relación con la presencia y el modo en que se aborda esta área en los currículos de los países. Una particularidad de los estudiantes chilenos es que en las subescalas *usar evidencia científica* e *identificar cuestiones científicas*, tienden a presentar mayores dificultades en aquellas preguntas referidas al área *fronteras de la ciencia y la tecnología*. Esto puede constituir un hallazgo interesante en la línea

de fortalecer los vínculos de la ciencia y la tecnología en la enseñanza, tanto para comprender el impacto recíproco que se da en el desarrollo de ambas, como para advertir sus efectos en la vida de las personas. En cuanto a las áreas de *contenido*, las preguntas que demandan un manejo de conocimientos sobre *sistemas vivos*, tienden a resultar más difíciles tanto para los estudiantes chilenos como para los de la OCDE. Nuevamente, una posible línea a indagar, para buscar una explicación a esto, consiste en la revisión de los currículos de ciencias naturales y la forma en que se implementan.

En otro sentido, se aprecia que para los estudiantes chilenos, resulta particularmente difícil la habilidad de *interpretar pruebas científicas* y *formular conclusiones*, propia de la subescala *usar evidencia científica*. Esta habilidad se relaciona con la capacidad de los estudiantes de establecer conclusiones fundadas y de poder apoyarse en la evidencia disponible. Este hallazgo reafirma lo observado en los niveles de desempeño descritos por PISA, donde un alto porcentaje de estudiantes chilenos (66,78%) no alcanza el nivel 3, nivel en que se describen tareas como distinguir entre conclusiones adecuadas e inadecuadas, a partir de datos científicos disponibles. Respecto de la subescala *identificar cuestiones científicas*, en el caso de Chile, los resultados más bajos se concentran en la habilidad de *reconocer cuestiones susceptibles de ser investigadas científicamente*. Este hallazgo también se observa en los niveles de desempeño, donde 67,7% de los estudiantes de Chile no alcanza el nivel 3 de esta subescala, nivel en que se describe el dominio de esta habilidad. Estos resultados muestran la necesidad de enfatizar y fortalecer más en Chile el desarrollo de habilidades de pensamiento científico entre los estudiantes. Esto se relaciona con la importancia de la enseñanza de las ciencias naturales como un área de conocimiento que tiene características propias, en tanto disciplina científica, y que tiene una forma de construir conocimiento que la distingue de otras formas del saber. A su vez, destaca la importancia de que la comunidad, y especialmente los profesores de



ciencias, tengan internalizadas estas particularidades, por ejemplo, para estos últimos, a través del contacto, durante su formación, con la investigación científica y su proceder. De allí la conveniencia de prever los efectos de una excesiva separación entre instituciones a cargo de la formación de docentes y las facultades universitarias donde se realiza investigación en ciencia y tecnología.

“La educación en ciencias para la vida o alfabetización en ciencias, se entiende como la formación de personas capaces de tomar decisiones fundadas y de comprender y participar de sociedades donde el conocimiento científico esta inmensamente presente” (OCDE, 2006). Así, el problema de una pobre alfabetización científica o el desconocimiento de una parte importante de la población del quehacer de la ciencia, conlleva, tanto inequidad respecto de las oportunidades de participación de las personas, como una posible falta de apoyo y confianza en el desarrollo del conocimiento científico, por parte de estos sectores de la población. Esto es delicado en la medida en que la participación, el apoyo y la confianza constituyen conductas y actitudes deseables tanto para el desarrollo de la ciencia como para la eficaz promoción y apropiación de sus avances por parte de la sociedad.

Respecto de los resultados del análisis según género, al comparar los resultados generales de las estudiantes chilenas con sus pares OCDE versus los de los estudiantes hombres chilenos con sus pares OCDE, se observa que las mujeres son las que presentan la mayor brecha respecto de sus pares. En Chile, se observan diferencias significativas en los PRC promedio a favor de los hombres, tanto en la *escala general*, como en las subescalas *usar evidencia científica* y *explicar fenómenos científicamente*, siendo esta última la que presenta las mayores diferencias entre hombres y mujeres. Esto se confirma en la distribución de preguntas por grupo de dificultad, donde se observa que aquellas preguntas que conllevan la necesidad de aplicar conocimientos científicos a diversas situaciones, presentan mayores dificultades para las mujeres. Las preguntas referidas a sistemas vivos, por su parte, también resultan especialmente difíciles para las estudiantes mujeres. Cabe destacar que la diferencia en los resultados según género es concordante con lo observado en la prueba TIMSS del 2003 (MINEDUC-UCE, 2004, p.88), donde se reportó, para el caso de Chile, mejores rendimientos en Matemática y en Ciencias para los hombres. Por su parte, en las últimas mediciones SIMCE de Ciencias Naturales, aplicadas en 8° Básico (SIMCE, 2008, p.46) también se observaron diferencias a favor de los hombres. Es importante destacar que estas diferencias no son frecuentes a nivel internacional, por el contrario, en el área de Ciencias, en la mayoría de los países, no hay diferencias de rendimiento por género. Diversos estudios como el de Guiso (Guiso y otros, 2008) sugieren que las diferencias de rendimiento observadas entre hombres y mujeres en estas áreas no tienen sustrato biológico, sino que se relacionan fundamentalmente con el desarrollo social y cultural de las comunidades y los países, acortándose la brecha en aquellas sociedades donde hay más igualdad de género y más expectativas sobre las posibilidades de desarrollo y participación de las mujeres en la sociedad. En efecto, como sostienen diversos autores (González García, 1998) esto resta fuerza a la idea de que “cerebros diferentes” entre hombres y mujeres, son la causa de

las desigualdades existentes. De esta manera, estos resultados fortalecen la necesidad de profundizar en estudios enfocados a las diferencias de género en los aprendizajes de los estudiantes chilenos. Los análisis desde la perspectiva de género constituyen un área de relevancia crucial que revela interesantes aspectos de las interacciones entre la sociedad y la actividad científico-tecnológica y puede plantear desafíos a las políticas públicas de educación en Chile. Pueden existir mecanismos sutiles e implícitos que contribuyan a mantener y a legitimar las diferencias de género en los resultados de aprendizaje en el área de Ciencias, como la socialización que experimentan hombres y mujeres. A nivel escolar, el “currículum oculto” y las aspiraciones, expectativas y comportamientos de apoderados, profesores y estudiantes son factores que pueden influir en las diferencias de género que se observan en los resultados de aprendizaje.

Un dato interesante en el ámbito del desarrollo profesional de las mujeres chilenas en Ciencias, es el aportado recientemente por CONICYT (2008) en una revisión de la participación femenina en los concursos del programa de capital humano avanzado en el período 2001-2007. En este estudio se constató una menor participación femenina en la postulación a becas, la que se traspasa al número de aprobados, donde las mujeres se mantienen con un porcentaje inferior de participación que los hombres. Este es otro elemento que apoya la idea de analizar con mayor detalle y profundidad aquello que puede explicar en Chile las diferencias en los resultados de aprendizaje encontradas. Así, lo observado en los resultados de la prueba PISA Ciencias 2006 invita a realizar un profundo análisis de los factores socioculturales, educativos y psicológicos que pueden estar afectando los logros de las mujeres en Ciencias en nuestro país. Dado que las diferencias de género son trascendentales desde el punto de vista de la equidad, hay lecciones que estudiar de países que han logrado disminuir las brechas en los desempeños de hombres y mujeres. Del mismo modo, las brechas en los logros de apren-

dizaje encontradas en las distintas pruebas señaladas, ameritan un detallado monitoreo hacia el futuro.

Otro elemento que aportan los resultados la Prueba PISA Ciencias 2006, se refiere a que las debilidades en la formación en el área de Ciencias Naturales, son diversas entre los estudiantes chilenos, más allá de las particularidades que hemos descrito. Esto implica grandes desafíos para la formación científica de los estudiantes en Chile. De acuerdo con el *Informe McKinsey* (Barber y Mourshed, 2007) la evidencia disponible sugiere que el principal impulsor de las variaciones en el aprendizaje escolar es la calidad de los docentes. Concordante con esto, un reciente estudio (Bravo y otros, 2008) cruzó los puntajes en Matemática de 4° Básico de cada curso en el SIMCE con las mediciones de calidad docente de los últimos tres años, medidos por los sistemas nacionales de acreditación. Los resultados señalan que aquellos estudiantes que tuvieron, durante los tres años anteriores a la medición, profesores evaluados como competentes o destacados, obtuvieron mejores resultados que sus compañeros que no habían tenido profesores con un historial de exitosas evaluaciones. Habría que ver si estos resultados son extrapolables a todas las áreas de aprendizaje. Por su parte, otro estudio sobre profesores de Ciencias en Chile (Claro, 2003) señala que aunque la gran mayoría de los docentes que ejercen en Enseñanza Media están titulados en educación, cerca de la mitad no tiene formación específica en la disciplina científica que enseña. Considerando que el currículum nacional se va actualizando y que los programas de estudio contienen material que incluye conceptos científicos de desarrollo reciente, aun cuando la formación inicial de los docentes fuese óptima, de no mediar un fuerte apoyo en la actualización de cada profesor, su labor no podrá llenar las expectativas de dichos programas (Claro, 2003). Lo anterior confirma la necesidad de fortalecer, tanto la calidad de los estudiantes que serán futuros profesores, como la calidad e idoneidad de los programas que cursan en su formación inicial y la



necesidad de asegurar una formación continua que entregue apoyo a los profesores en todas las áreas de las ciencias.

En síntesis, el análisis de los resultados de la Prueba PISA Ciencias 2006, entrega una abundante fuente de información sobre los aprendizajes alcanzados por los estudiantes de Chile en el área de Ciencias. Los resultados del presente estudio levantan preguntas que conviene complementar con estudios de carácter cualitativo, por ejemplo, un estudio sobre cómo se aplican los contenidos al momento de enseñar y evaluar los aprendizajes o identificar y describir experiencias exitosas de desarrollo de habilidades de pensamiento científico en los estudiantes, o estudios relativos a distintas prácticas sociales y escolares, que puedan dar luces sobre las diferencias de género encontradas. Asimismo, el presente estudio podría ser complementado con un análisis cualitativo de las preguntas de la prueba, usando las diversas respuestas de los estudiantes. Por último, los resultados y conclusiones de este estudio podrían contrastarse o corroborarse con nuevos análisis detallados de resultados de los estudiantes de Chile, en otras evaluaciones de aprendizaje, como también con resultados de estudiantes de grupos etarios distintos al evaluado por la prueba PISA.

Referencias

- Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. Nueva Jersey, Estados Unidos: John Wiley & Sons Inc.
- Barber, M. y Mourshed, M. (2007). *How the World's Best-Performing School Systems Come Out On Top*. Extraído desde www.mckinsey.com
- Bravo, D. y otros (2008). *La relación entre la evaluación docente y el rendimiento de los alumnos: Evidencia para el caso de Chile*. Extraído desde http://www.microdatos.cl/docto_publicaciones/Evaluacion%20docentes_rendimiento%20escolar.pdf
- Claro, F. (2003). "Panorama docente de las Ciencias Naturales en Educación Media", *Revista de Educación*, 307. Santiago, Chile: Ministerio de Educación (MINEDUC).
- CONICYT (Comisión Nacional de Investigación Científica y Tecnológica) (2008). *Participación por género en las becas CONICYT. Periodo 2001-2007*. Extraído desde <http://www.conicyt.cl/573/article-31252.html>
- González García, M.I. (1998). "Convergencia y conflicto de valores: el caso de las diferencias sexuales en habilidades cognitivas", Ambrogi, A. (Ed.) *La naturalización de la filosofía de la ciencia*. Palma de Mallorca: Universidad de las Islas Baleares.
- Guiso, L. y otros (2008). "Culture, Gender and Math", *Education Forum, Science*, 320. Extraído desde www.sciencemag.org.
- MINEDUC (Ministerio de Educación)- UCE (Unidad de Currículum y Evaluación) (2004). *Chile y el aprendizaje de las matemáticas y ciencias según TIMSS*. Santiago, Chile.
- _____ (2007). *PISA 2006: Rendimiento de estudiantes de 15 años en Ciencias, Lectura y Matemática*. Santiago, Chile.
- OCDE (Organización para la Cooperación y el Desarrollo Económicos) (2006). *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*. París, Francia.
- _____ (2008). *Informe PISA 2006. Competencias científicas para el mundo del mañana*. Madrid, España: Grupo Santillana de Ediciones.
- _____ (2009). *Top of the Class: High Performers in Science in PISA 2006*. París, Francia.
- SIMCE (Sistema de Medición de la Calidad de la Educación) (2008). *Resultados Nacionales SIMCE 2007*. Santiago, Chile: MINEDUC.

Anexo

1. Distribución de preguntas por los grupos de dificultad, de acuerdo a las distintas variables, para Chile y la OCDE, en cada una de las subescalas

Tabla A.1

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el formato de la pregunta, en la escala general de ciencias

Grupo de dificultad		Chile		OCDE	
		Preguntas Abiertas	Preguntas Cerradas	Preguntas Abiertas	Preguntas Cerradas
Baja	Cantidad de preguntas	1	19	2	24
	PRC	84,06	67,14	80,56	74,01
Media	Cantidad de preguntas	11	35	11	30
	PRC	44,66	46,20	55,22	56,31
Alta	Cantidad de preguntas	22	15	21	15
	PRC	23,94	26,55	34,92	36,20
Cantidad total de preguntas		34	69	34	69

Tabla A.2

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el formato de la pregunta, en la subescala explicar fenómenos científicamente

Grupo de dificultad		Chile		OCDE	
		Preguntas Abiertas	Preguntas Cerradas	Preguntas Abiertas	Preguntas Cerradas
Baja	Cantidad de preguntas	-	11	1	13
	PRC	-	70,10	75,95	76,20
Media	Cantidad de preguntas	6	16	6	15
	PRC	44,87	45,34	53,25	56,60
Alta	Cantidad de preguntas	9	7	8	6
	PRC	22,30	24,49	33,67	35,69
Cantidad total de preguntas		15	34	15	34

Tabla A.3

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el formato de la pregunta, en la subescala *usar evidencia científica*

Grupo de dificultad		Chile		OCDE	
		Preguntas Abiertas	Preguntas Cerradas	Preguntas Abiertas	Preguntas Cerradas
Baja	Cantidad de preguntas	1	-	2	5
	PRC	84,06	-	74,45	65,91
Media	Cantidad de preguntas	5	9	6	8
	PRC	44,41	51,98	52,15	52,78
Alta	Cantidad de preguntas	9	7	7	3
	PRC	25,83	28,25	34,43	33,22
Cantidad total de preguntas		15	16	15	16

Tabla A.4

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el formato de la pregunta, en la subescala *identificar cuestiones científicas*

Grupo de dificultad		Chile		OCDE	
		Preguntas Abiertas	Preguntas Cerradas	Preguntas Abiertas	Preguntas Cerradas
Baja	Cantidad de preguntas	-	1	-	10
	PRC	-	80,26	-	70,40
Media	Cantidad de preguntas	-	13	-	8
	PRC	-	53,22	-	48,56
Alta	Cantidad de preguntas	4	5	4	1
	PRC	23,35	36,91	33,56	40,49
Cantidad total de preguntas		4	19	4	19

Tabla A.5

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el contexto de las preguntas, en la escala *general de ciencias*

Grupo de dificultad		Chile			OCDE		
		Personal	Social	Global	Personal	Social	Global
Baja	Cantidad de preguntas	7	12	1	9	16	1
	PRC	71,44	65,57	72,87	76,19	73,3	78,35
Media	Cantidad de preguntas	12	26	8	9	24	8
	PRC	45,9	46,37	43,97	57,99	55,26	56,05
Alta	Cantidad de preguntas	9	19	9	10	17	9
	PRC	24,93	27,48	19,82	37,6	38,16	31,3
Cantidad total de preguntas		28	57	18	28	57	18

Tabla A.6

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el contexto de las preguntas, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile			OCDE		
		Personal	Social	Global	Personal	Social	Global
Baja	Cantidad de preguntas	6	4	1	7	6	1
	PRC	73,77	64,04	72,87	78,13	73,59	78,34
Media	Cantidad de preguntas	5	13	4	5	12	4
	PRC	44,19	46,31	42,92	55,35	55,12	57,59
Alta	Cantidad de preguntas	6	5	5	5	4	5
	PRC	23,85	27,97	17,84	36,96	37,99	29,35
Cantidad total de preguntas		17	22	10	17	22	10

Tabla A.7

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el contexto de las preguntas, en la subescala *usar evidencia científica*

Grupo de dificultad		Chile			OCDE		
		Personal	Social	Global	Personal	Social	Global
Baja	Cantidad de preguntas	-	1	-	2	5	-
	PRC	-	84,06	-	65,22	69,6	-
Media	Cantidad de preguntas	4	7	3	3	7	4
	PRC	51,46	50,74	42,97	56,09	51,59	51,44
Alta	Cantidad de preguntas	2	10	4	1	6	3
	PRC	25,24	29,06	22,3	27,08	37,34	29,85
Cantidad total de preguntas		6	18	7	6	18	7

Tabla A.8

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el contexto de las preguntas, en la subescala *identificar cuestiones científicas*

Grupo de dificultad		Chile			OCDE		
		Personal	Social	Global	Personal	Social	Global
Baja	Cantidad de preguntas	-	1	-	1	9	-
	PRC	-	80,26	-	72,05	70,22	-
Media	Cantidad de preguntas	2	10	1	2	5	1
	PRC	53,17	53,44	51,2	50,64	45,91	57,68
Alta	Cantidad de preguntas	3	6	-	2	3	-
	PRC	35,32	28,67	-	38,02	32,91	-
Cantidad total de preguntas		5	17	1	5	17	1

Tabla A.9

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el área de aplicación de las preguntas, en la escala *general de ciencias*

Grupo de dificultad		Chile					OCDE				
		Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y Tecnología	Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y Tecnología
Baja	Cantidad de preguntas	5	3	4	4	4	6	4	5	6	5
	PRC	66,23	68,44	66,16	65,07	74,58	73,61	74,38	71,57	74,32	78,88
Media	Cantidad de preguntas	15	7	6	8	10	14	6	5	7	10
	PRC	44,25	48,76	45,74	43,91	47,74	55,82	57,63	56,31	53,89	56,64
Alta	Cantidad de preguntas	9	7	8	3	10	9	7	8	2	9
	PRC	24,33	27,26	20,71	29,41	26,11	37,07	38,46	32,76	38,39	35,79
Cantidad total de preguntas		29	17	18	15	24	29	17	18	15	24

Tabla A.10

Distribución de las preguntas en los grupos de dificultad, de acuerdo con el área de aplicación de las preguntas, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile					OCDE				
		Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y Tecnología	Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y Tecnología
Baja	Cantidad de preguntas	5	2	-	1	3	6	2	1	2	3
	PRC	66,22	71,77	-	72,23	79,09	73,6	77,06	67,59	74,09	85,09
Media	Cantidad de preguntas	3	3	3	6	7	4	4	2	5	6
	PRC	46,83	47,02	47,42	43,70	44,09	54,89	56,6	58,05	53,86	56,18
Alta	Cantidad de preguntas	4	5	4	1	2	2	4	4	1	3
	PRC	25,13	27,70	16,84	23,18	21,29	33,07	36,16	28,21	41,24	39,55
Cantidad total de preguntas		12	10	7	8	12	12	10	7	8	12

Tabla A.11
Distribución de las preguntas en los grupos de dificultad, de acuerdo con el área de aplicación de las preguntas, en la subescala *usar evidencia científica*

Grupo de dificultad		Chile					OCDE				
		Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y Tecnología	Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y Tecnología
Baja	Cantidad de preguntas	-	-	-	1	-	1	2	-	2	2
	PRC	-	-	-	84,06	-	63,74	66,52	-	75,93	64,92
Media	Cantidad de preguntas	5	2	2	2	3	6	1	3	2	2
	PRC	48,09	56,35	40,51	48,55	52,91	55,34	46,63	49,97	53,96	50,83
Alta	Cantidad de preguntas	5	2	3	1	5	3	1	2	-	4
	PRC	27,08	26,17	26,08	33,37	26,18	35,92	39,00	32,28	-	32,33
Cantidad total de preguntas		10	4	5	4	8	10	4	5	4	8

Tabla A.12
Distribución de las preguntas en los grupos de dificultad, de acuerdo con el área de aplicación de las preguntas, en la subescala *identificar cuestiones científicas*

Grupo de dificultad		Chile					OCDE				
		Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y Tecnología	Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y Tecnología
Baja	Cantidad de preguntas	-	-	1	-	-	-	1	5	2	2
	PRC	-	-	80,26	-	-	-	72,05	70,32	72,93	67,25
Media	Cantidad de preguntas	2	3	4	2	2	5	2	-	-	1
	PRC	46,28	49,78	58,89	54,14	53,03	48,2	49,82	-	-	47,86
Alta	Cantidad de preguntas	5	-	1	1	2	2	0	1	1	1
	PRC	34,68	-	20,12	31,67	26,41	35,8	0	38,91	35,54	28,71
Cantidad total de preguntas		7	3	6	3	4	7	3	6	3	4

Tabla A.13
Distribución de las preguntas en los grupos de dificultad, de acuerdo con el área de contenido de las preguntas, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile				OCDE			
		Sistemas Vivos	Sistemas Físicos	Sistemas de la Tierra y el Espacio	Sistemas Tecnológicos	Sistemas Vivos	Sistemas Físicos	Sistemas de la Tierra y el Espacio	Sistemas Tecnológicos
Baja	Cantidad de preguntas	5	3	2	1	6	4	3	1
	PRC	60,22	75,98	71,77	69,06	73,60	81,02	73,90	79,36
Media	Cantidad de preguntas	6	9	7	-	7	7	6	1
	PRC	40,40	44,38	45,26	-	55,49	55,65	56,64	50,68
Alta	Cantidad de preguntas	10	3	2	1	8	4	2	-
	PRC	24,10	24,38	13,34	31,30	34,45	38,19	27,58	-
Cantidad total de preguntas		21	15	11	2	21	15	11	2

Tabla A.14
Distribución de las preguntas en los grupos de dificultad, de acuerdo con las habilidades evaluadas en las preguntas, en la subescala *usar evidencia científica*

Grupo de dificultad		Chile			OCDE		
		Identificar supuestos, pruebas y razonamientos que subyacen a las conclusiones	Interpretar pruebas científicas y formular conclusiones	Reflexionar sobre implicancias sociales de los avances científicos y tecnológicos	Identificar supuestos, pruebas y razonamientos que subyacen a las conclusiones	Interpretar pruebas científicas y formular conclusiones	Reflexionar sobre implicancias sociales de los avances científicos y tecnológicos
Baja	Cantidad de preguntas	-	-	1	1	4	2
	PRC	-	-	84,06	63,74	66,33	74,71
Media	Cantidad de preguntas	4	8	2	3	10	1
	PRC	45,60	51,45	48,00	57,60	51,01	52,20
Alta	Cantidad de preguntas	3	13	-	3	7	-
	PRC	20,01	28,48	-	26,88	37,15	-
Cantidad total de preguntas		7	21	3	7	21	3

Tabla A.15

Distribución de las preguntas en los grupos de dificultad, de acuerdo con las habilidades evaluadas en las preguntas, en la subescala *identificar cuestiones científicas*

Grupo de dificultad		Chile			OCDE		
		Identificar términos clave para la búsqueda de información científica	Reconocer cuestiones susceptibles de ser investigadas científicamente	Reconocer los rasgos clave de la investigación científica	Identificar términos clave para la búsqueda de información científica	Reconocer cuestiones susceptibles de ser investigadas científicamente	Reconocer los rasgos clave de la investigación científica
Baja	Cantidad de preguntas	1	-	-	3	4	3
	PRC	80,26	-	-	74,41	68,45	68,99
Media	Cantidad de preguntas	2	7	4	-	4	4
	PRC	59,83	51,46	52,99	-	53,52	43,61
Alta	Cantidad de preguntas	-	3	6	-	2	3
	PRC	-	23,86	34,40	-	29,91	38,31
Cantidad total de preguntas		3	10	10	3	10	10

2. Distribución de las preguntas por grupos de dificultad, según género, para Chile y la OCDE, en la subescala “explicar fenómenos científicamente”.

Tabla A. 16

Distribución de las preguntas en los grupos de dificultad para mujeres, de acuerdo con el formato de la pregunta, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile		OCDE	
		Preguntas Abiertas	Preguntas Cerradas	Preguntas Abiertas	Preguntas Cerradas
Baja	Cantidad de preguntas	-	6	2	13
	PRC	-	75,07	69,42	75,90
Media	Cantidad de preguntas	3	19	5	16
	PRC	45,64	46,98	49,54	53,78
Alta	Cantidad de preguntas	12	9	8	5
	PRC	22,14	28,86	32,14	31,37
Cantidad total de preguntas		15	34	15	34

Tabla A.17

Distribución de las preguntas en los grupos de dificultad para hombres, de acuerdo con el formato de la pregunta, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile		OCDE	
		Preguntas Abiertas	Preguntas Cerradas	Preguntas Abiertas	Preguntas Cerradas
Baja	Cantidad de preguntas	1	12	1	16
	PRC	60,36	72,28	78,00	74,58
Media	Cantidad de preguntas	5	16	10	16
	PRC	45,99	46,55	50,53	52,96
Alta	Cantidad de preguntas	9	6	4	2
	PRC	24,93	24,89	25,45	31,03
Cantidad total de preguntas		15	34	15	34

Tabla A.18

Distribución de las preguntas en los grupos de dificultad para mujeres, de acuerdo con el contexto en que se plantean las preguntas, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile			OCDE		
		Personal	Social	Global	Personal	Social	Global
Baja	Cantidad de preguntas	3	2	1	7	7	1
	PRC	82,50	64,89	73,21	77,88	71,63	78,92
Media	Cantidad de preguntas	7	13	2	5	12	4
	PRC	48,73	45,83	46,26	52,88	51,58	56,25
Alta	Cantidad de preguntas	7	7	7	5	3	5
	PRC	22,63	26,51	20,21	34,64	36,02	26,54
Cantidad total de preguntas		17	22	10	17	22	10

Tabla A.19

Distribución de las preguntas en los grupos de dificultad para hombres, de acuerdo con el contexto en que se plantean las preguntas, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile			OCDE		
		Personal	Social	Global	Personal	Social	Global
Baja	Cantidad de preguntas	7	5	1	8	8	1
	PRC	75,28	65,63	72,59	77,2	71,99	77,81
Media	Cantidad de preguntas	4	13	4	8	12	6
	PRC	43,59	47,41	46,03	49,08	53,30	53,4
Alta	Cantidad de preguntas	6	4	5	1	2	3
	PRC	26,12	29,46	19,82	23,45	32,44	25,17
Cantidad total de preguntas		17	22	10	17	22	10

Tabla A.20

Distribución de las preguntas en los grupos de dificultad para mujeres, de acuerdo con el área de aplicación en que se plantean las preguntas, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile					OCDE				
		Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y la Tecnología	Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y la Tecnología
Baja	Cantidad de preguntas	2	2	-	-	2	7	2	1	2	3
	PRC	71,73	69,50	-	-	84,01	72,38	76,77	66,90	71,38	85,23
Media	Cantidad de preguntas	6	3	2	5	6	3	5	2	5	6
	PRC	51,43	42,37	51,49	44,14	45,00	51,49	53,14	56,28	51,08	53,36
Alta	Cantidad de preguntas	4	5	5	3	4	2	3	4	1	3
	PRC	23,41	24,90	17,68	26,77	24,67	32,02	31,73	26,79	38,42	36,40
Cantidad total de preguntas		12	10	7	8	12	12	10	7	8	12

Tabla A.21

Distribución de las preguntas en los grupos de dificultad para hombres, de acuerdo con el área de aplicación en que se plantean las preguntas, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile					OCDE				
		Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y la Tecnología	Salud	Recursos Naturales	Medio Ambiente	Riesgos	Fronteras de la Ciencia y la Tecnología
Baja	Cantidad de preguntas	5	2	-	2	4	6	4	1	2	4
	PRC	69,81	73,68	-	61,03	77,31	73,59	71,05	68,27	76,72	80,98
Media	Cantidad de preguntas	3	3	4	5	6	5	4	3	6	8
	PRC	46,46	51,00	45,96	45,08	45,53	52,91	46,53	54,16	54,49	51,57
Alta	Cantidad de preguntas	4	5	3	1	2	1	2	3	-	-
	PRC	26,64	30,15	13,23	26,55	25,08	23,45	32,44	25,17	-	-
Cantidad total de preguntas		12	10	7	8	12	12	10	7	8	12

Tabla A.22

Distribución de las preguntas en los grupos de dificultad para mujeres, de acuerdo con el área de contenido que evalúan las preguntas, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile				OCDE			
		Sistemas Vivos	Sistemas Físicos	Sistemas de la Tierra y el Espacio	Sistemas Tecnológicos	Sistemas Vivos	Sistemas Físicos	Sistemas de la Tierra y el Espacio	Sistemas Tecnológicos
Baja	Cantidad de preguntas	2	2	2	-	7	4	3	1
	PRC	71,73	84,01	69,50	-	72,38	80,36	73,48	77,00
Media	Cantidad de preguntas	8	8	5	1	7	7	6	1
	PRC	49,99	43,82	43,61	60,81	51,89	52,74	54,35	49,74
Alta	Cantidad de preguntas	11	5	4	1	7	4	2	-
	PRC	22,52	24,66	21,83	27,15	32,11	34,08	26,46	-
Cantidad total de preguntas		21	15	11	2	21	15	11	2

Tabla A.23

Distribución de las preguntas en los grupos de dificultad para hombres, de acuerdo con el área de contenido que evalúan las preguntas, en la subescala *explicar fenómenos científicamente*

Grupo de dificultad		Chile				OCDE			
		Sistemas Vivos	Sistemas Físicos	Sistemas de la Tierra y el Espacio	Sistemas Tecnológicos	Sistemas Vivos	Sistemas Físicos	Sistemas de la Tierra y el Espacio	Sistemas Tecnológicos
Baja	Cantidad de preguntas	5	4	3	1	7	5	4	1
	PRC	69,81	72,53	70,73	76,33	72,23	78,39	73,00	81,75
Media	Cantidad de preguntas	7	8	6	-	10	10	5	1
	PRC	46,41	47	45,63	-	50,39	51,27	56,89	51,56
Alta	Cantidad de preguntas	9	3	2	1	4	-	2	-
	PRC	25,14	27,68	14,74	34,95	26,34	-	28,66	-
Cantidad total de preguntas		21	15	11	2	21	15	11	2

¿Qué nos dice PISA sobre la educación de los jóvenes en Chile?

Análisis comparativo de los resultados chilenos en las pruebas de Matemática SIMCE y PISA

Por
Luis Alfaro
Raúl Gormaz¹

Resumen

En este trabajo se presenta un análisis de los resultados para Chile de la prueba PISA Matemática 2006, la cual se compara con la prueba SIMCE Matemática de 2° Medio aplicada el mismo año. El objetivo de investigación consistió en entender qué elementos de la prueba PISA la hacen difícil para los estudiantes chilenos, lo cual se espera sea un aporte para orientar los esfuerzos para mejorar el aprendizaje de Matemática en Chile.

La unidad de observación fueron los ítemes de Matemática de ambas pruebas, y sus respectivos porcentajes de respuesta correcta. El análisis consistió primeramente, en identificar la presencia o ausencia de conocimientos y habilidades propias del Marco Curricular Chileno en las preguntas de ambas pruebas, y determinar cuán difíciles resultaban estas preguntas. En una segunda etapa, se analizaron algunas características de las preguntas de PISA que pudieran explicar el bajo rendimiento de los estudiantes chilenos.

Del análisis, se obtuvo que los contenidos evaluados por PISA están considerados en el Marco Curricular Chileno, por lo que la dificultad de sus preguntas no radicaría en una falta de alineación.

Por otra parte, se distinguieron en las preguntas de PISA algunos elementos de forma que podrían aumentar la dificultad para responder por parte de los estudiantes chilenos. Estos elementos se clasificaron principalmente como: la resolución de problemas en situaciones novedosas, el trabajo simultáneo con números de variados ámbitos numéricos y el trabajo con razonamientos de mayor demanda cognitiva. También pareciera estar afectando, el trabajo con información visual y, en particular, recurrir a situaciones reales.

A partir del análisis, también se pudo determinar que la profundidad en el manejo de conceptos matemáticos representa un factor importante en la dificultad que reviste PISA para los estudiantes chilenos. De ello, se deriva la necesidad de incorporar en el trabajo de aula, la resolución de problemas matemáticos de manera más profunda y sistemática.

¹ Luis Alfaro y Raúl Gormaz son miembros del Equipo Disciplinario de Matemáticas del SIMCE.

Introducción

El énfasis de la prueba PISA está puesto en la comprensión de ciertos contenidos y/o conceptos matemáticos, y su aplicación a situaciones reales, y no en el conocimiento detallado de los mismos. La prueba PISA declara evaluar la Competencia Matemática², que define como:

Una capacidad del individuo para identificar y entender la función que desempeñan las matemáticas en el mundo, emitir juicios fundados y utilizar y relacionarse con las matemáticas de forma que se puedan satisfacer las necesidades de la vida de los individuos, como ciudadanos constructivos, comprometidos y reflexivos (OCDE, 2006a, p.74).

Al revisar los resultados de la participación de Chile en la prueba PISA, sección matemática, aplicada el año 2006 (en adelante “PISA”), en relación a la competencia matemática, lo que se observa inmediatamente es que el rendimiento de Chile está por debajo de la mayoría de los países que rindió esta prueba. Específicamente, Chile ocupa algún lugar³ entre el 43 y el 48 del ranking de los 57 países que participaron, es decir, se sitúa en el último 20% del ranking de los países evaluados. A nivel local, está entre el 2^{do} y 3^{er} lugar en comparación con los 6 países latinoamericanos que participaron, lo que contrasta con los buenos resultados en Ciencias y Lectura, en los que Chile destaca a nivel latinoamericano. Lo anterior implica que a los estudiantes chilenos les resultó más difícil la prueba PISA Matemática que a los estudiantes de la mayor parte de los otros países.

2 Para mayores detalles del constructo evaluado, se puede consultar el Marco de Evaluación de la prueba PISA (OCDE, 2006a), donde se describen, entre otros aspectos, los contenidos Matemáticos (espacio y forma, cambio y relaciones, cantidad e incertidumbre), las capacidades (de reproducción, de conexiones y de reflexión) y las situaciones o contextos (personal, educacional/profesional, público y científico).

3 El error estadístico propio de toda medición impide discernir con mayor precisión el ranking de Chile en esta prueba.

Adicionalmente, considerando la muestra de estudiantes chilenos de 2° Medio que rindió el año 2006 tanto PISA como SIMCE 2° Medio Matemática (en adelante “SIMCE”), se pudo establecer que las preguntas PISA resultan más difíciles que las de la prueba SIMCE (véase el detalle de esta información, en el Anexo).

En efecto, para la muestra de estudiantes chilenos, las preguntas SIMCE mostraron una distribución bastante simétrica en torno a una dificultad de 50%, siendo su mediana⁴ 48%. Las preguntas de la prueba PISA, en cambio, se concentraron mayoritariamente en el primer tramo (menores porcentajes de respuestas correctas, por ende, mayor dificultad), y tuvieron una mediana de 30%. Asimismo, se pudo observar que 35% de las preguntas PISA fueron respondidas correctamente solo por menos de 20% de los estudiantes chilenos, lo que contrasta con SIMCE, donde las preguntas que respondieron correctamente menos de 20% de los estudiantes son solo 2 de las 64 preguntas (correspondientes a 3%). En este sentido, cabe destacar que, puesto que la prueba PISA es una evaluación internacional, lo que se busca es que la dificultad media corresponda al logro medio de la comunidad evaluada (los países de la OCDE y no Chile). De allí que, para la muestra total PISA, la mediana haya sido de 47%. En otras palabras, los alumnos chilenos rindieron una prueba diseñada para un grupo que, en promedio, muestra mayor desempeño que ellos.

Considerando todo lo anteriormente expuesto, cabe preguntarse entonces: ¿En qué radica la mayor dificultad que presentaron los estudiantes chilenos en la prueba PISA, tanto comparados con los demás países que la rindieron, como con los logros que demuestran en la evaluación nacional SIMCE?

La primera hipótesis explicativa, acerca de la mayor dificultad de PISA en comparación con SIMCE, fue que la primera evalúa contenidos temáticos y

4 La mediana usada aquí es la mediana de los PRC de los ítems de cada prueba.

habilidades no cubiertos por el Marco Curricular Nacional, en cambio SIMCE evalúa exclusivamente el Marco Curricular Nacional. De esta forma, aquellos contenidos temáticos y habilidades no cubiertos por el Marco Curricular y que son requeridos para responder adecuadamente una pregunta PISA, serían la principal causa de su dificultad.

Una segunda hipótesis explicativa consistió en que el tipo de preguntas al que están habituados los estudiantes chilenos difiere del encontrado en PISA. En este contexto, se compararon los énfasis observados en las evaluaciones de aula, en SIMCE y en PISA, tanto en contenidos temáticos, como en categorías de habilidades. Un análisis adicional en relación a esta hipótesis explicativa consistió en describir elementos o aspectos presentes en las preguntas de PISA, que pudiesen aparecer con mayor o menor frecuencia, o con una mayor o menor intensidad, en las preguntas SIMCE, elementos que comúnmente se asocian a una mayor dificultad.

A continuación, se describe la metodología utilizada en este estudio; luego se presentan los resultados obtenidos del análisis comparativo de las pruebas SIMCE y PISA, entre sí y con el Marco Curricular chileno, y finalmente, se presentan conclusiones y recomendaciones generales para orientar la enseñanza de la Matemática en los establecimientos educacionales de Chile, así como preguntas orientadoras de futuras investigaciones.

1. Metodología

Como se señaló, una primera hipótesis de la dificultad de los ítems PISA para los estudiantes chilenos, se basa en la diferencia que existiría entre lo que evalúan los ítems PISA y lo que se exige en el Marco Curricular. Más precisamente, si para responder un ítem de PISA se requiere el dominio de contenidos temáticos y/o habilidades no contemplados en el Marco Curricular chileno, al menos hasta el nivel de 2° Medio, tendríamos



aquí un elemento que explicaría la mayor dificultad que reporta el ítem para los estudiantes chilenos.

Para confirmar o refutar esta primera hipótesis, fue necesario analizar cada ítem PISA, determinar los contenidos temáticos y habilidades que requería para su resolución, y determinar si estos contenidos temáticos y habilidades estaban presentes en el Marco Curricular chileno, al menos entre 1° Básico y 2° Medio.

Para cumplir con este propósito, se utilizó una adaptación del “Marco Curricular Analítico” (MCA), preparado por el Consorcio de Evaluación e Investigación Evaluativa (CEIE) de República Dominicana, bajo la dirección de G. A. Valverde (EERC, 2005)⁵ (véase esta adaptación en el Anexo). Con este MCA se analizó cada ítem de PISA, asociándole uno o más contenidos temáticos y expectativas de desempeño, los que a su vez, fueron contrastados con el Marco Curricular chileno, para determinar (en los casos en que fue posible) el nivel escolar en que se señala que deben ser adquiridos por los estudiantes.

⁵ El que, a su vez, corresponde a una extensión y revisión del utilizado por el *Third International Mathematics and Science Study* (Robitaille y otros, 1993). El uso del MCA, en lugar del Marco Curricular (MC) se consideró necesario por dos motivos: En la medida que el MC no es exhaustivo (hay contenidos y/o habilidades que no aparecen mencionados explícitamente en él), resulta necesario disponer de una lista que garantice la presencia de todo contenido/habilidad que pueda demandar un ítem PISA. Por otro lado, es práctico que cada contenido y/o habilidad aparezca solo una vez en el MCA. Esto no ocurre en el Marco Curricular chileno, pues, por ejemplo, los contenidos temáticos pueden aparecer en distintos niveles educacionales.

Con el objeto de contrastar la segunda hipótesis de este estudio, se establecieron algunos puntos de comparación entre PISA, SIMCE e información sobre cómo se evalúa en las salas de clases. Esta última, se obtuvo del análisis de una muestra de evaluaciones de 19 establecimientos (representativa de las diferentes regiones, rendimientos SIMCE y tipos de dependencia), en seis niveles escolares (2°, 4°, 6° y 8° básicos; y 2° y 4° años de Enseñanza Media)⁶.

Con estos tres antecedentes (PISA, SIMCE y evaluaciones de aula), se pudo obtener la distribución de ítemes según los cuatro ejes temáticos (Números, Geometría, Álgebra, y Datos y Azar). Esto permitió comparar los énfasis con que cada uno de estos temas es considerado en las evaluaciones, y establecer algunas diferencias. En el caso de PISA y SIMCE, además se pudo identificar las dificultades reportadas por eje, lo que entregó una dimensión de comparación adicional.

Para establecer diferencias en la distribución de ítemes según habilidades, fue necesario establecer una taxonomía especial, por cuanto en las tres evaluaciones analizadas se utilizan distintos criterios. En PISA se distingue Reproducción, Aplicación y Reflexión (OCDE, 2006a); en SIMCE, Conocimiento, Resolución de problemas y Razonamiento matemático, y, en las evaluaciones de aula, se utiliza la clasificación de TIMSS (Mullis y otros, 2003), a saber: Manejar conocimientos y procedimientos, Usar conceptos, Resolver problemas de rutina, y Razonar. Dada esta diversidad de taxonomías, en este estudio se establecieron dos categorías: una primera que considera Reproducción (PISA), Conocimiento (SIMCE), Manejar Conocimientos y Procedimientos/Usar Conceptos (Aula, TIMSS) y una segunda categoría con el resto, esto es, Aplicación y Reflexión (PISA), Resolución de Problemas y Razonamiento Matemático (SIMCE) y Resolver Problemas de Rutina y Razonar (Aula, TIMSS).

6 El análisis de la muestra se obtuvo del estudio *Evaluación de aula en enseñanza básica y media. Subsector Educación Matemática* (MINEDUC, 2009).

Finalmente, con el objeto de descubrir algunos posibles elementos propios de las preguntas PISA que pudieran relacionarse con su dificultad, se analizaron las distintas orientaciones de la prueba PISA, declaradas en su Marco de Evaluación, y se consideraron las indicaciones de estudios sobre elementos explicativos de dificultad, como los de Grønmo y Olsen, (2006) y Nohara (2001). En este contexto, se determinó analizar los siguientes elementos, que se asocian tradicionalmente a una mayor dificultad:

- Situaciones reales
- Información visual
- Razonamiento de múltiples pasos
- Situaciones novedosas
- Cálculos numéricos
- Pensamiento reversible

Las **situaciones reales**, se refieren a aquellas centradas en los problemas del mundo real, por tanto su uso permite conocer la capacidad del estudiante de aplicar un determinado conocimiento con efectividad, para resolver una situación problemática, del tipo de las que se encuentran en el mundo real⁷. Se busca que las matemáticas se empleen para resolver el problema planteado, y que tenga sentido hacerlo. El uso de situaciones reales en una pregunta tiene el efecto de hacerla más interesante (pues invita a pensar aspectos del mundo que posiblemente no conocemos) y a la vez, desafiante, principalmente por la novedad.

Esta característica se ilustra en el ejemplo de la prueba PISA de la Figura 1.


7 PISA utiliza también el término "situaciones auténticas". Esto se hace para incluir, además de situaciones reales, situaciones inventadas, pero con elementos que se rescatan de una problemática real.

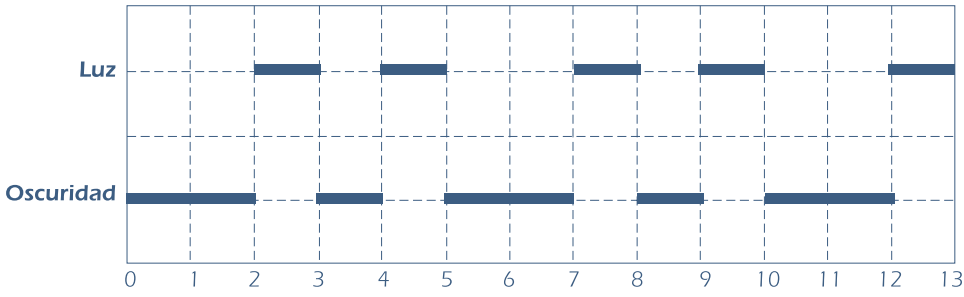
Figura 1.
Ejemplo de pregunta de Matemática PISA

Los faros son torres que disponen de un foco luminoso en su parte superior. Los faros ayudan a los barcos a seguir su rumbo de noche cuando navegan cerca de la costa.

Un faro emite destellos luminosos según una secuencia regular fija. Cada faro tiene su propia secuencia.

En el diagrama que figura a continuación puede verse la secuencia de un determinado faro. Los destellos de luz alternan con períodos de oscuridad.

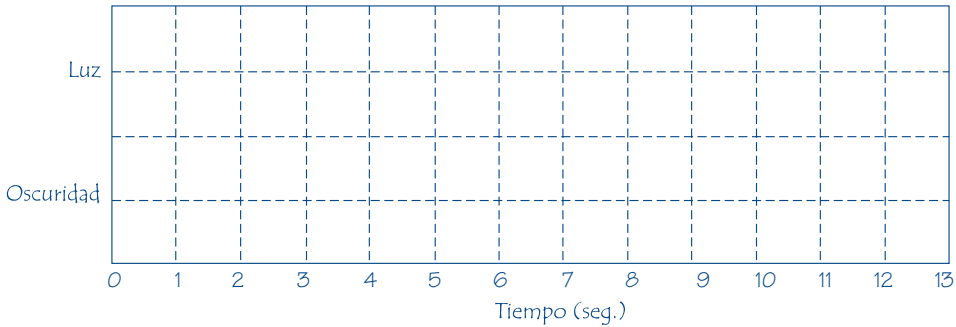




Se trata de una secuencia regular. Al cabo de un tiempo, la secuencia se vuelve a repetir. El tiempo que tarda en completarse un ciclo completo, antes de volver a iniciarse la secuencia, recibe el nombre de período. Una vez que se ha hallado el período de la secuencia es fácil ampliar el diagrama para los siguientes segundos, minutos e incluso horas.

Pregunta 1: EL FARO

Realiza un gráfico en el diagrama de abajo, indicando una secuencia posible de destellos de un faro que emita destellos durante 30 segundos por minuto. El período de esta secuencia debe ser igual a 6 segundos.



Pauta de Corrección	
Puntaje Completo	El estudiante grafica 3 segundos de destellos luminosos en 6 segundos. Por ejemplo, concentrado en 3 segundos de luz, bien en dos segundos de luz, seguido por otro destello de 1 segundo de luz. Si aparecen graficados más de 6 segundos, este esquema debe repetirse, de la misma forma, en cada tramo de 6 segundos.
Puntaje Parcial	El estudiante grafica un esquema de tres destellos de 1 segundo de luz, separados cada uno por un segundo de oscuridad.
Sin Puntaje	Otras respuestas u omitida.

Fuente: OCDE, 2006a

En este ítem, se pide a los estudiantes que elaboren una secuencia posible de destellos de luz. Se trata de una problemática real en el mundo de los faros, y que para la mayoría de los estudiantes no será familiar. Muchos estudiantes no habrán nunca enfrentado en el colegio una pregunta similar, por lo que se trata de una situación novedosa. Al estudiante se le entrega la iniciativa para proponer una nueva solución, es decir, una nueva secuencia de destellos. Este aspecto propositivo⁸ de la respuesta es esencial para poder decir que alguien es competente en matemáticas.

Habrá quien argumente que esta pregunta puede favorecer a los estudiantes que residan cerca del mar. A este respecto, no debe olvidarse que la competencia matemática conlleva la capacidad de utilizar las matemáticas en contextos distintos de los del ámbito local propio. Esta capacidad de transferencia es un elemento esencial de dicha competencia. La segunda característica, que se ha denominado **información visual**, se refiere al uso de información en forma de tablas, gráficos, diagramas y/o esquemas.

Un ejemplo de uso de información visual, se ilustra en la Figura 2. Se trata de planificar la mejor ruta para unas vacaciones.

8 PISA exige que las capacidades matemáticas no se empleen solo de una manera pasiva o siguiendo instrucciones, sino demostrando la capacidad de proponer y construir respuestas. La solución de este problema requiere que se satisfagan dos condiciones: las mismas cantidades de tiempo de luz y oscuridad (30 segundos por minuto) y un período de 6 segundos. Debido a esa combinación, es esencial que los estudiantes demuestren un manejo pleno del concepto de periodicidad. El estudiante debe saber cómo representar su solución en un gráfico, reconocer que en 6 segundos debe tener 3 segundos de destellos de luz, y que este esquema de destellos se debe repetir cada 6 segundos, para satisfacer la condición sobre el período. Los elementos que están en juego son cada uno muy simple, pero la dificultad consiste en ponerlos todos a interactuar. No hay un método de resolución que el estudiante aplique mecánicamente, más bien el estudiante debe organizar un método de solución propio.

Figura 2.
Ejemplo de pregunta de Matemática PISA

En las Figuras 1 y 2 se muestran un mapa de la zona y las distancias entre las distintas ciudades.

FIGURA 1: MAPA DE LAS CARRETERAS QUE HAY ENTRE LAS CIUDADES.

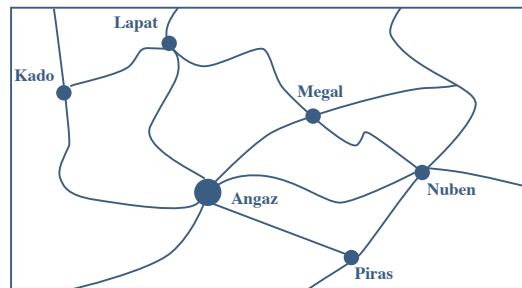


FIGURA 2: DISTANCIAS MÁS CORTAS ENTRE LAS CIUDADES EN KILÓMETROS.

Angaz						
Kado	550					
Lapat	500	300				
Megal	300	850	550			
Nuben	500		1.000	450		
Piras	300	850	800	600	250	
	Angaz	Kado	Lapat	Megal	Nuben	Piras

Calcula la distancia más corta por carretera, entre Nuben y Kado.

Distancia: kilómetros.

Pauta de Corrección

Puntaje Completo	El estudiante escribe 1050.
Sin Puntaje	Otras respuestas u omitida.

Fuente: OCDE, 2006a

En este ítem, la información se reparte entre 2 elementos visuales: un mapa simplificado y una tabla de distancias mínimas. La distancia solicitada (la más corta entre Nuben y Kado) no aparece en la tabla, por lo que debe ser deducida a partir del resto de la

información disponible. La estrategia de solución no es evidente, y requiere entender, pero también saber aprovechar, que las distancias indicadas en la tabla son las mínimas entre cada par de lugares.

En cuanto a la tercera característica, el **razonamiento de múltiples pasos**, cabe señalar que esta se refiere a que, si bien la información disponible para resolver el problema está en el enunciado, la resolución requiere que se transforme hacia una forma intermedia, desde donde la resolución se hace posible. En estos problemas, la información del enunciado no permite obtener la respuesta en forma inmediata: los datos pueden servir para generar nueva información, que permita obtener la solución. Estos datos intermedios, aunque no son solicitados, deben ser obtenidos para poder resolver el problema.

El ejemplo anterior también permite ilustrar esta característica. En efecto, un primer análisis muestra que todas las posibles rutas para llegar desde Nuben a Kado (en el plano), pasan a través de Megal, Angaz o Piras. En esta etapa, se puede generar información intermedia, que puede servir para responder. Se puede descubrir, por ejemplo, que para ir de Nuben a Kado hay tres alternativas: (a) ir de Nuben a Megal y desde Megal, a Kado; (b) ir de Nuben a Angaz y desde Angaz, a Kado, o (c) ir de Nuben a Piras y desde Piras, a Kado (aunque una mirada al mapa muestra que esta solución es mala, pues desde allí solo se puede pasar hacia Angaz, y la solución (b) será claramente superior). Aquí se debe trabajar con esta información intermedia y calcular las 3 sumas correspondientes a cada par de trayectos que llevan desde Nuben a Kado, trayectos todos, que están en la tabla. A partir de esta información, el problema se resuelve con una segunda etapa muy simple, pues la menor suma es la distancia mínima solicitada: $1050 = 500$ (menor distancia desde Nuben a Angaz) + 550 (menor distancia desde Angaz a Kado).

Una cuarta característica es la de las **situaciones novedosas** (en el primer ejemplo del faro, algo se

señaló respecto de esta característica). Las situaciones novedosas se refieren a aquellas que el estudiante enfrenta por primera vez, sin tener como referencia otros problemas similares que le permitan resolverlas de manera análoga. En el análisis de este estudio, para cada ítem se determinó si la situación presentada se consideraba novedosa para el estudiante. Aunque esto es claramente dependiente de cada estudiante, fue posible guiarse por lo que dice la tradición de la enseñanza de ciertos temas, y por lo que usualmente se encuentra en los libros de texto. En estos, el estudiante se familiariza con ciertas situaciones que son típicamente usadas para ilustrar un concepto matemático o la aplicación de un concepto. Ejemplos de esto son las situaciones de compra y venta en el uso de los números y las operaciones, y los juegos de lanzamiento de dados y monedas en relación a las probabilidades. Lo novedoso de un problema puede corresponder a pequeñas variantes de una situación ya conocida y trabajada, hasta situaciones nunca antes estudiadas, y por lo tanto, enfrentadas por primera vez.

En cuanto a la dificultad por **cálculos numéricos**, se puede señalar que esta característica se vincula a la contextualización de los problemas en situaciones reales. En efecto, un contexto real no requiere generalmente realizar cálculos con números “simples”, por ejemplo, operar con números enteros del 1 al 100. Es probable que aparezcan números grandes (expresados en miles de millones), cantidades con decimales o fracciones, todos los cuales requieren un manejo eficiente de los sistemas numéricos, tener un claro sentido de la cantidad y de los órdenes de magnitud, ser capaces de estimar o aproximar un resultado, ser capaces de pasar de una representación a otra, etc. Los ítems que se consideraron que presentan esta característica fueron aquellos que requerían operar con números grandes (mayor que los millones), usar decimales o usar representación fraccionaria.

En la Figura 3, se muestra una pregunta que tiene dificultad por cálculos numéricos.

Figura 3.
Ejemplo de pregunta de Matemática PISA

VUELO ESPACIAL

La estación espacial Mir permaneció en órbita 15 años y durante este tiempo dio alrededor de 86.500 vueltas a la Tierra.

La permanencia más larga de un astronauta en la Mir fue de 680 días.

PREGUNTA 32: VUELO ESPACIAL

La Mir daba vueltas alrededor de la Tierra a una altura aproximada de 400 kilómetros. El diámetro de la Tierra mide aproximadamente 12.700 km y su circunferencia es de alrededor de 40.000 km ($\pi \times 12.700$).

Calcula aproximadamente la distancia total recorrida por la Mir durante sus 86.500 vueltas mientras estuvo en órbita. Redondea el resultado a las decenas de millón.

Pauta de Corrección	
Puntaje Completo	Los cálculos deben ser hechos correctamente, con la aproximación pedida, aceptándose como válido un resultado entre 3.600 y 3.800 millones de Km.
Puntaje Parcial	Solo presenta un error de procedimiento en el cálculo, por ejemplo: Usar radio en lugar de diámetro, en la fórmula del perímetro. Sumar 400 Km en lugar de 800 Km, para obtener el diámetro de la órbita. No redondear a la decena de millón, como se solicita, sino por ejemplo, redondear al millón.
Sin Puntaje	Otras respuestas u omitida.

Fuente: ISEI-IVEI, 2005

El conocimiento matemático involucrado es aquí el cálculo del perímetro de una circunferencia. No es necesario recordar la fórmula, pues ésta se muestra indirectamente, al incluirse en el enunciado el cálculo del perímetro de la Tierra. La solución resulta de calcular el diámetro de la órbita del satélite, multiplicar el perímetro por 86.500 y aproximar a la unidad solicitada (decena de millón). Se trata de una matemática muy simple, aplicada a una situación real y novedosa. Pero llegar a buen destino con los cálculos, no resulta simple, justamente porque se debe operar con números grandes.

La última de las características que se analiza en este estudio, es el **pensamiento reversible**, que consiste en preguntar cómo se podría obtener tal respuesta (y por lo tanto, cuáles deberían ser los datos de origen para obtenerla). Una mayor complejidad corresponde a preguntar si un tal resultado es un resultado posible. En estos ítems, no se pregunta ¿cuál es el resultado para estos datos? (problema directo), sino ¿para qué grupo de datos se obtiene tal resultado? (problema inverso).

Un ejemplo de un ítem que requiere de pensamiento reversible se observa en la Figura 4.

Figura 4.
Ejemplo de pregunta de Matemática PISA⁹

PRESA-DEPREDADOR

En el gráfico que viene a continuación se muestra el crecimiento de dos organismos vivos: el Paramecium y el Saccharomyces.

Modelo Presa-Depredador

Días	Paramecium aurelia	Saccharomyces exiguus
0	90	160
1	180	40
2	120	20
3	60	10
4	20	30
5	10	60
6	50	120
7	130	110
8	70	50
9	40	20
10	20	10
11	10	30
12	20	70
13	30	130
14	40	130
15	80	50
16	170	20
17	90	30
18	30	20

Pregunta 1: PRESA-DEPREDADOR

Uno de los dos animales (el depredador) se come al otro (la presa) ¿Permite el gráfico identificar cuál es la presa y cuál, el depredador?

Una propiedad del fenómeno presa-depredador se puede expresar de la siguiente manera: la tasa de crecimiento de los depredadores es proporcional a la cantidad de presas disponibles ¿Es aplicable esta propiedad al gráfico anterior?

Fuente: OCDE, 2006a

En este ejemplo, a pesar de que se presenta bastante información, no es de respuesta evidente. El estudiante debe establecer ciertas relaciones entre las dos curvas. Por ejemplo, si hay muchos depredadores, las presas disminuyen (son comidas por los depredadores), y si hay pocas presas, los depredadores disminuyen (por falta de alimento). Una vez que se logra comprender esta relación, identificándola en el gráfico, se puede identificar quién es el depredador.

Es el hecho de partir desde los resultados y determinar desde dónde se obtuvieron, lo que convierte este ítem en uno de pensamiento reversible, pues se debe

determinar el dato (¿qué curva corresponde a la presa y cuál al depredador?), para explicar el resultado, en este caso, las curvas de población.

Otro ejemplo, sobre pensamiento reversible, es el que se muestra en la Figura 5.

Figura 5.
Ejemplo de pregunta de Matemática PISA

Pregunta 1: MEDIA DE EDAD

Si el 40% de la población de un país tiene al menos 60 años, ¿es posible que la media de edad sea de 30 años?

Fuente: OCDE, 2006a

⁹ Para este ítem y el siguiente no se dispone de la pauta de corrección.

Tabla 1.
Distribución de ítems PISA por nivel escolar y PRC comparado

Nivel (Marco Curricular)	Ítems (cantidad)	Ítems (porcentaje)	Promedio de PRC (Chile)	Promedio de PRC (Internacional)
Primer Ciclo Básico	5	10%	48%	62%
Segundo Ciclo Básico	30	63%	28%	46%
Primero Medio	6	13%	36%	49%
Segundo Medio	6	13%	46%	60%
Tercero Medio	1	2%	31%	47%
TOTAL	48	100%		

Para el ítem de la Figura 5 no se dispone de la pauta de puntuación, pero la respuesta correcta es: “Sí, es (matemáticamente) posible que la media sea de 30 años”. De paso, al analizar cómo es esto posible, se descubre que, si para el 60% restante de la población la media fuera 10 años, la media de edad resulta ser de 30 años. En este ejemplo, vemos que al tratar de explicar un dato, nos encontramos con información nueva y sorprendente (el 60% de la población tiene promedio de edad de ¿10 años o menos?). Esto llama a seguir reflexionando.

Aquí el estudiante, como en el ejemplo anterior, a pesar de tener un resultado, debe centrar su intención en los datos que llevan a ese resultado, ver qué condiciones cumplen y si le permiten respaldar nuevos valores.

En este estudio, usando estas seis características propuestas, se clasificaron los ítems de PISA y de SIMCE, según si cada ítem presentaba o no la característica.

2. Resultados

2.1 Análisis de los ítems de la prueba PISA, en función del Marco Curricular Analítico (MCA)

El análisis de los ítems de la prueba PISA, a la luz del MCA arrojó como resultado que solo 1 de los 48 ítems presenta contenidos temáticos que no corresponden a ninguno de los niveles escolares cursados por los estudiantes chilenos que rinden la prueba (véase la Tabla 1): se trata de un ítem que alude a contenidos temáticos de 3° Medio y que en Chile presentó un 30,7% de respuestas correctas.

Ahora bien, pese a que no se observaron diferencias en las habilidades y contenidos temáticos evaluados, es posible apreciar que la mayor parte de los ítems (73%), corresponde a aprendizajes de la Enseñanza Básica, al menos en lo que respecta a su temática. Esto no parece concordar con el hecho de que PISA evalúa a estudiantes que cursan principalmente 2° Medio (10° grado), pero apoya la idea de que esta evaluación no se centra en el conocimiento matemático más avanzado. En efecto, solo 27% de los ítems corresponden a aprendizajes de Enseñanza Media, lo que lleva a afirmar que PISA más bien promueve un uso en profundidad del conocimiento matemático general.

Aunque este hecho no se contrapone con el Marco Curricular chileno, el cual también promociona este énfasis en la profundización¹⁰, sí se puede afirmar que el currículum prescrito es distinto al implementado: baste como evidencia, lo señalado en el estudio sobre evaluación del aula del MINEDUC (2009).

Por otra parte, si se comparan las dificultades observadas en cada categoría, excluyendo el ítem que hace referencia a temas de 3° Medio, se observa que los ítems de 2° Medio y de Primer Ciclo Básico, muestran menor dificultad que el resto. Esto último no sorprende, por cuanto se trata de los años en que se aprende la base de los números, las operaciones y la geometría, base que estará en permanente uso durante el resto de la enseñanza escolar. Por su parte, la mayor dificultad que se observa en el Segundo Ciclo Básico (5° a 8° Básico), que además corresponde al ciclo con mayor concentración de ítems, sustenta la idea de que las preguntas de PISA se enfocan en la evaluación del uso de cuatro ideas clave, a saber: espacio y forma, cambio y relaciones, cantidad, e incertidumbre (ver OCDE, 2006a, p.83), para las que se exige una mayor profundización. Así, entre dos opciones que pueden orientar la formación matemática: una enseñanza enciclopédica que despliega una multitud de resultados, fórmulas, propiedades, métodos, pero sin profundidad; en contraste con una enseñanza de unas pocas ideas clave, pero priorizando la profundización (esto es, una misma idea entendida desde muchas perspectivas y conectada con otras ideas), la prueba PISA parece apostar por la segunda alternativa.

En efecto, PISA declara que “no excluye los conocimientos y la comprensión basados en el currículo, pero los evalúa en función de la adquisición de unos conceptos y habilidades de carácter más amplio que posibilitan la aplicación de los conocimientos adquiridos” (OCDE, 2006a p.12). No se trata del

10 Por ejemplo, para 2° Medio, el tercer Objetivo Fundamental declarado dice “Explorar sistemáticamente diversas estrategias para la resolución de problemas; profundizar y relacionar contenidos matemáticos” (MINEDUC, 2005, p.87)

aprendizaje de más elementos curriculares, sino de la forma en que estos son utilizados para la resolución de problemas.

2.2 Análisis de los énfasis de las evaluaciones de aula, respecto de PISA y SIMCE

En el ya mencionado estudio sobre la evaluación de aula en Chile¹¹ (MINEDUC, 2009), se concluye, entre otras cosas, que en relación a los contenidos temáticos evaluados en la Educación Básica, existe una mayor concentración de Números y operaciones, en desmedro de los otros temas (principalmente Geometría y Datos). En Educación Media el énfasis se traslada hacia el Álgebra, en desmedro principalmente de Datos y azar (véase la Tabla 2).

Tabla 2.
Distribución de las evaluaciones de aula según eje temático y según nivel (8° Básico y 2° Medio)

	Porcentaje de ítems en evaluaciones de 8° Básico	Porcentaje de ítems en evaluaciones de 2° Medio
Números	43,3%	5,2%
Álgebra	32,1%	47,4%
Geometría	24,6%	47,4%
Datos y azar	0%	0%
TOTAL	100%	100%

Fuente: MINEDUC, 2009

Por su parte, como se observa en la Tabla 3, la distribución de las preguntas según ejes de contenido temático, en PISA y SIMCE muestran algunas diferencias importantes.

11 Este estudio incluye de hecho un análisis muy completo del cual aquí solo usamos una pequeña parte. Incluye, por ejemplo, discrepancia entre el nivel de enseñanza de un tema según el Marco Curricular, y el nivel donde éste es evaluado, cuáles temas presentes en el currículum no son evaluados, ejemplos comentados de evaluaciones realizadas en aula, etc.

Tabla 3.

Distribución de ítemes PISA y SIMCE, según ejes de contenido temático y dificultad observada para la muestra de estudiantes chilenos y la muestra general (expresada en mediana del PRC)

	PISA 2006			SIMCE 2006	
	Porcentaje de ítemes	Mediana de PRC (*) (Chile)	Mediana de PRC (OCDE)	Porcentaje de ítemes	Mediana de PRC
Ítemes de Cantidad / Números / Operaciones	32%	31,5%	50%	26%	30,4%
Ítemes de Forma y Espacio / Geometría	23%	30,0%	43%	13%	42,5%
Ítemes de Cambio y Relaciones / Álgebra	13%	26,5%	44%	40%	50,2%
Ítemes de Incertidumbre / Datos y Azar	31%	29,5%	45%	20%	74,1%

(*) PRC: Porcentaje de respuestas correctas.

De la información de la Tabla 3, en primer lugar llama la atención que la presencia de “Cambio y Relaciones” (o Álgebra) es menor en PISA que en SIMCE, sin embargo la dificultad de los ítemes de este eje es mucho mayor en PISA. En efecto el PRC promedio de SIMCE es 50,2% y el de PISA 26,5%, para la muestra de estudiantes chilenos y 44% para el conjunto de países OCDE. En otras palabras, lo que son capaces de hacer los estudiantes chilenos en Álgebra es mucho menos de lo que se espera en PISA.

En el caso de “Datos y Azar”, nuevamente la dificultad de SIMCE es mucho menor que la de PISA (PRC 74,1% contra PRC 29,5%, respectivamente, en ambos casos, para la muestra de estudiantes chilenos). Esto se podría explicar por la existencia de una menor proporción de evaluaciones de aula en este eje temático. De hecho, según el estudio aludido (MINEDUC, 2009), no se observaron evaluaciones de este eje en 2° Medio. En este sentido, huelga decir que menos evaluaciones implican menos posibilidades de profundizar las habilidades a desarrollar, y en consecuencia, solo se desarrolla la capacidad de responder ítemes de menor dificultad.

En el eje de Números, se observa el mismo nivel de dificultad en PISA y SIMCE, para la muestra de estudiantes chilenos (aproximadamente PRC 30%). Estamos aquí en el tema que concentra la mayor atención de la evaluación de aula en la Enseñanza Básica; en desmedro de los otros ejes temáticos. Pese a ello, no se observa una disminución de la brecha entre los estudiantes chilenos y el grupo de la OCDE: frente al PRC de 31,5% de Chile, los estudiantes de los países OCDE muestran 50%. Por tanto, al igual que en todos los demás ejes, existe una diferencia de aproximadamente 20 puntos porcentuales entre la mediana de PRC para Chile y para la OCDE.

Ahora bien, respecto de las habilidades evaluadas, como se observa en la Tabla 4, en Educación Básica, la habilidad más frecuente en las evaluaciones de aula es la de Manejar *conocimientos y procedimientos* (en particular, asociada a la destreza de Calcular), y en Educación Media, el énfasis está en *Usar Conceptos* (en particular, asociado a las destrezas de *Conocer y Clasificar*). Las habilidades de *Resolver problemas de rutina y de Razonar* son mucho menos trabajadas.

Tabla 4.
Distribución de ítemes según ejes de habilidad¹²

Nivel	Manejar Conocimientos y Procedimientos / Usar Conceptos	Resolver Problemas de Rutina / Razonar
8°B	74,1%	26,0%
2°M	60,7%	39,4%

Fuente: MINEDUC, 2009

En contraste con esta realidad, PISA se orienta principalmente a la resolución de situaciones problemáticas, y por lo tanto su énfasis está justamente en las habilidades de Resolver problemas de rutina / Razonar, que concentran aproximadamente el 77% de los ítemes.

2.3 Análisis de los ítemes SIMCE y PISA, según las seis características asociadas a una mayor dificultad

Como se señaló en la sección “Metodología”, los ítemes de las pruebas PISA y SIMCE, fueron clasificados según seis características, cuya presencia tradicionalmente se asocia a una mayor dificultad, a saber:

- Situaciones reales
- Información visual
- Razonamiento de múltiples pasos
- Situaciones novedosas
- Cálculos numéricos
- Pensamiento reversible

A continuación se presenta, para cada una de estas características, el porcentaje de ítemes de las pruebas SIMCE y PISA que presenta esa característica y la mediana de las dificultades del grupo de ítemes con la característica y sin ella, para la muestra de estudiantes chilenos.

¹² Usando las categorías cognitivas de TIMSS, categorías con las que se trabajó en este estudio.

2.3.1 Situaciones reales

La clasificación de ítemes que presentan situaciones reales es la siguiente:

Tabla 5a.
Porcentajes de ítemes con situaciones reales, en las pruebas PISA y SIMCE

	PISA 2006	SIMCE 2006
Porcentaje de ítemes con situaciones reales	98%	70%

Tabla 5b.
Mediana de las dificultades por categoría (con/sin situaciones reales), en las pruebas PISA y SIMCE

Mediana de los PRC en:	PISA 2006	SIMCE 2006
Ítemes CON situaciones reales	31%	47%
Ítemes SIN situaciones reales	16%	48%

Como se observa en la Tabla 5a, los ítemes con situaciones reales son muchos más en PISA (98%) que en SIMCE (70%). De hecho, solo 1 de los 48 ítemes de PISA no presenta una situación real. Esta falta de evidencia, impide afirmar que la diferencia en la mediana de PRC (que se observa en la Tabla 5b) sea significativa. En SIMCE, no se observan diferencias entre las medianas de ítemes con y sin situaciones reales.

2.3.2 Información visual

La clasificación de ítemes que presentan información visual es la siguiente:

Tabla 6a.
Porcentajes de ítemes con información visual, en las pruebas PISA y SIMCE

	PISA 2006	SIMCE 2006
Porcentaje de ítemes con información visual	57%	31%

Tabla 6b.

Mediana de las dificultades por categoría (con/sin información visual), en las pruebas PISA y SIMCE

Mediana de los PRC en:	PISA 2006	SIMCE 2006
Ítemes CON información visual	25%	45%
Ítemes SIN información visual	31%	50%

En la Tabla 6a se observa que la proporción de ítemes con información visual es prácticamente el doble en PISA que en SIMCE (la presencia de estos ítemes en ambas pruebas difiere en 26 puntos porcentuales). Respecto del aumento de la dificultad que produce la inclusión de información visual, se pudo establecer que, en ambas pruebas, el porcentaje de respuesta correcta disminuye unos 5 puntos respecto de aquellos sin información visual (véase la Tabla 6b).

2.3.3 Razonamiento de múltiples pasos

La clasificación de ítemes que requieren razonamiento de múltiples pasos es la siguiente:

Tabla 7a.

Porcentajes de ítemes con razonamiento de múltiples pasos, en las pruebas PISA y SIMCE

	PISA 2006	SIMCE 2006
Porcentaje de ítemes con razonamiento de múltiples pasos	36%	27%

Tabla 7b.

Mediana de las dificultades por categoría (con/sin razonamiento con múltiples pasos), en las pruebas PISA y SIMCE

Mediana de los PRC en:	PISA 2006	SIMCE 2006
Ítemes CON razonamiento de múltiples pasos	23%	43%
Ítemes SIN razonamiento de múltiples pasos	31%	50%

En la Tabla 7a, se observa que la proporción de ítemes con razonamiento de múltiples pasos es levemente superior en PISA que en SIMCE (la presencia de estos ítemes aumenta en 9 puntos porcentuales). Respecto de la dificultad, en ambas pruebas el porcentaje de respuesta correcta disminuye cerca de 7 puntos en caso de requerir razonamiento de múltiples pasos (véase la Tabla 7b).

2.3.4 Situaciones novedosas

La clasificación de ítemes que presentan situaciones novedosas es la siguiente:

Tabla 8a.

Porcentajes de ítemes con situaciones novedosas, en las pruebas PISA y SIMCE

	PISA 2006	SIMCE 2006
Porcentaje de ítemes con situaciones novedosas	70%	44%

Tabla 8b.

Mediana de las dificultades por categoría (con/sin situaciones novedosas), en las pruebas PISA y SIMCE

Mediana de los PRC en:	PISA 2006	SIMCE 2006
Ítemes CON situaciones novedosas	22%	42%
Ítemes SIN situaciones novedosas	43%	52%

En la Tabla 8a, se observa que la proporción de ítemes con situaciones novedosas es muy superior en PISA, respecto de SIMCE (la presencia de estos ítemes en PISA es 26 puntos porcentuales superior que en SIMCE). Respecto del efecto sobre la dificultad, en ambas pruebas el porcentaje de respuesta correcta disminuye 21 puntos en el caso de PISA y 10 en el caso de SIMCE, para los ítemes con situaciones novedosas (véase la Tabla 8b).

2.3.5 Cálculos numéricos

La clasificación de ítems con dificultad por cálculos es la siguiente:

Tabla 9a.
Porcentajes de ítems con dificultad por cálculos numéricos, en las pruebas PISA y SIMCE

	PISA 2006	SIMCE 2006
Porcentaje de ítems con dificultad por cálculos	32%	27%

Tabla 9b.
Mediana de las dificultades por categoría (con/sin dificultad por cálculos), en las pruebas PISA y SIMCE

Mediana de los PRC en:	PISA 2006	SIMCE 2006
Ítems CON dificultad por cálculos	16%	43%
Ítems SIN dificultad por cálculos	31%	50%

En la Tabla 9a, se observa que la proporción de ítems con dificultad por cálculos es similar en PISA y en SIMCE (la presencia de estos ítems difiere solo en 5 puntos porcentuales). Respecto de la dificultad, en ambas pruebas el porcentaje de respuesta correcta disminuye, 15 puntos en el caso de PISA y 7 en el caso de SIMCE, cuando los ítems presentan dificultad por cálculos numéricos (véase la Tabla 9b).

2.3.6 Pensamiento reversible

La clasificación de ítems que requieren pensamiento reversible es la siguiente:

Tabla 10a.
Porcentajes de ítems con pensamiento reversible, en las pruebas PISA y SIMCE

	PISA 2006	SIMCE 2006
Porcentaje de ítems con pensamiento reversible	26%	8%

Tabla 10b.
Mediana de las dificultades por categoría (con/sin pensamiento reversible), en las pruebas PISA y SIMCE

Mediana de los PRC en:	PISA 2006	SIMCE 2006
Ítems CON pensamiento reversible	20%	50%
Ítems SIN pensamiento reversible	31%	48%

Como se observa en la Tabla 10a, la proporción de ítems con pensamiento reversible es prácticamente 3 veces superior en PISA que en SIMCE (la presencia de estos ítems es 18 puntos porcentuales superior en PISA). Respecto de la variación de la dificultad, en PISA el porcentaje de respuesta correctas disminuye 11 puntos, pero en SIMCE aumenta en 2 puntos porcentuales, cuando el ítem presenta pensamiento reversible (véase la Tabla 10b).

3. Conclusiones

Como primera hipótesis de este estudio se planteó que un escaso grado de alineación entre el Marco Curricular chileno y los contenidos temáticos evaluados en PISA podrían ser la causa lógica de la mayor dificultad presentada por los estudiantes chilenos al momento de responder la prueba; sin embargo, se observó que, prácticamente todas las preguntas de PISA pueden considerarse coherentes con el currículum chileno, y por lo tanto no es esto lo que explica la mayor dificultad observada.

Frente a la segunda hipótesis, sin embargo, relativa a los elementos característicos que parecen relacionarse con la dificultad que presentan las preguntas de PISA, los resultados sí permiten indicar cuáles son los que resultan especialmente difíciles para los estudiantes chilenos. En este contexto, de la primera comparación de PISA con SIMCE y las evaluaciones de aula, se observa que:

- Lo que se evalúa en PISA tiene su foco en la *Resolución de Problemas y el Razonamiento*, no así la evaluación de aula, cuyo énfasis está en *Manejar Conocimientos y Procedimientos y Usar conceptos*. Esto señala la necesidad de que en la escuela, exista una mayor profundización en el tratamiento de los contenidos temáticos. Un contenido temático debe ser trabajado más allá de un mero conocimiento de definiciones, propiedades y algunos procedimientos. Es necesario su uso en situaciones diversas y en la resolución de problemas, relacionarlo con otros contenidos temáticos, e idealmente en procesos más reflexivos (por ejemplo, en generalización o argumentación).
- Álgebra y Datos y Azar son los ejes de contenidos temáticos que aparecen particularmente deficitarios en la evaluación escolar, porque los problemas que se les pide resolver habitualmente a los estudiantes chilenos, son pocos o muy fáciles. Datos y Azar se evalúa muy poco en sala de clases y tanto este eje como Álgebra se evalúan a nivel nacional (en SIMCE), con preguntas más fáciles que en PISA. La exigencia de profundización es menor. Nótese que además, estos dos ejes de contenidos temáticos son los de aparición más tardía en los niveles de enseñanza. Según el Marco Curricular las medidas de tendencia central se estudian en 7° Básico; el Álgebra (expresiones algebraicas simples y ecuaciones de 1er grado), en 8° Básico, y el inicio de las probabilidades ocurre en 2° Medio.

Ahora bien, entre las 6 características estudiadas¹³, propias de cada ítem, se obtuvo que 5 de ellas efectivamente aumentan la dificultad de los ítems para la muestra de estudiantes chilenos. A continuación se señalan estas características, ordenadas según la diferencia que mostraron en el PRC de Chile, frente a los ítems que no presentaban el elemento analizado.

¹³ Aquí se presentan estas características ordenadas según la diferencia en PRC de Chile obtenidas entre los ítems PISA que tienen la característica y los ítems que no la tienen.

1. **Situaciones novedosas:** 21 puntos porcentuales más difíciles. ¿Estaremos formando estudiantes que evitan ser creativos?, ¿que no desean enfrentar situaciones nuevas? Puede ser que trabajar según una lista de CMO (Contenidos Mínimos Obligatorios) tenga por efecto que el sistema educacional se despreocupe de ver algo distinto de lo que allí aparece.
2. **Dificultades con cálculos numéricos:** 15 puntos porcentuales más difíciles. Es usual que consideremos que un problema debe ser presentado con números “simples”, y así debe ser también el resultado. ¿Deberíamos insistir más en el trabajo con números más difíciles? ¿Con números grandes o con decimales o que al dividirlos no den resultados “exactos”? ¿Cuáles son los efectos de estas alternativas, desde el punto de vista del aprendizaje? y ¿cómo influye en esto, el uso de calculadoras o computadores?
3. **Pensamiento reversible:** 11 puntos porcentuales más difíciles. “Devolverse” es más difícil (y no siempre posible). Esta característica está en la base del desarrollo de un análisis crítico de la información que nos entregan los medios de comunicación. En efecto, si por los medios de comunicación se muestra un resultado, se podría decir cuáles deberían haber sido los datos para llegar a ese resultado, y por lo tanto si tal resultados es posible.
4. **Razonamiento de múltiples pasos:** 8 puntos porcentuales más difíciles. Este elemento, como el anterior, es indicativo de un trabajo en profundidad. No se trata de preguntas de respuesta instantánea; requieren un tiempo de reflexión y de trabajo. ¿Será esto un efecto de la masificación de las preguntas de opción múltiple?
5. **Información visual:** 6 puntos porcentuales más difíciles. Estamos entrando en una era de las comunicaciones, en la cual la información visual

nos bombardea en forma continua. Es necesario fortalecer la capacidad de análisis de todas estas múltiples fuentes de información.

6. Situaciones reales. Aunque no hay suficiente información en PISA sobre el no uso de situaciones reales, se debe convenir en que su uso permite darle un sentido al estudio de la matemática. Su uso es talvez parte de la respuesta a la pregunta, ¿y esto, para que sirve?, pregunta recurrentemente escuchada en relación a la matemática.

Junto con las conclusiones anteriores, conviene destacar ciertas recomendaciones que podrían orientar la enseñanza de la matemática en Chile.

En primer lugar, cabe señalar que el uso de contextos en una pregunta de matemática, incluso si su descripción es extensa, resulta ser valioso para fortalecer el desarrollo de la resolución de problemas por parte de los estudiantes. Sin embargo, el uso de contextos en sí, no basta. Lo importante es poder utilizar información puesta en un contexto, de tal manera que los estudiantes no se limiten a reproducir procedimientos aprendidos bajo una instrucción académica orientada a resolver problemas de un determinado tipo. La repetición de procedimientos, transforman finalmente al problema en un problema de rutina, en una simple repetición de procesos ya aprendidos, que no aportan a una profundización de los aprendizajes. Enfrentar a los estudiantes a situaciones nuevas es importante, pero además, deben sentir la necesidad de usar procedimientos nuevos, donde tengan que organizar, analizar e incluso probar otras estrategias de solución. Adicionalmente, si los contextos son reales, no solo resultarán preguntas más interesantes, sino que de paso, servirán para mostrar la utilidad de la matemática.

Por otro lado, el cuantificar la dificultad de una pregunta en matemática, particularmente una pregunta del tipo resolución de problemas, es de gran interés, y hay mucho por hacer al respecto. Comprender de qué

depende la dificultad de un ítem, orienta sobre cómo enfocar los esfuerzos en la labor de la enseñanza de esta disciplina.

Consecuentemente con lo anterior, es aconsejable que cualquier posible cambio de orientación en la enseñanza de la matemática en el aula, se acompañe del correspondiente cambio en las evaluaciones, y esto tanto para las evaluaciones de aula, como para las evaluaciones SIMCE u otras evaluaciones estandarizadas.

Finalmente, dentro de las características de PISA analizadas, se puede encontrar inspiración que enriquezca el estudio de la matemática en su rol de herramienta para comprender el mundo¹⁴.

Referencias

- EERC (Educational Evaluation Research Consortium) (2005). *Report No. 3*. Extraído desde http://www.ceie.albany.edu/Quarterly%20Reports/Reports%20and%20Tech%20proposal/EERC%2003%2008%202005%20Report_Publica.doc
- Grønmo, L. y Olsen, R. (2006). *TIMSS versus PISA: The Case of Pure and Applied Mathematics*. Extraído desde http://www.iea.nl/fileadmin/user_upload/IRC2006/IEA_Program/TIMSS/Gronmo__Olsen.pdf
- ISEI-IVEI (Instituto Vasco de Evaluación e Investigación Educativa) (2005). *Proyecto PISA 2003. Ejemplos de ítems de Matemáticas y Solución de Problemas*. Extraído desde <http://www.isei-ivei.net/CAST/pub/pisaitemscastellano.pdf>
- MINEDUC (Ministerio de Educación) (2005). *Marco Curricular de la Educación Media. Objetivos Fundamentales y Contenidos Mínimos Obligatorios de la Educación Media*. Santiago, Chile.
- _____ (2009). *Evaluación de aula en enseñanza básica y media. Subsector Educación Matemática*. Santiago, Chile.
- Mullis, I. y otros (2003). *Assessment Frameworks and Specifications 2003*. Chestnut Hill, MA: International Study Center, Boston College.

14 Se pueden consultar los siguientes documentos con abundante material sobre preguntas utilizadas en pruebas PISA de aplicaciones anteriores: <http://www.oecd.org/dataoecd/48/23/39817028.pdf> y http://www.oecd.org/document/25/0,3343,en_32252351_3223573_1_39733465_1_1_1_1,00.html. Para otros documentos de interés con material relativo a ítems PISA ver OCDE, 2006b, 2006c, 2009.

Nohara, D. (2001). "A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)". *Working Paper*, 2001-07. Extraído desde <http://nces.ed.gov/pubns2001/200107.pdf>

OCDE (Organización de Cooperación y Desarrollo Económicos) (2006a). *PISA 2006. Marco de la evaluación. Conocimientos y habilidades en Ciencias, Matemáticas y Lectura*. Paris, Francia.

_____ (2006b). *PISA Released Items –Mathematic*. Extraído desde <http://www.pisa.oecd.org/dataoecd/14/10/38709418.pdf>

_____ (2006c). *PISA2006_MS Math ItemClassification_1.xls*. Paris, Francia.

_____ (2009). *Take the Test: Sample Questions from OECD's PISA Assessments*. Extraído desde http://www.oecd.org/document/31/0,3343,en_32252351_32236191_4194268_7_1_1_1_1,00.html

Robitaille, D. (Ed.) (1993). "Curriculum Frameworks for Mathematics and Science", *TIMSS Monograph Series*, 1. Vancouver, Canada: Pacific Educational Press.

Gráfico A.2: Distribución ítems PISA para la muestra de estudiantes chilenos, según rangos de dificultad expresados en PRC

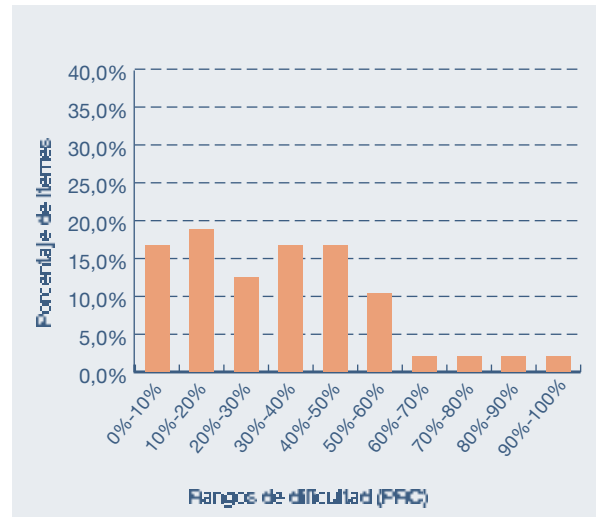
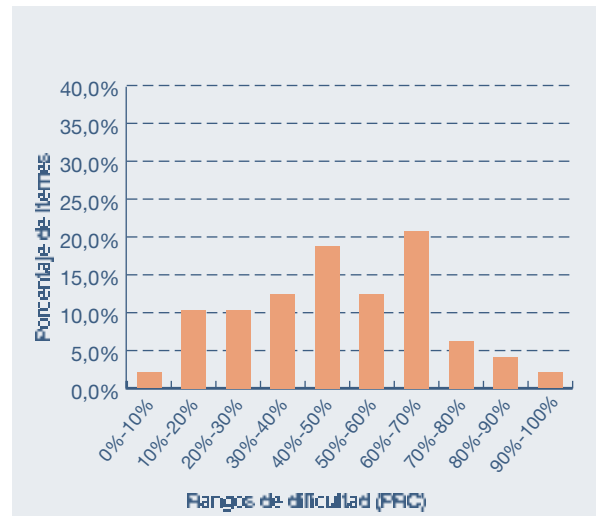


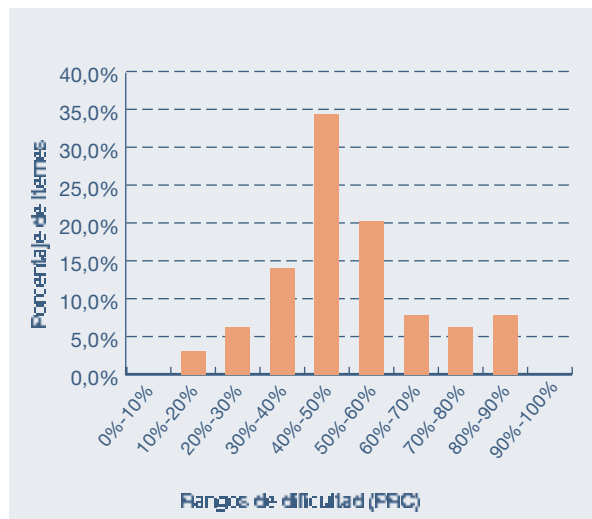
Gráfico A.3: Distribución ítems PISA para la muestra de países OCDE, según rangos de dificultad expresados en PRC



Anexo

1. Distribuciones de dificultad de las pruebas PISA y SIMCE 2006, Matemática

Gráfico A.1: Distribución ítems SIMCE, según rangos de dificultad expresados en PRC



Marco Curricular Analítico (MCA)¹⁵

Documento ajustado por los autores para los propósitos de este estudio, basado en el documento llamado “Marco Curricular Analítico” (MCA), preparado por el Consorcio de Evaluación e Investigación Evaluativo (CEIE) de República Dominicana, bajo la dirección de G. A. Valverde.

A - Clasificación de categorías de contenidos temáticos

1. **Ámbito: Familias de Números–Conjuntos numéricos**

- 1.1 Naturales
- 1.2 Fracciones comunes
 - 1.2.1 Fracciones propias
 - 1.2.2 Fracciones impropias
 - 1.2.3 Fracciones mixtas
 - 1.2.4 Fracciones decimales
- 1.3 Decimales
 - 1.3.1 Decimales finitos
 - 1.3.2 Decimales infinitos con periodo
 - 1.3.3 Decimales infinitos con semiperiodo
- 1.4 Fracciones – Racionales
- 1.5 Enteros
- 1.6 Irracionales
- 1.7 Complejos (imaginarios)
- 1.8 Otras categorías (pares/impares, primos, múltiplos, divisores, etc.)

2. **Ámbito: Representaciones numéricas**

(las diversas escrituras de los números)

- 2.1 Sistemas Aditivos (palitos, romanos, dedos)
- 2.2 Sistemas de Notación Posicional
- 2.3 Otros sistemas de numeración (horario, etc.)
- 2.4 Fracciones

¹⁵ MCA: Marco Curricular Analítico. Consiste en una presentación exhaustiva de elementos que caracterizan un currículum. Puede utilizarse, por ejemplo, para analizar un texto, una actividad pedagógica o un reactivo en una prueba. El MCA presenta una categorización de contenidos, que se conocen también como contenidos disciplinares y una categorización de expectativas de rendimiento que representan lo que se pretende que los estudiantes sean capaces de hacer con los contenidos.

2.5 Decimales

- 2.5.1 Decimales con periodo y con semiperiodo
- 2.6 Potencias – Números con Exponentes
- 2.7 Notación científica
- 2.8 Números negativos (significado del signo -)
- 2.9 Racionales
- 2.10 Irracionales
- 2.11 Representaciones gráficas: recta numérica y otras escalas graduadas, tablas, gráficos
- 2.12 Representaciones de datos en tablas

3. **Ámbito: Transformaciones, Operaciones, Funciones (cómo se obtiene un número que resulta al trabajar con uno o más números)**

- 3.1 Comparar
- 3.2 Ordenar
- 3.3 Máximo y mínimo
- 3.4 Redondear
- 3.5 Sumar y restar
- 3.6 Multiplicar – producto
- 3.7 Dividir
- 3.8 Potencias
- 3.9 Raíces
- 3.10 MCD – MCM
- 3.11 Múltiplos y divisores
- 3.12 Valor absoluto
- 3.13 Polinomios
- 3.14 Fracciones racionales

4. **Ámbito: Algoritmos, procedimientos estándares**

- 4.1 En Números:
 - 4.1.1 Procedimientos para cálculo mental
 - 4.1.2 Descomposición aditiva y factorización
 - 4.1.3 Representaciones gráficas en recta numérica
 - 4.1.4 Descomposición/composición de áreas (rectángulos)
 - 4.1.5 Procedimiento escrito estándar (para las 4 operaciones)
 - 4.1.6 Procedimiento escrito nuevo
 - 4.1.7 Procedimiento mecánico: dedos, ábaco, calculadora, computador

- 4.1.8 Estimaciones por “órdenes de magnitud” (redondear, aproximar)
- 4.1.9 Razón
- 4.1.10 Proporcionalidad
- 4.1.11 Porcentajes
- 4.2 En Álgebra – Dominio variacional:
 - 4.2.1 Variables (uso y comprensión)
 - 4.2.2 Graficar funciones, gráficos de puntos, de líneas, plano xy.
 - 4.2.3 Leer propiedades de funciones desde el gráfico (Máx., Mín., crecimiento)
 - 4.2.4 Gráfico e intersecciones (imagen, preimagen, simetría del gráfico, asíntotas verticales, horizontales, oblicuas)
 - 4.2.5 Tabular funciones
 - 4.2.6 Ecuaciones
- 4.3 Álgebra – Dominio simbólico
 - 4.3.1 Reducción de términos semejantes
 - 4.3.2 Factorización
 - 4.3.3 Evaluación de una fórmula, conocidos los argumentos
- 4.4 En Geometría:
 - 4.4.1 Área o perímetro de un rectángulo
 - 4.4.2 Estimaciones por visualización gráfica
- 5. Ámbito: Algoritmos, procedimientos no estándares**
- 6. Ámbito: Propiedades de las operaciones, transformaciones y funciones**
 - 6.1 Propiedades de la suma y del producto
 - 6.2 Distributividad
 - 6.3 Factorización
 - 6.4 Propiedades de los exponentes
 - 6.5 Propiedades de las raíces
 - 6.6 Propiedades raíces-exponentes
 - 6.7 $F+G$, aF
 - 6.8 $F(x \pm a)$, $F(ax)$
- 7. Ámbito: Medida, relación entre formas geométricas y cantidad**
 - 7.1 Medidas de longitud. Sumar y restar largos
 - 7.2 Medidas de longitud. Subdividir y contar. Fracciones y/o decimales
 - 7.3 Medidas de longitud. Estimar a partir de referentes
 - 7.4 Medidas de áreas. Rectángulos y productos. Distributividad
 - 7.5 Medidas de áreas. Triángulos
 - 7.6 Medidas de perímetros. Cuadrado y rectángulo
 - 7.7 Medidas de perímetros. Polígonos (triángulos)
 - 7.8 Medidas de volúmenes. Cubos y prismas
 - 7.9 Medidas de volúmenes. Cilindros
 - 7.10 Medidas de volúmenes. Conos y esferas
- 8. Ámbito: Coordenadas, relación entre posiciones y representaciones numéricas.**
 - 8.1 Coordenadas de un punto en el plano
 - 8.2 Coordenadas de un punto en el espacio
 - 8.3 Proporcionalidad y coordenadas de rectas que pasan por el origen
 - 8.4 Coordenadas de puntos sobre una recta y relaciones afines
 - 8.5 Pendiente de rectas y paralelismo
 - 8.6 Pendiente de rectas y ortogonalidad
 - 8.7 Distancia entre 2 puntos, a partir de sus coordenadas
 - 8.8 Ecuación de una circunferencia
 - 8.9 Ecuación de 2° grado y la parábola
 - 8.10 Cónicas y sus ecuaciones
 - 8.11 Ecuaciones rectas, planos y superficies en el espacio
 - 8.12 Intersección de rectas y sistemas de ecuaciones lineales
 - 8.13 Intersección de rectas y cónicas y sistemas de 2° grado
- 9. Ámbito: Representación de relaciones entre cantidades numéricas**
 - 9.1 Fórmulas de cálculo simple (perímetros, áreas, volúmenes)
 - 9.2 Relación funcional entre 2 cantidades (una es función de la otra)
 - 9.3 Relación entre 2 cantidades (tipo balance o equilibrio)

- 9.4 Sistemas de relaciones (ecuaciones) lineales
- 9.5 Relaciones (ecuaciones) cuadráticas
- 10. Ámbito: medidas de tendencia central**
 - 10.1 Promedio
 - 10.2 Moda
 - 10.3 Mediana
- 11. Ámbito: Probabilidad y frecuencia relativa**
 - 11.1 La probabilidad como límite de frecuencias relativas
 - 11.2 Caso equiprobable. Probabilidad y “reparto equitativo”
 - 11.3 Caso equiprobable. Ley de Laplace (casos favorables/casos totales)
- 12. Ámbito: Formas geométricas**
 - 12.1 Bidimensional
 - 12.1.1 Puntos, rectas, segmentos, semirectas y rayos
 - 12.1.2 Plano coordenado en 2D
 - 12.1.3 Ángulos
 - 12.1.4 Paralelismo y perpendicularidad
 - 12.1.5 Triángulos, cuadriláteros y su clasificación y propiedades
 - 12.1.6 Teorema de Pitágoras y sus aplicaciones
 - 12.1.7 Círculo, circunferencia y sus propiedades (métricas y angulares)
 - 12.1.8 Otros polígonos y sus propiedades
 - 12.2 Tridimensional
 - 12.2.1 Plano coordenado en 3D
 - 12.2.2 Planos y rectas en el espacio
 - 12.2.3 Paralelismo y perpendicularidad de planos
 - 12.2.4 Formas y superficies tridimensionales y sus propiedades
- 13. Ámbito: Mediciones**
 - 13.1 Bidimensional
 - 13.1.1 Concepto de medida (incluyendo unidades no estándar)
 - 13.1.2 Utilización de instrumentos
 - 13.1.3 Medida y clasificación de ángulos
 - 13.1.4 Cálculos, fórmulas y propiedades de longitud y perímetro
 - 13.1.5 Cálculos, fórmulas y propiedades del área
 - 13.2 Tridimensional
 - 13.2.1 Cálculos, fórmulas y propiedades del área de la superficie de un cuerpo
 - 13.2.2 Cálculos, fórmulas y propiedades del volumen
- 14. Ámbito: Posición y Visualización**
 - 14.1 Redes
 - 14.2 Perspectivas
 - 14.3 Vistas
 - 14.4 Cortes de un cuerpo por un plano
- 15. Ámbito: Simetría, Congruencia y Semejanza (como ámbito de transformaciones)**
 - 15.1 Simetrías
 - 15.2 Traslaciones
 - 15.3 Rotaciones
 - 15.4 Patrones, teselas, frisos, etc.
 - 15.5 Congruencia de triángulos
 - 15.6 Congruencia de otras figuras geométricas
 - 15.7 Triángulos semejantes y sus propiedades
 - 15.8 Otras figuras semejantes y sus propiedades
- 16. Ámbito: Geometría Constructiva**
 - 16.1 Construcciones geométricas, usando bordes rectos y compás
- B - Clasificación de Expectativas de Desempeño**
 - 1. Conocer**
 - 1.1 Reconocer una representación
 - 1.2 Construir una representación
 - 1.3 Reconocer equivalencias
 - 1.4 Construir un objeto o su descomposición en uno o más equivalentes
 - 1.5 Recordar objetos matemáticos y sus propiedades

- 1.6 Reconocer objetos matemáticos y sus propiedades
- 1.7 Seleccionar una unidad apropiada para una medida dada
- 1.8 Seleccionar una herramienta apropiada para una medida dada
- 1.9 Seleccionar un grado apropiado de exactitud para una medición en una situación y tarea dadas
- 1.10 Decidir cuándo un estimado o una respuesta correcta es apropiado
- 1.11 Leer, interpretar una presentación de datos y/o usarla para contestar una pregunta
- 1.12 Comparar (cantidades, figuras, etc.)
- 1.13 Clasificar objetos matemáticos (números, cantidades, figuras, etc.)
- 1.14 Situar, ubicar y ubicarse a partir de uno o más criterios
- 1.15 Reconocer la relación entre símbolo y otra representación del mismo objeto, por ejemplo, traduciendo de uno a otro

2. Resolver problemas

- 2.1 Asociar y aplicar conocimiento (contenidos) a preguntas
- 2.2 Formular y clarificar problemas y situaciones
- 2.3 Seleccionar o construir una representación matemática de un problema
 - 2.3.1 Representación numérica
 - 2.3.2 Representación en tablas
 - 2.3.3 Representación algebraica
 - 2.3.4 Diagrama tipo flujo
 - 2.3.5 Representación como máquina
 - 2.3.6 Representación tipo gráfico (XY, tablero, plano)
 - 2.3.7 Representación con vectores y/o transformaciones
 - 2.3.8 Representación con símbolos, con colores, etc.
- 2.4 Determinar datos o el rango de datos que se necesitan para resolver un problema
- 2.5 Seleccionar, adaptar y/o desarrollar una estrategia

- 2.5.1 Árbol de decisiones, búsqueda de (todas las) alternativas
 - 2.5.2 Reducción a ecuaciones
 - 2.5.3 Reducción a un problema de gráficos
 - 2.5.4 Reducción a un problema de geometría
 - 2.5.5 Estimación de cantidades
 - 2.5.6 Recolectar y organizar datos
 - 2.5.7 Reemplazar por un problema más simple
 - 2.5.8 Reemplazar por una combinación de problemas
 - 2.5.9 Aprovechar la solución de un problema más simple
- 2.6 Secuenciar, dar orden, a partir de uno o más criterios

3. Razonamiento matemático

- 3.1 Desarrollar un algoritmo formal para realizar cálculos
- 3.2 Generalizar una estrategia conocida para resolver un problema específico
- 3.3 Identificar o establecer una conjetura para explicar un patrón en los datos, una tabla, un gráfico, etc.
- 3.4 Establecer y verificar una conjetura
- 3.5 Justificar
- 3.6 Identificar una contradicción
- 3.7 Conexiones y traducción entre representaciones

4. Comunicar

- 4.1 Desarrollo de notación y vocabulario matemático
- 4.2 Seleccionar un uso correcto de terminología y vocabulario
- 4.3 Seleccionar el uso correcto de notación matemática
- 4.4 Escribir una expresión correcta, utilizando notación matemática apropiada
- 4.5 Expresar una relación entre cantidades, utilizando expresiones matemáticas, oraciones abiertas, ecuaciones, etc.
- 4.6 Describir/discutir

Análisis comparativo de la estructura interna y objetivos de evaluación de las pruebas de Lectura SIMCE y PISA

Por
Cristián Donoso
Pablo Lima¹

Resumen

El objetivo de este estudio es comparar las pruebas SIMCE y PISA, aplicadas en 2001 y 2006. A partir de esta comparación, se intenta plantear una hipótesis respecto de las diferencias de tendencia observadas entre ambas evaluaciones.

Para confrontar los instrumentos se consideraron cinco aspectos: los textos-estímulo, los contextos de la lectura, el formato de las respuestas, los procedimientos de lectura involucrados y los contenidos conceptuales evaluados.

A partir de lo anterior, se pudo establecer que las tendencias que muestran los resultados de ambas evaluaciones (alza significativa en PISA y diferencia no significativa en SIMCE) no son comparables, principalmente por la diferencia que existe en el constructo evaluado de ambas pruebas.

En efecto, mientras PISA evalúa exclusivamente la habilidad de comprensión lectora desde un enfoque principalmente funcional; los resultados del SIMCE dan cuenta de la habilidad lectora, considerando distintas perspectivas, lo que la define como una evaluación más general del subsector curricular de Lengua Castellana y Comunicación.

¹ Cristián Donoso y Pablo Lima son miembros del Equipo Disciplinario de Lectura del SIMCE. Los autores de este estudio quieren destacar la participación de Andrés Anwandter, Coordinador del Equipo de Lectura SIMCE durante el año 2008, en el diseño de esta investigación.

Introducción

La prueba PISA está referida a niveles internacionales de logro de aprendizaje establecidos por los países desarrollados que se agrupan en la OCDE. Estos estándares se consideran necesarios para el óptimo desenvolvimiento del individuo en una sociedad moderna y del conocimiento. Por su parte, la prueba SIMCE es una prueba nacional chilena de carácter censal, que se aplica desde 1988, y que está referida al Marco Curricular nacional, el cual establece los Objetivos Fundamentales y Contenidos Mínimos que todo establecimiento educativo debe asegurar en el individuo para su pleno desarrollo en nuestra sociedad. Sus resultados están asociados a consecuencias en el sistema educacional chileno.

Las pruebas PISA de Lectura, aplicadas en Chile en los años 2001 y 2006 y las pruebas nacionales SIMCE aplicadas en los mismos años, compartieron buena parte de los estudiantes en sus muestras². Pese a ello, los resultados de ambas evaluaciones fueron muy distintos: mientras en PISA los puntajes aumentaron de 409 a 442 puntos entre 2001 y 2006³; en SIMCE hubo una variación no significativa de apenas dos puntos (de 252 en el año 2001 a 254 en el 2006)⁴.

Aunque las escalas de ambas pruebas no son comparables, sería razonable encontrar una tendencia similar, lo que no ocurre. Ante esta diferencia de resultados, cabe preguntarse hasta qué punto los instrumentos son comparables. Si bien existe una serie de variables (sociológicas, pedagógicas, métricas,

entre otras) que podrían contribuir a explicar por qué se observan resultados diferentes en dos mediciones de una misma área, en este estudio se plantea una comparación de dos aspectos centrales en toda evaluación: la estructura interna de los instrumentos y su propósito evaluativo. Así, en este estudio se planteó como propósito describir y analizar las semejanzas y diferencias de estas pruebas en cuanto a la forma de los instrumentos (cómo están construidas) y lo que miden (sus objetivos de evaluación). Estas dos líneas iniciales de investigación se materializaron en cinco variables: tipos textuales, contextos de lectura, formatos de respuestas (primera línea), procedimientos de lectura y contenidos conceptuales implicados (segunda línea).

A continuación, en la primera parte de este artículo, se describe la metodología utilizada, luego se señalan los principales resultados obtenidos, y posteriormente se exponen las conclusiones y se propone una discusión a partir de lo expuesto.

1. Metodología

Este estudio comprende dos tipos de análisis: uno descriptivo de los instrumentos PISA y SIMCE, de manera de llevar a cabo una comparación en cuanto a sus características estructurales generales, y otro descriptivo de los objetivos de evaluación, para lo cual, como se explicará más adelante, se utilizó como criterio de comparación el Marco Curricular chileno.

El corpus que se utilizó fueron las pruebas SIMCE de 2° Medio de Lengua Castellana y Comunicación, aplicadas en los años 2001 y 2006, y las pruebas de comprensión lectora PISA 2000 (aplicada en Chile el 2001) y 2006. Textos e ítemes de cada sistema de evaluación fueron considerados como un conjunto, como se muestra en la Tabla 1.

2 PISA se aplica a una muestra de estudiantes de 15 años, independientemente del nivel educativo que cursen, mientras que SIMCE se aplica a todos los estudiantes de 2° Medio, cuyo promedio de edad es de 15 años. Una buena parte de la muestra de estudiantes chilenos que participó en PISA 2000 y 2006, también lo hizo en SIMCE, durante esos mismos años.

3 Estos resultados acercan a Chile al promedio de la OCDE (492 puntos) y lo ponen significativamente por encima del resto de los países de América Latina, cuyo promedio es de 403 puntos.

4 En la mitad de este período, en el año 2003, hubo una aplicación más de SIMCE en 2° medio, la cual arrojó un promedio nacional de lenguaje de 253.

Tabla 1.
Textos e ítemes en PISA (2000 y 2006) y
SIMCE (2001 y 2006)⁵

	PISA	SIMCE
Número de textos-estímulo	37	29
Número de ítemes 2000/2001	135	83
Número de ítemes 2006	2 (28)	31 (41)
Total ítemes	137	114

El primer paso fue clasificar los textos-estímulo de ambas pruebas. Para ello, se planteó una aproximación individual a cada prueba, teniendo en cuenta que PISA y SIMCE usan diferentes taxonomías. Vale decir, los textos-estímulo de SIMCE fueron descritos según su clasificación propia, y lo mismo se realizó con PISA. Posteriormente se intentó un cruce de clasificaciones, como se verá más adelante.

En segundo lugar, se definieron los contextos de lectura de cada texto-estímulo, utilizando las categorías aportadas por el marco de la prueba PISA, porque SIMCE no declara esta dimensión.

El tercer paso fue identificar los formatos de respuesta a las preguntas de ambas pruebas, para lo cual se tomó la taxonomía que usa PISA, y que es ampliamente utilizada en el contexto de las mediciones de gran escala.

Para la cuarta etapa (procedimientos y contenidos conceptuales), fue necesario definir cuál sería la base de comparación para ambas pruebas y por qué se adoptaba una determinada decisión. En este sentido, mientras la prueba SIMCE pretende estar alineada al currículum nacional, la prueba PISA se basa en un marco de evaluación que se sostiene en la idea de competencias adquiridas a través de un

aprendizaje a lo largo de la vida (OCDE-INCE, 2000, p.20). Dada esta diferencia de principio, se acordó la relevancia de mirar ambos instrumentos desde la óptica del Marco Curricular chileno vigente a la fecha (Decreto 220, 1998), para analizar cuáles son las áreas específicas que toman en cuenta ambas pruebas, en cuáles áreas convergen y en cuáles se diferencian, por cuanto el interés del estudio es analizar lo que los estudiantes chilenos deberían haber aprendido a la edad de quince años.

Con el objeto de comparar ambos instrumentos desde una base común, se construyó un Marco Curricular Analítico (MCA) basado en el Marco Curricular chileno vigente en los niveles de 1° y 2° año de Enseñanza Media (MINEDUC, 1998). Esta versión del Marco Curricular es la que rigió las pruebas SIMCE de 2° Medio, aplicadas en los años 2001 y 2006.

En términos generales, como señala Valverde (2005) un Marco Curricular Analítico (MCA) es “una herramienta taxonómica de análisis en la cual se enumeran y se distinguen los diferentes elementos del currículo de una específica materia o curso”. Se utiliza frecuentemente para analizar las tres dimensiones de un marco curricular nacional (currículum prescrito, currículum implementado y currículum logrado) e identifica contenidos conceptuales, procedimientos y perspectivas (metas no-cognoscitivas o contenidos actitudinales).

Para este estudio, el modelo de MCA se adaptó de manera que fuera funcional al objetivo de comparar los ítemes de ambas pruebas. Por ello, se construyó un MCA que, basado en el Marco Curricular chileno de 1° y 2° Medio, permitiera distinguir entre procedimientos, contenidos conceptuales y perspectivas (véase este MCA en el Anexo).

Antes de aplicar el MCA, se procedió a describir los ítemes de ambos instrumentos, considerando básicamente los siguientes elementos: descripción de

⁵ En el caso de las aplicaciones 2006, el primer número refiere a los ítemes que aparecieron en ellas por primera vez (para efectos de este estudio) y el número entre paréntesis apela al total de ítemes. La diferencia que se observa entre los totales de ítemes entre PISA 2000 y 2006 no constituye una dificultad, puesto que se resguarda que se mantenga la comparabilidad entre las aplicaciones.

la habilidad implicada (procedimiento) y mención del contenido conceptual que se requiere para responder el ítem. Además, la descripción incluyó una mención al texto-estímulo sobre el que se realiza la tarea de lectura. Finalmente, y solo en los casos en que resultó relevante, se incluyó alguna información sobre la extensión del texto-estímulo (véanse ejemplos de la metodología empleada en el Anexo).

En el análisis según el MCA, cada ítem podía vincularse con un contenido y un procedimiento, con más de un contenido y más de un procedimiento, o con ninguno de los contenidos o procedimientos (esto último solo ocurrió con algunos ítems de la prueba PISA, por cuanto el MCA se basó en el currículum chileno, al cual PISA no está sujeto). No se consideraron las perspectivas presentes en el MCA, ya que no fueron identificables valoraciones claras en los ítems analizados.

A continuación, se presentan los resultados generales de este estudio.

2. Resultados

2.1 Análisis de los tipos de textos-estímulo

A. Descripción de los tipos de textos-estímulo posibles de ser encontrados en PISA y SIMCE

En una prueba de comprensión lectora, las estructuras, formatos, longitud y dificultad de los textos-estímulo son variables. En cuanto los ítems se refieren a esos textos, los marcos de evaluación definen tipologías textuales para evaluar la lectura. En este ámbito de las tipologías textuales, lo primero que se puede observar es que las clasificaciones empleadas por las pruebas SIMCE y PISA no coinciden.

Así, como se señala en una publicación del Equipo de Estudios Internacionales SIMCE (2004, pp. 28 y 29), en PISA se establece una clasificación de base, que está “en el centro mismo de la evaluación de la lectura”,

según la cual se distingue entre textos *continuos* y *discontinuos*. Vale decir, clasifica los textos mediante criterios diferentes para lograr una amplia diversidad textual en la evaluación.

Los textos continuos son aquellos compuestos de estructuras lingüísticas mayores, es decir, oraciones y párrafos. Los discontinuos tienen una configuración (estructural o gráfica) especial, que permite una distribución de la información distinta de la lineal y, en consecuencia, permite un manejo igualmente no lineal de la información por parte del lector. Los textos discontinuos son clasificados de acuerdo a su estructura, mientras que los continuos, en función de su contenido y el propósito del autor. En el caso de los textos continuos, se distinguen siete tipos: narración, exposición, descripción, argumentación y persuasión, instrucción, registro e hipertexto. Para los textos discontinuos se postulan nueve tipos: cuadros y gráficos, tablas, diagramas, mapas, formularios, hojas informativas, convocatorias y anuncios, vales, y certificados (OCDE, 2006).

En SIMCE, por su parte, los textos se clasifican inicialmente según un criterio de ficción/no ficción,



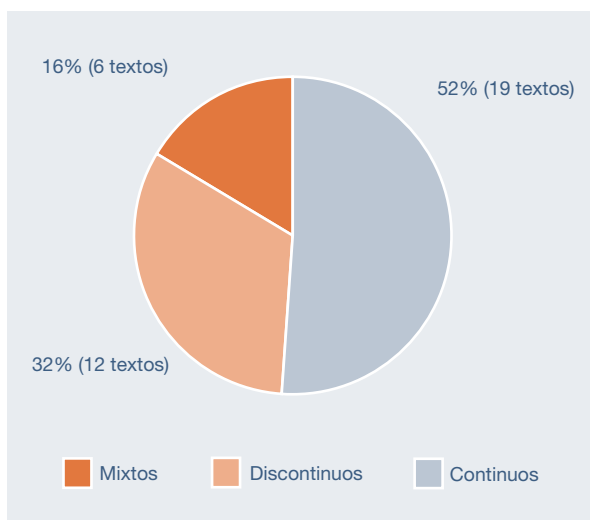
lo que origina una primera distinción entre textos literarios y no literarios. Entre los primeros, se cuentan narraciones y poemas. Los segundos se dividen en textos expositivos (principalmente de divulgación científica), textos instruccionales y textos provenientes de los medios masivos de comunicación, lo que en la práctica quiere decir que son textos con contenidos persuasivos o que manifiestan puntos de vista (SIMCE, 2001).

B. Comparación de los tipos de textos-estímulo incluidos en las pruebas PISA y SIMCE

En PISA, de los siete tipos de textos continuos, hay dos que no aparecen en las aplicaciones analizadas: registro e hipertexto. De los tipos de textos discontinuos, que son nueve, hay uno que no aparece: “vale”. Además, se incluyeron textos mixtos, que integran textos continuos y discontinuos (por ejemplo, una exposición acompañada de un mapa).

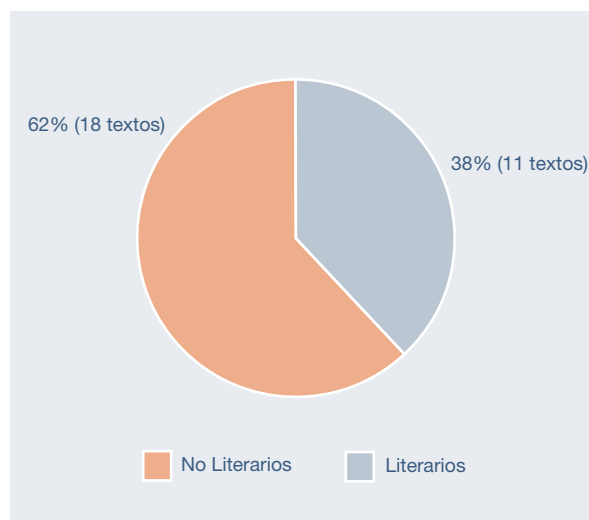
En términos de la distribución, en el Gráfico 1 es posible observar que existe un predominio de textos continuos, seguidos de los textos discontinuos y finalmente los textos mixtos.

Gráfico 1: Distribución efectiva de los tipos de textos-estímulo incluidos en las pruebas PISA



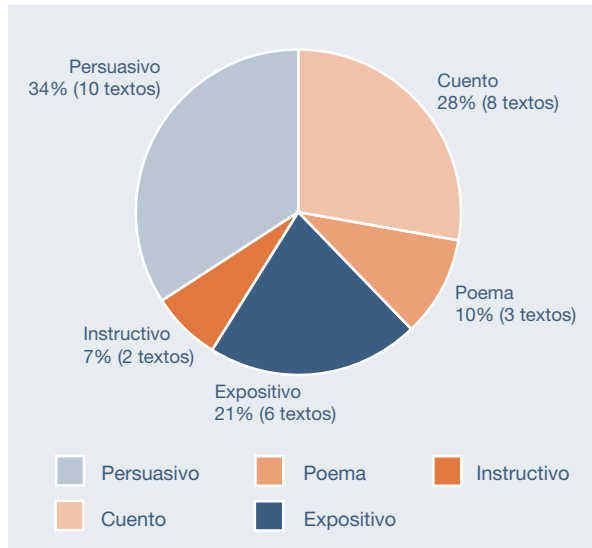
En el caso de las pruebas SIMCE, como se observa en el Gráfico 2, se incluyen las dos categorías contempladas en su marco de evaluación: literarios y no literarios, existiendo un predominio de los del segundo tipo.

Gráfico 2: Distribución efectiva de los tipos de textos-estímulo incluidos en las pruebas SIMCE, según categorías generales



Aunque en el Gráfico 2 se muestra un predominio de textos no literarios, hay que considerar que esta proporción gruesa puede entregar una idea algo engañosa, pues la categoría de textos literarios consta de solo dos subtipos: poemas y narraciones. Esto quiere decir que la prueba es abundante en cuentos y poemas. En cambio, los textos no literarios reparten su 62% entre tipos textuales bastante distintos y de diversas procedencias: textos expositivos de variada índole, textos provenientes de los medios masivos de comunicación (donde se puede apreciar convergencia y divergencia interpretativa o textos provenientes del mundo de la publicidad, como afiches o avisos) y textos instruccionales. En el Gráfico 3 se puede apreciar la distribución de los textos según estas subcategorías.

Gráfico 3: Distribución efectiva de los tipos de textos-estímulo incluidos en las pruebas SIMCE, según categorías específicas

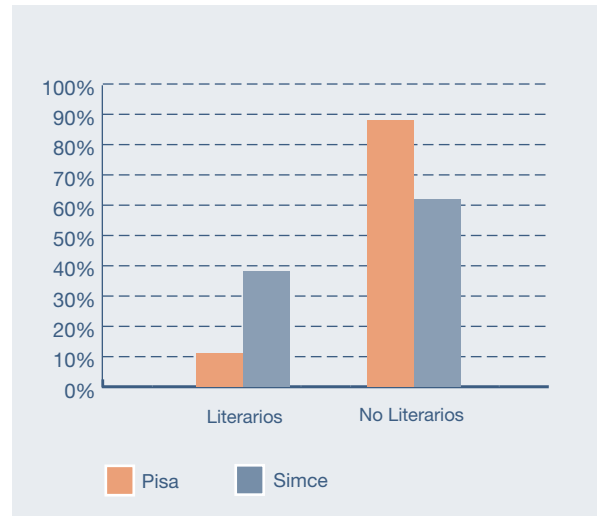


Como se observa en el Gráfico 3, los textos de intención persuasiva son los más utilizados en la prueba SIMCE. Además, es posible observar que entre los textos literarios los que tienen más presencia son los cuentos, que conforman además el segundo tipo más usado del total.

Otro dato que es posible recoger es que la separación entre textos no literarios es bastante diversa, y más bien corresponde a propósitos textuales, frente a la separación por géneros que se aprecia en los literarios. Si la categoría de textos no literarios es tan heterogénea, cabría hacer la pregunta de hasta qué punto esta no se trata de una categoría abierta, en la que eventualmente podrían entrar varios otros tipos, como lo hace PISA.

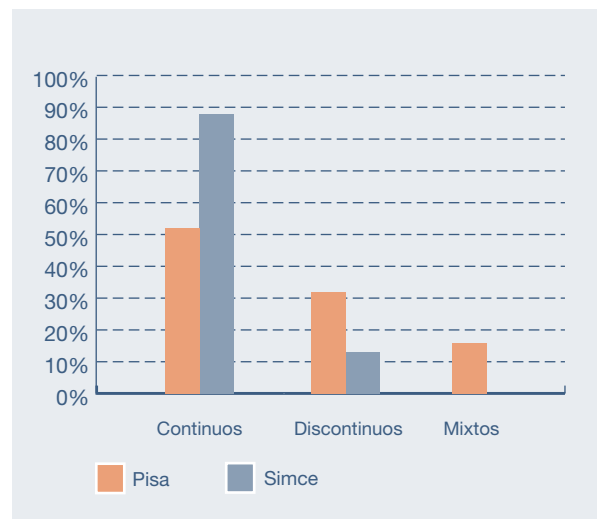
Ahora bien, con el objeto de poder realizar una comparación en términos de tipos de textos-estímulo, en los Gráficos 4 y 5 se presenta el análisis de ambas pruebas, según cada una de las tipologías propuestas.

Gráfico 4: Distribución de textos-estímulo de las pruebas PISA y SIMCE, según la taxonomía SIMCE



En esta comparación es interesante observar que si aplicamos el modelo de SIMCE a PISA, encontramos que de los 37 textos que utiliza esta última solo 4 son literarios, lo que constituye el 11% de la prueba. En SIMCE, el porcentaje de textos literarios también es menor al de no literarios, pero alcanza el 38%.

Gráfico 5: Distribución de textos-estímulo de las pruebas PISA y SIMCE, según la taxonomía PISA



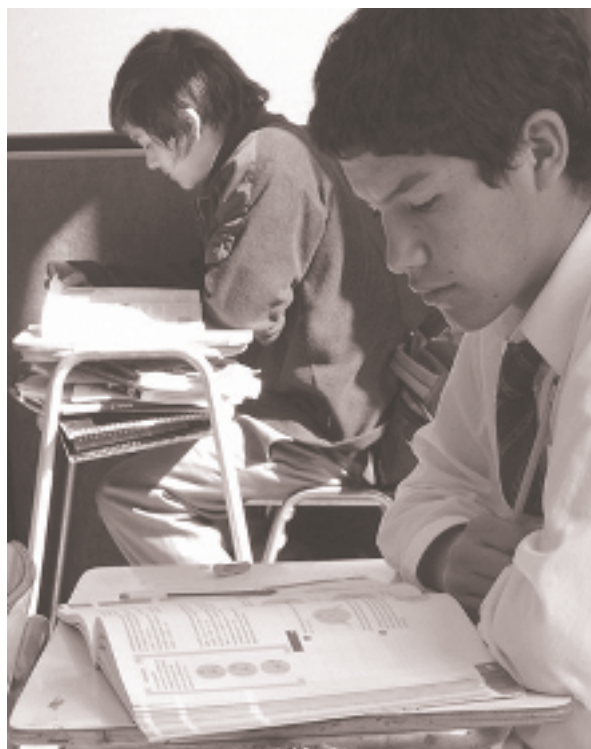
Vista la prueba SIMCE desde el modelo de PISA, encontramos que la mayoría de los textos de SIMCE son continuos, salvo en el caso de los avisos publicitarios, observándose la siguiente distribución: de los 29 textos, solo 4 son discontinuos (13% de la prueba). En ambas pruebas el porcentaje de textos continuos es mayor, pero en PISA la diferencia con los otros tipos es más estrecha.

Hay que observar, por lo tanto, que PISA enfatiza la lectura de textos no literarios para la evaluación, lo que puede relacionarse con la perspectiva más funcional que le da a la lectura (Sanz Moreno, 2005). Por su parte, es evidente que SIMCE destaca la presencia de textos continuos, lo que se explica en la relevancia que tiene la literatura en esta prueba (contrastándola con su escasa presencia en PISA), pero también en la utilización de formatos “conservadores” de textos no literarios, sin darle demasiada importancia a los formatos discontinuos que corresponden generalmente a textos como formularios, hojas informativas o mapas. Además, SIMCE no utiliza textos mixtos, lo que provoca que presente un espectro de textos más restringido. En definitiva, lo relevante que resulta de la comparación con PISA, es observar que esta última levanta una categoría paralela de textos discontinuos muy bien aprovechada, que origina una selección de tipos textuales diferentes de los “tradicionales” (léase “continuos”) para evaluar la comprensión lectora. Esto hace que, en definitiva, sea más variada en tipos textuales que la prueba SIMCE.

2.2 Análisis de los contextos de lectura

A. Descripción de los contextos de lectura posibles de ser encontrados en PISA y SIMCE

La distinción de contextos de lectura solo se encuentra en las pruebas PISA, donde se los define como un aspecto que caracteriza los textos en forma genérica, y que da cuenta de la intención de su uso. Por lo tanto es importante comprender que PISA no repara solo en



el tipo de texto, sino también, de manera igualmente relevante, en “el propósito con que se puede usar la información entregada por el texto” (SIMCE, 2004, pp.29-31).

PISA distingue cuatro contextos de lectura (OCDE, 2006):

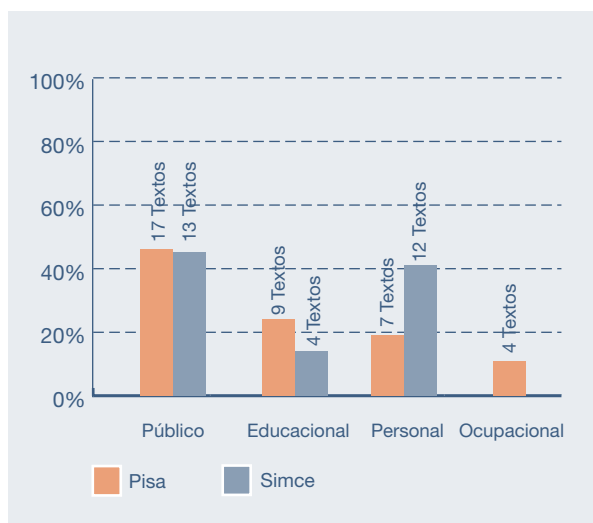
- a) Lectura para uso personal: textos cuya lectura permite satisfacer intereses personales, que se leen por curiosidad o recreación, principalmente.
- b) Lectura para uso público: textos cuya lectura permite participar en las actividades de la sociedad.
- c) Lectura para fines educacionales: textos generalmente no elegidos en forma autónoma. Su lectura permite recibir instrucción.
- d) Lectura para uso ocupacional: la lectura de estos textos está vinculada a la realización de una actividad. PISA incluye estos textos como una forma de evaluar a los jóvenes con respecto a su integración al mundo laboral.

Aunque en SIMCE no se distinguen contextos de lectura, sí es posible clasificar sus textos según la taxonomía propuesta en PISA.

B. Comparación de los contextos de lectura considerados en las pruebas PISA y SIMCE

Según puede verse en el Gráfico 6, los cuatro contextos de lectura postulados por el marco de evaluación de PISA aparecen efectivamente en la prueba, con predominancia del contexto público. En SIMCE, por su parte, no aparece el contexto ocupacional y la mayor presencia la tiene el público y el personal.

Gráfico 6: Comparación de los contextos de lectura de las pruebas PISA y SIMCE



En los contextos de lectura, SIMCE no trabaja con textos que se vinculen a una intención de uso ocupacional. PISA, por su parte, sí presenta textos de uso ocupacional y uso público, por lo que aparecen algunos textos principalmente “útiles” o “funcionales” como formularios, boletas, solicitudes, enfatizando una dimensión práctica, útil y funcional de la lectura.

2.3 Análisis de los formatos de respuesta

A. Descripción de los formatos de respuesta posibles de ser encontrados en PISA y SIMCE

El formato de las respuestas que dan los alumnos a los ítemes, si bien no está definido como una dimensión de la lectura en el marco de PISA, sí se plantea como un aspecto relevante. Por ello, se propone una clasificación que puede tener influencia en la tarea que supone cada ítem.

En PISA se definen cinco formatos de respuesta (SIMCE 2004):

- Selección múltiple simple (SMS): elección de una alternativa de respuesta entre varias.
- Selección múltiple compleja (SMC): elección de más de una respuesta. Por ejemplo, rellenar una tabla.
- Respuesta cerrada (RC): construcción de una respuesta, que tiene solo una posibilidad de correcta.
- Respuesta abierta corta (RAcor): producción de un texto breve, con más de una posibilidad de respuesta correcta, con posibilidad de recibir puntaje completo o parcial.
- Respuesta abierta compleja (RAcom): producción de un texto extenso, el cual es evaluado detenidamente por un corrector, y con posibilidad de recibir puntaje completo o parcial.

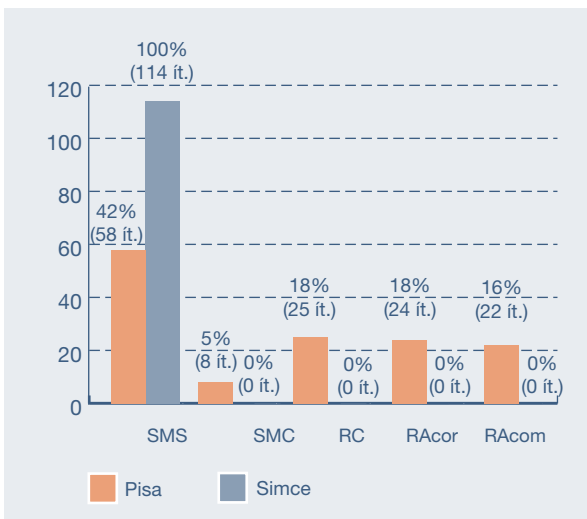
El formato de respuesta utilizado por SIMCE puede analizarse de acuerdo con el modelo propuesto por PISA, como se ve en el apartado siguiente.

B. Comparación de los formatos de respuesta incluidos en las pruebas PISA y SIMCE

Como puede verse en el Gráfico 7, las pruebas PISA utilizan todos los formatos que se definen en su Marco

de Evaluación, con predominio de los ítems de selección múltiple simple (SMS). Por su parte, todos los ítems que en SIMCE evalúan comprensión de lectura tienen este único formato de respuesta⁶.

Gráfico 7: Comparación de los formatos de respuesta de los ítems de las pruebas PISA y SIMCE



Desde la perspectiva de los formatos de respuesta, se puede observar que la prueba PISA presenta una mayor variedad de formatos de respuestas que la prueba SIMCE. De todas formas, el único formato de esta última (selección múltiple simple) es el predominante en PISA.

La diferencia antes expuesta puede ser más significativa que un mero asunto de forma. A la luz de estos datos, en SIMCE sería posible pensar que la presencia exclusiva del formato de selección múltiple en la evaluación de la lectura se conecta con la presencia dominante de la habilidad de *reconocer* (con representación de 46% en los ítems de esta prueba), ya que otros tipos de tareas de lectura pueden ser más complejos de evaluar con este único formato. Así, la variedad de

⁶ En SIMCE, el formato de respuesta abierta compleja solo se usa para evaluar producción escrita, por lo que no se incluye en esta comparación de pruebas de Lectura.

formatos de PISA, sumada a la mayor diferenciación de textos, puede relacionarse con la mayor cantidad de procedimientos del MCA que incorpora esta prueba. Además, el uso de formatos complejos es coherente con la presencia de procedimientos combinados, fenómeno que se presenta en PISA y no en SIMCE. En definitiva, esta diferencia de nivel estructural tiene implicancias sustantivas para las cualidades de una y otra evaluación.

2.4 Análisis de los procedimientos y contenidos conceptuales

A. Descripción de los procedimientos y contenidos conceptuales posibles de ser encontrados en PISA y SIMCE

En el Marco Curricular Analítico (MCA), cuyo detalle puede verse en el Anexo, se distinguen 37 procedimientos y 75 contenidos conceptuales.

Del análisis de ambas pruebas, surgieron además algunos procedimientos y contenidos de PISA no contenidos en el MCA, pues, como se mencionó al comienzo, este se basa en el Marco Curricular chileno vigente (MINEDUC, 1998).

A continuación, se compara la distribución efectiva de los procedimientos y contenidos conceptuales en las pruebas SIMCE y PISA.

B. Comparación de los procedimientos evaluados por los ítems de las pruebas SIMCE y PISA

En las Tablas 2 y 3 se presentan los procedimientos del MCA que se encuentran en las pruebas SIMCE y PISA, y la cantidad de ítems asociados a cada uno.

Tabla 2.
Procedimientos en PISA y SIMCE

Procedimiento	Cantidad de ítems		Porcentaje de ítems	
	SIMCE	PISA	SIMCE	PISA
Procedimiento no contemplado en el MCA	0	7	0%	5,10%
Reconocer	52	32	45,61%	23,35%
Opinar	0	11	0%	8,02%
Seleccionar	1	1	0,87%	0,72%
Identificar	3	34	2,63%	24,81%
Diferenciar	1	0	0,87%	0%
Evaluar	13	3	11,40%	2,18%
Interpretar	37	21	32,45%	15,32%
Comparar	0	2	0%	1,45%
Relacionar	5	12	4,38%	8,75%
Apreciar	2	0	1,75%	0%
Reflexionar	0	2	0%	1,45%

Tabla 3.
Ítems con dos procedimientos (solo en PISA)

Procedimientos	Cantidad de ítems	Porcentaje de ítems
Identificar/Interpretar	3	2,18%
Identificar/Comparar	1	0,72%
Identificar/Relacionar	4	2,91%
Interpretar/Relacionar	1	0,72%
Comparar/Interpretar	2	1,45%
Relacionar/Interpretar	1	0,72%

Respecto de los procedimientos que considera cada prueba, PISA cubre un procedimiento más que SIMCE (nueve contra ocho), y algunos procedimientos evaluados por PISA (presentes en siete ítems) no se encuentran en el MCA, como por ejemplo *formular una hipótesis* sustentable en el texto u *organizar* la información del mismo en una tabla. Una lectura de este dato podría concluir que PISA considera una

mayor variedad de aplicaciones de la competencia lectora, por lo que diversifica sus alcances y evalúa un espectro más amplio de la misma.

La prueba SIMCE se concentra principalmente en los procedimientos lectores de *reconocer* e *interpretar* y, en menor medida, en los de *evaluar* y *relacionar*, observándose en general una reducida gama de procedimientos.

Por último, cabe destacar que, como se muestra en la Tabla 3, en más de un caso, los ítems de PISA suponen una respuesta elaborada por etapas: son ejercicios que constan de dos partes. Es decir, el estudiante tiene que aplicar primero un procedimiento y luego otro (primero *identificar*, luego *interpretar*, por ejemplo) para cumplir la tarea de lectura que se le propone. Esto evidentemente involucra un proceso más complejo en el ejercicio de la competencia lectora. SIMCE no considera respuestas por etapas, por lo que sus preguntas apuntan a un resultado final del proceso lector, sin distinción de etapas previas.

C. Comparación de los contenidos conceptuales evaluados en los ítems de las pruebas SIMCE y PISA

En las siguientes tablas, se muestran los contenidos conceptuales⁷ del MCA que están presentes en las pruebas SIMCE y PISA, y la cantidad de ítems asociados a cada uno.

⁷ Hemos considerado "contenido conceptual" todos aquellos términos que aparecen en el Marco Curricular como tales. Aunque la idea de "mensaje" puede también pertenecer a un espectro semántico cotidiano, la decisión de análisis nos lleva a considerarlo igualmente como contenido conceptual.

Tabla 4.
Contenidos conceptuales en PISA y SIMCE

Contenido conceptual	Cantidad de ítems		Porcentaje de ítems	
	SIMCE	PISA	SIMCE	PISA
Contenido conceptual no considerado en el MCA	5	103	3,64%	75,18%
Efectos comunicativos	4	0	2,91%	0%
Propósitos del emisor	4	3	2,91%	2,18%
Modalizaciones discursivas	5	0	3,64%	0%
Formas discursivas adecuadas	4	2	2,91%	1,45%
Niveles de habla	1	0	0,72%	0%
Receptor	4	1	2,91%	0,72%
Mensaje	28	9	20,43%	6,56%
Tipos de texto	1	0	0,72%	0%
Propósitos de un texto	6	8	4,37%	5,83%
Modos de tratar los temas	0	1	0%	0,72%
Estructuras textuales	1	0	0,72%	0%
Organización interna de los textos	3	0	2,18%	0%
Principios de composición de un texto	1	0	0,72%	0%
Recursos de composición de un texto	2	4	1,45%	2,91%
Elementos de composición de un texto	2	0	1,45%	0%
Estructuras gramaticales	1	0	0,72%	0%
Recursos verbales	3	0	2,18%	0%
Hechos	6	0	4,37%	0%
Opiniones	2	5	1,45%	3,64%
Tipos de mundos literarios	2	0	1,45%	0%
Elementos básicos del mundo literario	11	0	8,02%	0%
Géneros	1	0	0,72%	0%
Elementos distintivos de las obras literarias	2	0	1,45%	0%
Elementos constitutivos de las obras literarias	0	1	0%	0,72%
Similitudes y diferencias entre medios	3	0	2,18%	0%
Elementos de medios masivos	3	0	2,18%	0%
Recursos de medios masivos	2	0	1,45%	0%
Propósitos de medios masivos	5	0	3,64%	0%
Versiones de la información	2	0	1,45%	0%



En contraste con lo que sucede con los procedimientos, es posible observar que la prueba SIMCE considera más contenidos conceptuales o propiamente disciplinarios que la prueba PISA. En nuestro análisis, fue posible constatar que la prueba SIMCE considera veintisiete contenidos conceptuales presentes en el Marco Curricular chileno, mientras que la prueba PISA solo considera nueve. Además, la prueba SIMCE presenta la mayor parte de sus ítemes ligados a algún contenido conceptual del Marco Curricular, y la presencia de algunos de ellos, como mensaje de un texto o elementos básicos del mundo literario, se aprecia en veintiocho y once ítemes respectivamente. Es decir, muchos ítemes repiten los contenidos conceptuales.

Por su parte, la mayor parte de los ítemes de PISA no considera contenidos conceptuales disciplinarios presentes en el MCA, sino que evalúa principalmente procedimientos. Esto no quiere decir que los ítemes PISA carezcan de contenidos, sino solamente que estos no están presentes en el Marco Curricular chileno.

Muchos de los contenidos que sí están presentes aparecen solo una vez en la prueba PISA (por ejemplo receptor, o modo de tratar temas) y ninguno de ellos aparece en más de nueve ocasiones (los contenidos con mayor representatividad son mensaje y propósito de un texto, con nueve y ocho apariciones, respectivamente).

Esto lleva a considerar la naturaleza de los contenidos que aborda una y otra prueba, en la que es posible encontrar diferencias: así, mientras la prueba SIMCE involucra el manejo de algunos contenidos conceptuales de la disciplina de manera exclusiva⁸, que se nominan preferentemente por medio de metalenguaje (párrafo, tipo de mundo literario, tipos de texto, etc.), la prueba PISA evita los tecnicismos o el lenguaje específico y prefiere parafrasear estos contenidos mediante formulaciones menos técnicas (numera e indica rango de líneas que hay que leer, o se refiere a “parte del cuento” en vez de párrafo, por ejemplo), solo abordando contenidos conceptuales estrictamente relacionados con la comprensión. Esto facilita el acceso del estudiante a poder contestar la prueba, sin que ello esté condicionado por el grado de dominio disciplinario que ha alcanzado en 1° o 2° Medio.

En síntesis, la diferencia entre ambas pruebas se observa en la diversidad de contenidos que se consideran en la evaluación, la presencia de los mismos en la cantidad total de ítemes, y la naturaleza de los contenidos.

Para ilustrar las diferencias más marcadas posibles de encontrar entre ambas pruebas, se presentan los siguientes ejemplos.

⁸ Un ejemplo ilustrador de un contenido conceptual del Marco Curricular, evaluado en la prueba SIMCE y que no es considerado en PISA, es el de textos expositivos que señala: “Lectura de textos expositivos, para percibir: a) la variedad de tipos de textos expositivos; b) los rasgos relacionados con factores determinantes de la situación de comunicación; c) la estructura global y la organización interna de este tipo de textos; d) fórmulas y recursos verbales y no verbales en los textos expositivos” (MINEDUC, 1998). Con respecto a los contenidos de Literatura, también es posible encontrar algunos ejemplos ilustradores: “Lectura de un mínimo de seis obras literarias de diversas épocas, culturas y géneros, en las que se configuren mundos literarios de diversos tipos (cotidiano, fantástico, onírico, mítico, utópico, marginal, etc.) dando oportunidad para: a) la apreciación de la capacidad de la literatura de crear múltiples mundos posibles y el valor de ella para el conocimiento y comprensión de la diversidad humana, así como de la permanencia de valores estéticos y humanos universales; b) la comparación de los mundos creados en las obras leídas con el mundo en que vivimos, con los que se manifiestan en otras obras artísticas y con las imágenes que entregan los medios masivos de comunicación y la publicidad, percibiendo las similitudes y diferencias y proponiendo explicaciones para ellas; c) el reconocimiento de los elementos que constituyen el mundo literario y la comparación de obras de distintos géneros para percibir lo distintivo del mundo narrativo, del mundo lírico, del mundo dramático” (MINEDUC, 1998).

TEXTO SIMCE: “LA SOMBRA DEL HUÉSPED”

LA SOMBRA DEL HUÉSPED

Aquella noche sostenía con mi huésped una interesante conversación. Advertí que estaba preocupado; su mirada, dirigida hacia un ángulo de la habitación, manifestaba cierta angustia, pero como su sombra daba en ese punto, no descubrí nada extraño.

Aprovechando un silencio, le pregunté por qué había dejado la milicia.

—¿Acaso enfermó usted?

—Sí, respondió con tristeza, clavando sus ojos en un rincón del aposento. —Fue algo nefasto. Un misterio. Pronto hará cuarenta años. Si lo contara, me creerían loco. En aquel tiempo, perdí toda mi alegría. Habrá oído usted hablar de los *yoghis*, a quienes se les atribuye poderes prodigiosos, ¿sabe usted en qué consisten?

—Creo que en provocarse el autosonambulismo, volviéndose así insensibles, videntes.

—Exactamente. Cuando los vi actuar, quise desarrollar idénticas facultades. Puse manos a la obra, sin pensar las consecuencias. Prodigiosamente, logré ciertas capacidades. Pude dormirme y manejar mi voluntad en ese estado. De a poco, alcancé el desdoblamiento. Por curiosidad, una noche resolví ver mi doble. Ver qué era lo que salía de mí. Lo hice. Cuando recuperé la conciencia, vi ante mí una forma al fondo de la habitación. Era la de un mono, un animal que me miraba fijamente. Desde entonces no se aparta de mí. Lo veo constantemente. Siempre está ahí, mirándome, sin hacer nada. Es negro y melancólico. Su cara es como todas las caras de mono. No obstante, siento que se parece a mí.

Notando mi perplejidad, se puso de pie: —Voy a caminar para que usted lo vea. Observe mi sombra.

Entonces, tuve la más grande de las sorpresas. ¡La sombra de aquel sujeto no se movía! Sintiendo la condena que sufría, tomé un papel y lo puse sobre la mancha. Con un lápiz describí la silueta de mi amigo. Cuando terminé, ambos palidecimos horriblemente. La raya trazada por mi mano describía una frente hundida, una nariz chata, un hocico bestial.

Leopoldo Lugones
(adaptación)

Para este texto⁹, además de preguntas como reconocimiento de tema o interpretación de determinados fragmentos, se presentó la siguiente pregunta:

¿Qué tipo de mundo muestra el final del texto?

- A. Un mundo fantástico.
- B. Un mundo incomunicable.
- C. Un mundo cotidiano.
- D. Un mundo mítico.

En ella se deben aplicar conocimientos conceptuales —conceptos, definiciones y características de los tipos de mundo en literatura— que permiten discriminar entre las alternativas, y seleccionar la que corresponde al tipo de mundo del texto en cuestión.

Por su parte, como se señaló, la prueba PISA se focaliza en evaluar exclusivamente la competencia lectora, definida desde una perspectiva funcional y amplia, es decir, como una herramienta útil para el individuo, que le permite aplicar una serie de conocimientos más bien prácticos para actuar convenientemente en distintas tareas, propias de quien ha alcanzado un grado de alfabetización que lo hace competente para las demandas habituales de la sociedad contemporánea.

⁹ Dado que no es posible publicar ítemes no liberados, se presenta este texto que no forma parte del corpus analizado, pero es muy similar, en tema y estructura, a los que sí formaron parte.

En efecto, la competencia lectora evaluada por PISA se desprende del concepto de “alfabetización en lectura”, entendida como “la comprensión, el uso y la reflexión sobre textos escritos, con el fin de alcanzar las metas propias, desarrollar los propios conocimientos y potencialidades y participar en la sociedad” (OCDE, 2000).

Se puede interpretar que este concepto se relaciona, a su vez, con el concepto de “alfabetización funcional”, el cual, desde sus primeros planteamientos realizados por William Gray en 1957, ha ido evolucionando en los términos que explica Berta Braslavsky (2003):

El concepto de alfabetización funcional evolucionó, y en el año 1978 la UNESCO definió al “alfabetizado” como “alguien que puede comprender o escribir un texto corto y breve de su vida diaria” (Ryan, 1995). Para ser “alfabetizado funcional” [el individuo] debía ser capaz de “intervenir en todas las actividades en las cuales la alfabetización es requerida para el efectivo funcionamiento de su grupo en la comunidad y también para capacitarlo en la continuidad del uso de la lectura, la escritura y el cálculo para sí mismo y para el desarrollo de la comunidad”. Era una definición esencialmente relativa, ya que depende del contexto social, al incluir dos condiciones: 1) la habilidad de la lectura y la escritura, y que 2) ese dominio permita cumplir con las exigencias de la

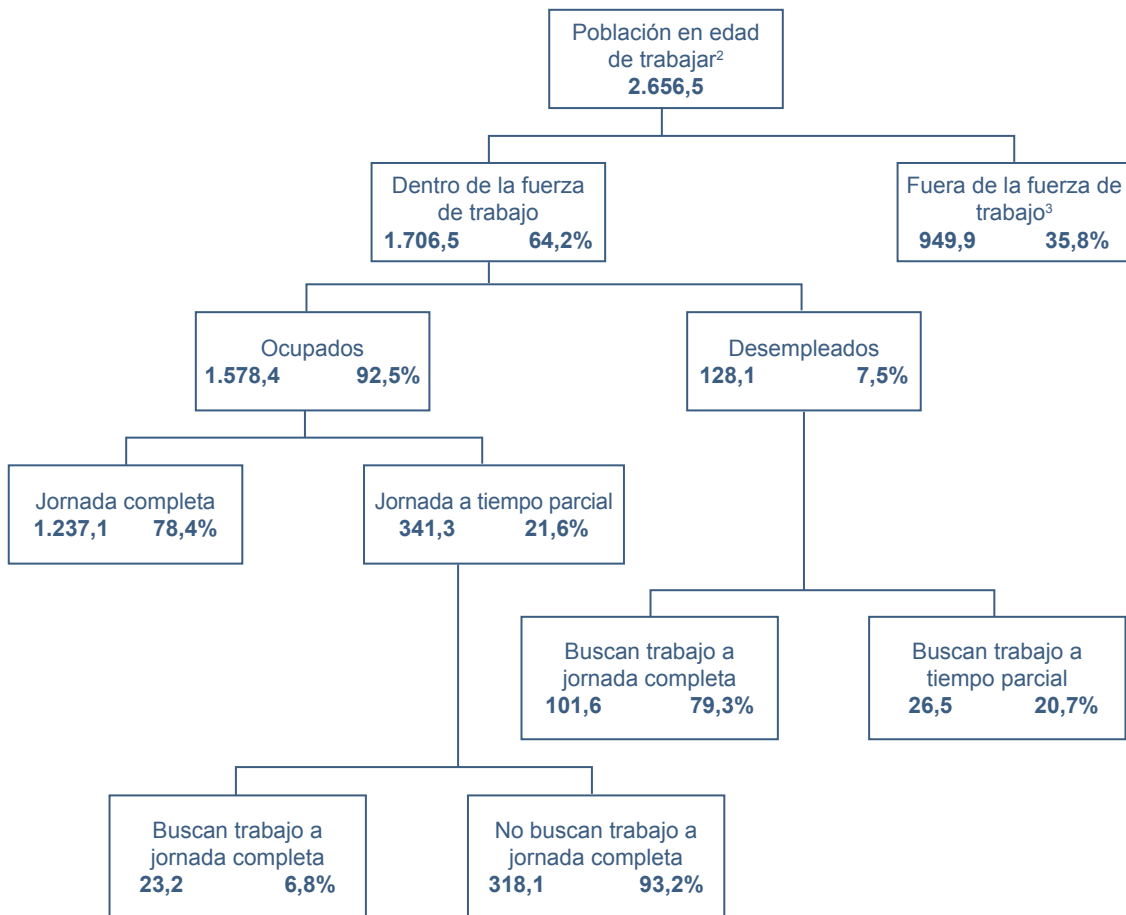
comunidad. No se trata de la importancia que tiene la alfabetización en cuanto a la relación personal, subjetiva, del alfabetizado con el texto, sino en cuanto le sirve a la persona para ser eficaz en el trabajo, consciente en el voto, responsable en todos sus comportamientos y en las transacciones con su medio. Es decir, en todas las actividades que se relacionan con la vida pública en la adultez.

Por lo tanto, pese a que en PISA el enfoque que se le da a la competencia lectora incluye más aplicaciones cotidianas y prácticas de la lectura que en SIMCE, el dominio que evalúa PISA es, desde cierto punto de vista, más reducido, por específico: PISA es una prueba exclusivamente de comprensión lectora, aplicada en contextos de lectura de preferencia prácticos. A las preguntas de comprensión lectora tradicionales, del tipo “¿Qué pasó en tal situación?”, o “¿Cuántos gramos de harina debe tener tal receta?”, se agregan otras como evaluar la utilidad de disponer información en un tipo de diagrama determinado (por ejemplo, la utilidad de poner ciertos datos en un diagrama de árbol), trabajar con gráficos reconociendo sus partes variables e invariables, reconocer el propósito de un cuadro dentro de un texto, o clasificar la información de un texto en una tabla, a partir de un criterio dado. Todas estas formas de preguntar implican un ejercicio amplio de la competencia lectora, sin requerir del manejo o apropiación de contenidos específicos, lenguaje disciplinario o metalenguaje para ser respondidas, como se observa en el siguiente ejemplo.

TEXTO PISA: “ESTRUCTURA DE LA FUERZA DE TRABAJO”

El esquema de árbol que aparece a continuación, muestra la estructura de la fuerza de trabajo de un país o la “población en edad de trabajar”. La población total de este país en 1995 era aproximadamente de 3,4 millones.

Estructura de la fuerza de trabajo al 31 de marzo de 1995
(en miles de personas)¹



Notas

1. Las cifras referentes a la población se expresan en miles de personas (x 1.000).
2. La población en edad de trabajar se define como las personas entre 15 y 65 años.
3. Se considera que “No forman parte de la fuerza de trabajo” a aquellos que no buscan trabajo activamente y/o que no están dispuestos a trabajar.

Para este texto, se presentó la siguiente pregunta en formato de respuesta selección múltiple compleja:

¿En qué parte del esquema de árbol, si corresponde, incluirías a cada una de las personas de la lista que aparece a continuación?

Marca con una cruz en el casillero correspondiente de la tabla.

La primera está contestada, a modo de ejemplo.

	“Dentro de la fuerza de trabajo: ocupados”	“Dentro de la fuerza de trabajo: desocupados”	“Fuera de la fuerza de trabajo”	No incluido en ninguna categoría
Un camarero a tiempo parcial de 35 años	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Una empresaria de 43 años, que trabaja 60 horas a la semana	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Un estudiante de 21 años	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Un hombre de 28 años, que acaba de vender su tienda y está buscando trabajo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Una mujer de 55 años, que nunca ha trabajado ni ha querido trabajar fuera del hogar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Una abuela de 80 años, que aún trabaja algunas horas al día, en el puesto de la familia en el mercado	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

La aplicación de la habilidad lectora que hay detrás de esta pregunta rehúye la aplicación de contenidos conceptuales como la pregunta de SIMCE, y más bien apunta a saber trabajar con la información recibida con una finalidad eminentemente práctica. Se trata de un trabajo complejo de aplicación de la información recibida a diferentes situaciones —en este caso, perfiles laborales— para clasificarlas.

3. Conclusiones

Las pruebas SIMCE y PISA tienen una indiscutible base común de carácter general: ambas están relacionadas por tratarse de evaluaciones de la Lectura. Más allá de esta relación general, subyacen diferencias significativas en cuanto al objetivo de la evaluación y a la aplicación de la misma. La comparación realizada en este estudio ha permitido caracterizar cinco tipos de diferencias de forma de los instrumentos y de alineación con el currículum chileno:

- diferencias en las características de los tipos textuales
- diferencias en las características de los contextos de lectura
- diferencias en los formatos de las respuestas a los ítemes
- diferencias en la cantidad de procedimientos de lectura que cubre cada prueba
- diferencias en la cantidad de contenidos conceptuales que abarca cada prueba

Todo esto llevó a concluir que, en definitiva, las pruebas tienen una diferencia de enfoque y aplicación sustantiva. En efecto, ambos instrumentos poseen una diferencia de principio: SIMCE, si bien no de manera predominante, considera una mayor variedad de aspectos propios del área de Lenguaje y Comunicación, y PISA, por su parte, se concentra en la evaluación de aplicaciones variadas para la habilidad de comprensión de lectura y ejecuta esta evaluación cuidando de hacerlo de diferentes maneras (formatos) y en situaciones de lectura diversas (contextos).

La explicación para esta diferencia puede radicar en que, como ya se señaló, la prueba SIMCE de 2° Medio es una prueba que evalúa el Marco Curricular chileno (MINEDUC, 1998) en el subsector Lengua Castellana y Comunicación. Esta prueba se focaliza en lo que es posible evaluar con una prueba de lápiz y papel, y considera los contenidos y habilidades presentes en las “tareas de desarrollo” que conforman el subsector (Comunicación oral, Comunicación escrita, Literatura y Medios masivos de comunicación). Exceptuando Comunicación oral, que no es posible de abordar en una prueba de las características de SIMCE, los otros tres ejes sí son evaluados en esta prueba.

El mismo rótulo de estas “tareas de desarrollo” de Educación Media, centradas en temas de literatura y comunicación y medios, indica una organización que tiende a darse en función de contenidos conceptuales de la disciplina y no de habilidades transversales como

es la lectura. En ese escenario, es imprescindible reparar que en nuestro Marco Curricular¹⁰ la comprensión lectora no constituye un eje en sí mismo, sino que es una habilidad que se reparte entre las “tareas de desarrollo”. Es por eso que en SIMCE coexisten preguntas de lectura “pura” (formulación de síntesis, extracción de datos discretos o interpretación a partir de información), con ítemes de aplicación de contenidos conceptuales disciplinarios presentes en el currículum (identificación de formas discursivas básicas, organización textual e identificación de tipos de mundo o tipos de efectos de la literatura).

Lo anterior permite observar que el marco de evaluación de la prueba PISA coincide solo parcialmente con la enseñanza regular establecida en nuestro Marco Curricular para un estudiante hasta 2° Medio. Obviamente, la habilidad lectora, eje central de la evaluación de PISA, también es parte del Marco Curricular chileno, pero PISA presenta un enfoque más definido, considerándola una competencia transversal y que comprende la lectura funcional como habilidad práctica, muchas veces con fines ocupacionales, que permite el desarrollo de la persona en la sociedad. Prueba de esto es lo observado a propósito de los textos-estímulo utilizados, en que se prefieren los no literarios de variados formatos.

Aunque la habilidad de lectura que se desprende del Marco Curricular chileno también busca plantearse como una competencia, está permeada por la presencia de aspectos relevantes en el currículum: la presencia considerable que tiene la literatura o la importancia que se le da a la adquisición de aspectos técnicos de la disciplina, intencionan paralelamente la enseñanza del lenguaje como una disciplina de campo relativamente específico. Estos aspectos, al ser propios y distintivos del subsector de Lengua Castellana y Comunicación, son considerados en la evaluación SIMCE.

¹⁰ Esta referencia considera la orientación que se desprende del marco curricular vigente hasta la fecha, y que coincidió para las aplicaciones de 2001 y 2006. Por lo tanto, no considera la orientación que puede tomar la lectura en el ajuste curricular en desarrollo (2009).

En conclusión, las pruebas SIMCE y PISA no son equivalentes en cuanto al constructo que pretenden evaluar. Esta diferencia se observa en la forma que tiene cada prueba para evaluar la comprensión de lectura y en las diferencias que se observan al ser puestas en moldes comunes.

4. *Discusión*

Ahora bien, el incremento en la puntuación de nuestros estudiantes en la prueba PISA, visto desde el prestigio del que goza globalmente esta como instrumento de evaluación, podría suscitar ideas e inquietudes respecto del camino a seguir para las evaluaciones chilenas. En este sentido, es indudable que el impacto de la evaluación que realiza PISA, generando datos que orientan globalmente decisiones y políticas educativas en los distintos países en que se aplica, da pie para ello.

En el marco de este estudio, la aproximación comparativa ha permitido ver la prueba PISA como un instrumento muy variado en aspectos formales y de contenido. Esta diversidad va en beneficio de la precisión con la que PISA evalúa la lectura como competencia transversal, lo que se transparenta en un marco de evaluación bien delineado, que tiene un propósito claro y único y que, a diferencia de la prueba SIMCE, no tiene el objetivo vasto de medir en determinado grado de la escolarización, el desempeño en una disciplina de múltiples alcances como es el subsector Lengua Castellana y Comunicación. En síntesis, la prueba PISA tiene más delineado su objeto a evaluar: es más específico y por lo mismo se puede medir con más precisión. Si a esto se suma la mayor cantidad de recursos¹¹, se confirma la impresión de que esta prueba internacional, en cuanto instrumento, podría señalar algunos rumbos.

En relación con esto, podría afirmarse que la prueba SIMCE se podría apropiarse de una mayor diversidad de tipos textuales, formatos de respuesta y contextos de lectura al modo de PISA, lo que desembocaría, a nuestro entender, en un instrumento con más posibilidades de precisión en evaluar lo que se propone, sin que por ello fuera necesario cambiar el currículo ni el constructo evaluado.

Dicho lo anterior, cabe hacerse la siguiente pregunta: ¿hasta qué punto es posible o resultaría conveniente transformar la prueba de 2° Medio en una prueba similar a PISA? La pregunta es pertinente al reparar en que la prueba SIMCE está sujeta a la evaluación de un Marco Curricular vigente. Por ello no resulta fácil decidirse a ir en la dirección de PISA sin evaluar previamente si esa decisión no redundaría en una reducción curricular.

Más allá de esto, desde el análisis del currículo vigente es posible afirmar que el concepto de lectura que subyace en nuestro Marco Curricular se distancia de una definición equivalente a la de PISA. Por ejemplo, sería posible señalar que el enfoque en Enseñanza Media está más centrado en la comprensión y el análisis crítico de los mensajes provenientes de los medios masivos de comunicación que en la lectura funcional y ocupacional, y como tal, es entendible que aparezcan textos publicitarios en vez de solicitudes en la prueba. En definitiva, sería necesario considerar que la lectura en el Marco Curricular chileno supone una serie de contenidos conceptuales, procedimentales y valorativos que son parte de una perspectiva idiosincrática de la enseñanza del lenguaje en Educación Media y que conforman un perfil de egreso para los estudiantes en función de objetivos fundamentales que comprenden y proyectan tal vez algo más —o tal vez algo distinto— que un alfabetizado funcional. En definitiva, si la prueba SIMCE desea avanzar en medir el currículo nacional de mejor forma, tal vez puede tomar en cuenta los modelos internacionales validados de evaluación, pero sin obviar estas diferencias detectadas y otras que puedan encontrarse.

11 Estos facilitan decisiones técnicas como la elección del formato más adecuado para evaluar una tarea de lectura determinada, aunque esto signifique, por ejemplo, mayor cantidad de formatos de preguntas abiertas, que implican un operativo más costoso para su corrección.

Inicialmente, esta investigación se propuso realizar un estudio comparativo que permitiera proponer una explicación tentativa a la diferencia de resultados entre los estudiantes que rindieron PISA y SIMCE en 2001 y 2006. Un esbozo de respuesta, a la luz de estos resultados, podría ser el siguiente: PISA recoge un ejercicio de la lectura que se ha potenciado en los últimos años con la masificación de las tecnologías de la información: este es el de una lectura que excede el ámbito estrictamente escolar y se ramifica a aplicaciones y utilidades múltiples, necesarias para la participación activa en la sociedad. Asimismo, este tipo de lectura, requiere de un sujeto activo y versátil, capaz de situarse en diferentes “escenarios de lectura” y capaz de utilizar diferentes tipos de información¹². Por cierto, no sería raro que en medio de las voces que llaman la atención respecto de la crisis de la lectura —manifiesta por ejemplo en los índices deficientes de demanda de libros— los lectores actualmente estén leyendo más y tal vez mejor, pero nuevos contenidos expresados en nuevos formatos. La idea que podría resumir esa ejercitación actual de la lectura es la de una actividad funcional con fines específicos, ejecutada para satisfacer necesidades inmediatas y eminentemente prácticas y, en este sentido, concordante con la que evalúa PISA.

En términos de proyección de este estudio, se podría realizar una investigación psicométrica que construya una subescala con los ítemes de la prueba SIMCE que corresponden exclusivamente a comprensión de lectura sin aplicación de contenidos conceptuales de la disciplina, vale decir, los ítemes que se acercan más a la forma de evaluar de PISA, y comparar los resultados entre 2001 y 2006 para ver si en ellos se observa la misma o similar variación significativa que muestra la prueba internacional. Asimismo, siempre desde un análisis psicométrico de construcción de subescalas, sería interesante investigar qué aspectos de la comprensión lectora recogidos en algunos

ítemes de PISA y no evaluados por el SIMCE pueden explicar el incremento significativo en el puntaje. Por último, se podría comparar a sujetos que rindieron la misma prueba (*equating* entre sujetos), para ver si existe la posibilidad de desplegar mejor la habilidad de comprensión lectora en una prueba que en otra. Vale decir, verificar si la prueba nacional implica un límite o “techo” en la demostración de la habilidad, que no existiría en PISA, prueba en la que podrían existir más posibilidades de desplegar la misma. En este sentido, la mayor cantidad de procedimientos observados en PISA, unida a las preguntas que combinan habilidades, podrían explicar por qué un sujeto hábil puede demostrar mejor su competencia lectora en esa prueba.

En todo caso, la pregunta de fondo sería cuán satisfechos nos deja el que nuestros estudiantes estén progresando en la adquisición de una lectura funcional, pero contrariamente a este logro, muestren estancamiento en la medición de otros aprendizajes intencionados en nuestro currículum nacional, sobre todo si esta tendencia vuelve a observarse en la aplicación de PISA 2009. Responder esta pregunta podría acarrear varias consecuencias, desde tomar decisiones sustantivas sobre el enfoque curricular de la enseñanza del lenguaje, hasta lograr una mayor conciencia sobre cuáles son los fundamentos idiosincrásicos que sustentan nuestra enseñanza regular del mismo. De todas maneras, el plantearnos seriamente esta pregunta parece un buen comienzo.

¹² Para una exposición clara y concisa de la naturaleza, finalidades y aplicaciones de la lectura en el proyecto PISA, cf. Sanz Moreno, 2005.

Referencias

- Braslavsky, B. (2003). "¿Qué se entiende por alfabetización?". *Lectura y Vida, Revista Latinoamericana de Lectura*, 24 (2). Buenos Aires, Argentina: Asociación Internacional de Lectura.
- MINEDUC (Ministerio de Educación) (1998). *Objetivos Fundamentales y Contenidos Mínimos Obligatorios para la Formación General en la Educación Media, Decreto 220*. Santiago, Chile.
- OCDE (Organización para la Cooperación y el Desarrollo Económico) (2000). *Measuring Student Knowledge and Skills. The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy*. París, Francia.
- _____. (2006). *PISA 2006. Marco de la evaluación. Conocimientos y habilidades en Ciencias, Matemáticas y Lectura*. París, Francia.
- OCDE-INCE (Instituto Nacional de Calidad y Educación) (2000). *La medida de los conocimientos y destrezas de los alumnos. Un nuevo marco de evaluación*. Madrid, España: Ministerio de Educación, Cultura y Deporte.
- Sanz Moreno, A. (2005). "La lectura en el proyecto PISA". *Revista de Educación*. Extraído desde www.oei.es/evaluacioneducativa/lectura_proyecto_pisa_sanz.pdf
- SIMCE (Sistema de Medición de la Calidad de la Educación) (2001). *Orientaciones para la medición, Segundo año de Educación Media*. Santiago, Chile: MINEDUC.
- _____. (2004). *Competencias para la vida. Resultados de los estudiantes chilenos en el estudio PISA 2000*. Santiago, Chile: MINEDUC.
- Valverde, G. (2005). "Marco Analítico de Comprensión Lectora", documento de trabajo preparado para Consorcio de Evaluación e Investigación Educativa CEIE. República Dominicana.

Anexo

1. Marco Curricular Analítico 1º y 2º Medio, Marco Curricular de Educación Media (Lectura)

Procedimientos (A)

1. interactuar
2. comprender
3. reconocer
4. utilizar
5. incrementar
6. afianzar
7. fortalecer interés
8. fortalecer gusto
9. relacionar (con diversas manifestaciones)
10. descubrir
11. opinar
12. expresarse (creatividad)
13. explorar (propia creatividad escribiendo)
14. seleccionar
15. recepcionar
16. identificar
17. diferenciar
18. evaluar
19. leer
20. percibir
21. producir textos
22. producir medios
23. interpretar
24. crear
25. comparar
26. investigar
27. crear (escribiendo)
28. tomar apuntes
29. identificar (percibir)
30. reflexionar sobre el lenguaje
31. explicar
32. relacionar
33. representar
34. apreciar (observar)
35. reflexionar

36. producir (medio)
37. expresar creatividad a través de un medio

Contenidos conceptuales (B)

1. procesos de comunicación
2. comunicación oral
3. exposición oral
4. estrategias para captar y mantener atención
5. procedimientos para apuntar
6. actos de habla
7. estructuras básicas de la comunicación
8. situaciones de comunicación
9. factores de la situación de comunicación
10. rasgos asociados a factores de la comunicación
11. eficacia comunicativa
12. factores de eficacia comunicativa
13. efectos comunicativos
14. discursos
15. propósitos del emisor
16. modalizaciones discursivas
17. formas discursivas adecuadas
18. niveles de habla
19. situación
20. receptor
21. relaciones de simetría
22. relaciones de complementariedad
23. mensaje
24. tipos de mensaje
25. tipos de texto
26. propósitos de un texto
27. finalidad de un texto
28. modos de tratar temas
29. estructuras textuales
30. organización interna de los textos
31. principios de composición de un texto
32. recursos de composición de un texto
33. elementos de composición de un texto
34. léxico
35. estructuras gramaticales
36. elementos de gramática oracional y textual
37. recursos verbales
38. recursos no verbales
39. elementos paraverbales

40. ortografía
41. texto informativo
42. texto expositivo
43. discurso expositivo
44. tipos de texto expositivo
45. hechos
46. opiniones
47. aportes
48. obras literarias
49. tipos de mundos literarios
50. elementos básicos del mundo literario
51. imágenes de mundo de la literatura
52. géneros
53. rasgos de género
54. épocas
55. culturas
56. manifestaciones de la cultura
57. diversidad humana
58. valores estéticos
59. movimientos artísticos
60. tendencias artísticas
61. contexto histórico-cultural
62. contextos
63. visiones de mundo
64. elementos distintivos de las obras literarias
65. elementos constitutivos de las obras literarias
66. investigación literaria
67. medios masivos
68. similitudes y diferencias entre medios
69. funciones de medios masivos
70. elementos de medios masivos
71. recursos de medios masivos
72. propósitos de medios masivos
73. versiones de la información
74. transmisión de la información
75. textos de difusión radial o televisiva

Perspectivas (C)

1. valorar
2. valorar la lectura
3. dar importancia
4. apreciar
5. valorar la diversidad

2. Ejemplos de la descripción que se hizo de los textos e ítemes en este estudio

A. PISA

a) Texto: Las armas científicas de la policía

Las armas científicas de la policía

Se ha cometido un asesinato, pero el sospechoso niega todo. Afirma no conocer a la víctima. Dice que nunca lo vio, que nunca estuvo cerca de él, que nunca lo tocó... La policía y el juez están convencidos de que no está diciendo la verdad. Pero ¿cómo probarlo?

En la escena del crimen, los investigadores han reunido hasta la más mínima evidencia: fibras de tela, cabellos, huellas dactilares, colillas de cigarrillos... Los pocos cabellos encontrados en la chaqueta de la víctima son pelirrojos. Y se parecen extrañamente a los del sospechoso.

Si se pudiera probar que estos cabellos son realmente suyos, sería una evidencia de que él efectivamente conocía a la víctima...

Cada persona es única

Los especialistas se pusieron a trabajar. Examinaron algunas células de la raíz de esos cabellos y algunas células de la sangre del sospechoso. En el núcleo de cada célula de nuestro cuerpo se encuentra el ADN. ¿Qué es eso? El ADN es como un collar formado por dos corridas de perlas entrelazadas. Imagine

que estas perlas son de cuatro colores diferentes y que miles de estas perlas de colores (que forman un gen) están dispuestas en un orden muy específico. El orden es exactamente el mismo en todas las células del cuerpo de un individuo; tanto en las células de las raíces del cabello como en las del dedo gordo del pie, en las del hígado, como en las del estómago o de la sangre. Pero el orden de las perlas varía de una persona a otra. Dado el número de perlas dispuestas de este modo, hay muy pocas probabilidades de que haya dos personas con el mismo ADN, a excepción de los gemelos idénticos. Como el ADN es único para cada individuo, es una especie de carné de identidad genético.

Por lo tanto, los genetistas son capaces de comparar el carné de identidad genético del sospechoso (determinado por su sangre) con el de la persona

pelirroja. Si el carné genético es el mismo, se sabrá que el sospechoso en realidad sí estuvo cerca de la víctima, a quien según él, nunca había conocido.

Sólo una prueba

En casos de agresión sexual, asesinato, robo u otros delitos, la policía está haciendo, cada vez con mayor frecuencia, análisis genéticos. ¿Por qué? Para intentar encontrar pruebas de contacto entre dos personas, entre dos objetos o entre una persona y un objeto. Probar este contacto suele ser muy útil para la investigación. Pero no proporciona necesariamente la prueba de un delito. Es sólo una prueba más entre muchas otras.

Anne Versailles

¿Genético qué?


El ADN está formado por un gran número de genes, y cada uno de éstos está formado a su vez por miles de "perlas". El conjunto de estos genes forma el carné de identidad genético de una persona.

¿Cómo se descubre el carné de identidad genético?

El genetista toma unas pocas células de la base de los cabellos encontrados en la víctima, o de la saliva dejada en una colilla de cigarrillo. Las sumerge en un producto para eliminar de las células todo lo que rodea al ADN. Después hace lo mismo con algunas células de la sangre del sospechoso. Luego, el ADN se prepara especialmente para su análisis. Se introduce en un gel especial y se hace pasar una corriente eléctrica a través de este gel. Al cabo de algunas horas, se obtienen unas barras similares al código de barras (como los que se encuentran en los artículos que compramos), que son visibles bajo una lámpara especial. Se compara entonces, el código de barras del ADN del sospechoso con el de los cabellos encontrados en la víctima.

Estamos formados por billones de células

Todo ser viviente está formado por muchísimas células. Una célula es realmente muy pequeña. Incluso puede decirse que es microscópica porque sólo puede verse con la ayuda de un microscopio que la aumenta múltiples veces. Cada célula tiene una membrana exterior y un núcleo, en el que se encuentra el ADN.



Microscopio en un laboratorio de la policía

b) Descripción del texto

Tipo de texto/Contexto (Tipo de texto según clasificación cruzada)

Exposición / Educativo (No literario)

c) Ítemes

R100Q04

Para explicar la estructura del ADN, el autor habla de un collar de perlas. ¿Cómo varían estos collares de perlas de una persona a otra?

- A Varían en longitud.
- B El orden de las perlas es diferente.
- C El número de collares es diferente.
- D El color de las perlas es diferente.

R100Q05

¿Cuál es el propósito del recuadro titulado “¿Cómo se descubre el carné de identidad genético”?
Explicar...

- A lo que es el ADN.
- B lo que es un código de barras.
- C cómo se analizan las células para encontrar el patrón del ADN.
- D cómo se puede probar que se ha cometido un crimen.

R100Q06

¿Cuál es el objetivo principal del autor?

- A Prevenir
- B Entretener
- C Informar
- D Convencer

R100Q07

Al final de la introducción (en el primer recuadro sombreado) se dice: “Pero ¿cómo probarlo?”
Según el texto, los investigadores intentan encontrar una respuesta a esta pregunta...

- A interrogando a los testigos.
- B realizando análisis genéticos.
- C interrogando concienzudamente al sospechoso.
- D revisando nuevamente todos los resultados de la investigación.

d) Descripción de los ítems

CÓDIGO	FORMATO	DESCRIPCIÓN
R100Q04	SMS	Localiza información en un texto expositivo de mediana extensión.
R100Q05	SMS	Reconoce el propósito de un fragmento destacado de un texto expositivo de mediana extensión.
R100Q06	SMS	Reconoce el propósito de un texto expositivo de mediana extensión.
R100Q07	SMS	Reconoce de qué se trata un texto expositivo de mediana extensión.

B. SIMCE

a) Texto: El ADN ayuda a la policía inglesa

El ADN ayuda a la policía inglesa

Nunca imaginó Alec Jeffrey lo que ocurriría andando el tiempo, mientras tipificaba, en el año 1984 en la Universidad de Leicester, el ADN (ácido desoxirribonucleico). En la actualidad, la policía de Inglaterra posee la mayor colección de ADN humano, perfectamente clasificada en una base de datos que cada día se está incrementando. Y es que durante los últimos cuatro años la policía ha estado ejerciendo su derecho legal de tomar muestras de saliva a cualquier sospechoso que detenga, habiendo ya acumulado más de un millón de perfiles individuales de ADN. El método para tipificar el ADN consiste en identificar los genes que configuran nuestra identidad genética. No cabe duda de que la policía, al disponer de esta enorme información que permite identificar a cualquier individuo con mucha



mayor exactitud que las huellas digitales, tiene en sus manos una poderosa herramienta contra el crimen. Para los detectives basta encontrar en el sitio del suceso un simple pelo, una pizca de sangre, un espermatozoide o una

de personas que le den un uso indebido, ya que puede arrojar datos que en un futuro cercano agudicen la discriminación; por ejemplo, se puede negar el derecho al trabajo de una persona por razones de salud o aun de carácter.

escama de la caspa del cuero cabelludo, para poder tipificar su ADN. Este proceso demora menos de una hora, y comparar la información del sospechoso con la base de datos de la policía no toma más de diez minutos.

Todos estos y otros argumentos se esgrimen actualmente para que al menos la tipificación de inocentes sea destruida. La policía replica que, aún cuando se trate de personas inocentes, el saber que su ADN está tipificado prevendría posibles delitos. Estos son algunos de los problemas que seguramente se discutirán en el Parlamento. De lo que no cabe duda, es de que esta experiencia puede ser muy útil para nuestra propia policía.

Hay que considerar, sin embargo, que esta información puede caer en manos



Fuente Revista Creces, junio de 2000. Adaptación.

b) Descripción del texto

Tipo de texto/Contexto (Tipo de texto según clasificación cruzada)

No literario (Expositivo científico) / Público (Continuo)

c) Ítemes

B1

¿Qué actitud muestra el emisor sobre la tipificación del ADN, en el tercer párrafo?

- A Enojo por el uso indebido que se le da a la tipificación del ADN.
- B Optimismo por la influencia de la tipificación del ADN en el mundo del trabajo.
- C Interés por las posibilidades que tiene la tipificación del ADN en la discriminación de la gente.
- D Preocupación por el uso que se le puede dar a la tipificación del ADN.

B2

En el último párrafo ¿cómo refuerza el emisor la utilidad de ADN para la policía inglesa?

- A Comentando que otras policías pueden obtener beneficios.
- B Explicando los argumentos en contra de la tipificación del ADN de inocentes.
- C Narrando los posibles problemas que se discutirán en el Parlamento.
- D Mostrando la postura de la policía inglesa frente a la tipificación de inocentes.

d) Descripción de los ítemes

CÓDIGO	FORMATO	DESCRIPCIÓN
B1	SMS	Identifica la actitud del emisor en un fragmento de un texto expositivo breve.
B2	SMS	Reconoce mecanismos textuales utilizados por un emisor como apoyo a la información entregada en una sección de un texto expositivo breve.

Metodología 4



- Evaluación de la eficiencia de los sistemas educativos nacionales en su objetivo de proveer calidad y equidad
- Distribución de puntajes SIMCE en PISA y SIMCE 2006, en perspectiva comparada
- Rendimiento en la prueba PISA: ¿es posible entender los alcances y límites de las comparaciones entre países?

Evaluación de la eficiencia de los sistemas educativos nacionales en su objetivo de proveer calidad y equidad

Por
Claudio Thieme
Víctor Giménez
Diego Prior¹

Resumen

El presente estudio evaluó el desempeño de 54 países participantes de PISA 2006. Se utilizó para ello indicadores de eficiencia que relacionan variables de resultados obtenidos con variables de recursos utilizados en la provisión de educación. Como variables de resultados se consideraron de manera conjunta, outputs deseables de logro académico y outputs no deseables de inequidad educativa. Como variables de recursos se utilizaron constructos que dan cuenta de la calidad y cantidad de recursos educativos y humanos puestos a disposición de la educación por cada sistema educativo. De igual forma, se incorporó en el modelo de evaluación de la eficiencia, variables del entorno de cada sistema educativo, que si bien no son recursos controlables por los gestores, afectan los resultados.

Una vez obtenidos los indicadores de desempeño se determinaron las variables estructurales y de proceso que explicaban las diferencias de eficiencia encontradas.

En concreto, se plantearon tres preguntas de investigación: i) ¿Cuál es el grado de eficiencia que exhiben los distintos sistemas educativos?, ii) El no alcanzar el máximo resultado posible de un país en particular ¿se explica por problemas de gestión?, ¿de dotación de recursos?, ¿de entorno?, ¿o por una combinación de estos factores?, y iii) ¿Qué características de proceso explican las diferencias de desempeño de los distintos sistemas educativos?

Los resultados mostraron que el conjunto de países evaluados se clasifica en 7 grupos, de acuerdo con el tipo de problemática que exhibe. Chile se ubica en el grupo de países que muestra ineficiencias, producto de una problemática mixta de gestión y reducida dotación de recursos educativos y humanos puestos a disposición de la educación.

Por otra parte, dos factores de proceso de los sistemas educativos resultan estadísticamente significativos para explicar estas diferencias de desempeño: la capacidad de los establecimientos para la toma de decisiones autónomas o de manera conjunta con la autoridad en materia académica y la capacidad de tomar de decisiones autónomas en la formulación de presupuestos.

¹ Claudio Thieme es Ingeniero Civil Industrial, Doctor en Ciencias Económicas y Empresariales (Universidad Autónoma de Barcelona). Académico de la Facultad de Economía y Empresa de la Universidad Diego Portales. Víctor Giménez y Diego Prior son Licenciados y Doctores en Ciencias Económicas y Empresariales (Universidad Autónoma de Barcelona). Académicos del Departamento de Economía de la Empresa de la Universidad Autónoma de Barcelona.

Los autores agradecen al Ministerio de Educación y Ciencia de España. Programa Nacional del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2004-2007, Programa de Ciencias Sociales, Económicas y Jurídicas (Economía) (código: SEJ2007-60995/ECON), y proyecto Fondecyt N° 11085061.

Introducción

Los resultados educativos obtenidos por un país en una prueba estandarizada internacional son el primer referente del nivel de educación que reciben sus estudiantes. Sin embargo, no son estos resultados, por sí solos, indicadores de la calidad del desempeño de su sistema educativo o de la labor de sus gestores. Para evaluar el desempeño de una organización, cualquiera ella sea, se suelen utilizar indicadores de economía, productividad, eficiencia o eficacia. Cada uno de ellos mide y relaciona distintos aspectos de la función de producción que es necesario abordar al iniciar el reporte de este estudio.

Sobre la función de producción en educación, el conocimiento acumulado proveniente de la investigación educativa (principalmente del Programa de Investigación sobre Eficacia y Mejora Escolar), indica que esta función es un modelo complejo en el que se interrelacionan distintas variables y actores (Teddlie y Reynolds, 2000).

Como se aprecia en la Figura 1, los resultados educativos (*outputs*), son consecuencia tanto de los recursos provistos (*inputs*), como del proceso, las variables del entorno y las múltiples relaciones entre ellas. Por tanto, resultados de pruebas estandarizadas —como PISA— que incorporan los efectos de todos estos componentes, son un buen indicador del logro académico, pero no dan cuenta de aspectos específicos, como la calidad del desempeño que tienen los gestores del sistema educativo.

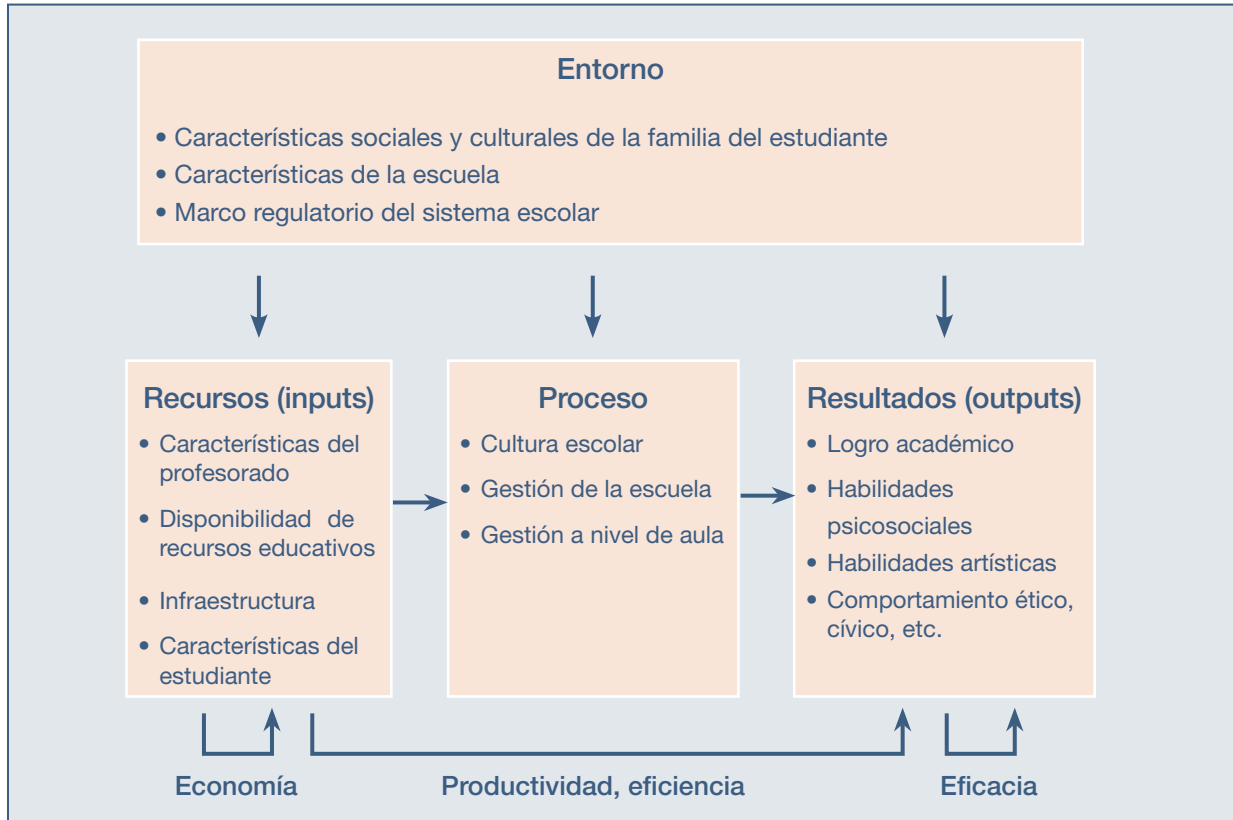
La definición de estos indicadores no resulta fácil, principalmente por el gran número de acepciones que estos conceptos involucran. Sin embargo, es generalmente aceptado que la noción de *eficacia* se relaciona sólo con los resultados del proceso evaluado. Suele expresarse como la *ratio* entre los resultados esperados u óptimos y los resultados efectivamente

conseguidos por la organización o institución. De manera análoga, el concepto de *economía* se relaciona sólo con los recursos, y compara los presupuestados u óptimos, con los recursos utilizados en dicha provisión. Por su parte, en un sentido clásico, a diferencia de los conceptos anteriores, la *productividad* es definida como una *ratio* bajo la cual el *output* observado es dividido por los *inputs* (trabajo, energía, etc.) requeridos para producirlo. Desde la perspectiva microeconómica, la *eficiencia* analiza la utilización de recursos en la producción de bienes o servicios y compara esta relación con otras organizaciones similares. En este escenario, surge el concepto de *eficiencia técnica* que se refiere a la capacidad de una unidad en producir el máximo *output* posible con un nivel dado de *inputs*, o de utilizar el menor *input* posible para obtener un determinado *output*.

Dado lo anterior, la relación entre productividad y eficiencia es muy estrecha. Ambos conceptos buscan expresar la capacidad de una organización en transformar *inputs* en *outputs*. Sin embargo, el elemento que marca la diferencia entre ambos, corresponde a que sólo la eficiencia incorpora la noción de comparación con un óptimo, ya sea teórico o empírico.



Figura 1.
Función de producción en educación e indicadores de desempeño



Como señalan Prior, Verges y Vilardell (1993), los modelos de evaluación de la eficiencia pueden clasificarse en dos grandes grupos: los que no utilizan una función de producción frontera y los que sí la utilizan. Los modelos frontera intentan estimar una función de producción extrema, a la luz de los *inputs* y *outputs* de la evaluación. Por su parte, los modelos no frontera, no requieren de la estimación de una frontera de producción, ya sea porque intentan estimar una función de producción de tipo medio o porque evalúan la eficiencia de manera absoluta, sin necesidad de compararse con otras unidades.

Sin lugar a dudas, los modelos frontera son los que han concentrado la mayor atención de la comunidad investigadora sobre el tema de la eficiencia. Ello, por cuanto la concepción frontera da cuenta fiel de la

característica esencial de la medición de la eficiencia que trata de “evaluar hasta qué punto una determinada organización está obteniendo la máxima producción con el menor consumo posible de factores” (Prior, 1992, p.113).

En particular, en este tipo de estudios es tradicional el uso de métodos de estimación no paramétricos, principalmente el denominado DEA (*Data Envelopment Analysis*). Existen innumerables ejemplos de aplicación del DEA al sector de la educación, aunque, en su mayoría, se limitan al estudio de las diferencias encontradas entre centros educativos, de una misma región o país².

² Como ejemplos de este tipo de trabajos podemos citar a Bessent y Bessent (1980), Färe, Grosskopf y Weber (1989), Smith y Mayston (1987), Bessent y otros (1982), Thanassoulis y Dunstan (1994), Pedraja y Salinas (1996), Mancebón (2003) y Thieme (2005).

Los estudios sobre eficiencia, y en especial los que utilizan modelos frontera, coinciden en la importancia clave de una adecuada selección de *inputs* y *outputs* para el correcto desempeño del modelo, y en la conveniencia de incorporar *outputs* ajustados en calidad o adicionar una medida de este indicador. Las variables socioeconómicas, culturales y familiares del entorno del estudiante, que no están bajo el control del gestor, tienen un considerable impacto sobre los resultados del proceso educativo. Si se considera que el objetivo de la medición de la eficiencia es evaluar la gestión de los recursos puestos a disposición de una entidad, sería un error considerar que dicha asignación de recursos es obra exclusiva de los gestores. Sin embargo, peor aún sería ignorar que dichas variables existen, ya que, de estar positivamente relacionadas con los resultados, se sobrevaloraría la eficiencia de aquellas unidades expuestas a condiciones favorables en detrimento de aquellas menos favorecidas.

Una dificultad adicional es que la medida del resultado académico obtenido por un país no se debería circunscribir únicamente a los conocimientos adquiridos durante la etapa escolar, sino que habría de incluir otros resultados asociados con la equidad de dicha provisión. Efectivamente, incorporar variables de equidad educativa en el modelo de evaluación del desempeño, es altamente relevante al evaluar el éxito relativo de un sistema educativo. Asegurar una educación justa y equitativa tiene implicancias, no sólo en el ámbito social, sino también en el político y en el económico. Desde el punto de vista social, posibilita avanzar hacia una sociedad que garantiza la igualdad de oportunidades entre sus ciudadanos. Desde un punto de vista político, disminuye las polarizaciones de los grupos antagónicos, posibilitando la convergencia en materia de políticas públicas. Finalmente, desde el punto de vista económico, favorece el desarrollo, al generar una mejor y más amplia base desde donde seleccionar a los mejores para cada tarea específica (Puryear, 2003; Gross, 2003 y Peña, 2004).

Así, cuando se evalúa el desempeño de los sistemas educativos nacionales, no basta con considerar como variables de resultados el logro académico obtenido por sus estudiantes, porque el éxito relativo se asocia a la provisión de un equitativo sistema que brinde similares oportunidades para todos los estudiantes. Se deben considerar como resultados del proceso educativo tanto los resultados de logro académico como los índices de equidad educativa, obtenidos por cada uno de los sistemas educativos. El informe PISA 2006 brinda esta oportunidad, al incorporar ambos indicadores en cada una de sus pruebas.

En este contexto, el presente trabajo tiene como principales objetivos: evaluar la eficiencia técnica de gestión y el máximo mejoramiento potencial alcanzable por los sistemas educativos de 54 países, y medir el efecto negativo que tienen los factores socioeconómicos y la provisión de recursos de cada país, en su objetivo de alcanzar el máximo resultado posible en la obtención de logro académico y equidad educativa.

La “eficiencia técnica de gestión” se refiere al grado en que un sistema educativo obtiene el mayor rendimiento académico con la menor inequidad, comparado con otros sistemas educativos de similares recursos disponibles y condiciones socioeconómicas y culturales. Sin embargo, en general, a un país no le basta ser eficiente en la gestión de sus recursos; le interesa también, obtener los máximos resultados posibles. Sin embargo, este rendimiento potencial depende de las condiciones de entorno de cada país, y exige destinar al sistema educativo los recursos económicos necesarios. El indicador de “máximo mejoramiento potencial”, por su parte, se refiere al grado de eficiencia que se obtendría si: a) se contase con la dotación de recursos óptimos (a cuyo indicador se denomina “máximo mejoramiento potencial de mediano y largo plazo”, ya que los cambios en la dotación de recursos del sistema se producen en este horizonte de tiempo), y b) se contase con la dotación óptima de recursos y condiciones del entorno (a cuyo indicador se denomina “máximo mejoramiento potencial de

muy largo plazo”, ya que cambios estructurales del entorno, principalmente el nivel socioeconómico de la población, son sólo posibles de obtener en ese horizonte de tiempo).

A partir de este estudio, se pudo establecer cuánto debería modificarse la dotación de recursos y de condiciones del entorno de cada país, para alcanzar estas dotaciones óptimas, y cuál es el impacto negativo total sobre los resultados, de las condiciones socioeconómicas y de las dotaciones de recursos, cuando dichas condiciones no son las óptimas.

Para ello, se utilizaron las denominadas funciones distancia direccionales (*Directional Distance Functions, DDF*). Este método de estimación frontera se diferencia del tradicional DEA debido a que, simultáneamente, expande los efectos deseados (*outputs*) a la vez que contrae los no deseados, mediante el consumo de *inputs* o de la producción de efectos indeseados —*bad outputs* (Briec, 2000 y Färe y Grosskopf, 2000). Ello resulta fundamental para el tratamiento directo (es decir, aquel que no requiere transformaciones en los datos), de los *outputs* deseados (el logro académico) y de los *bad outputs* (la inequidad).

A continuación, se describe la metodología utilizada, luego, se analizan los resultados, y finalmente, se señalan las principales conclusiones.

1. Metodología

1.1 Descripción de la metodología propuesta

Acorde al concepto de eficiencia, el cual incorpora la noción de comparación con un óptimo, en este trabajo se utilizaron modelos de frontera no paramétrica³, que

3 Los modelos de frontera no paramétrica se caracterizan por no utilizar una forma funcional determinada para construir la función de producción, sino construir una función de producción empírica, a partir de los datos de los *inputs* empleados y de los *outputs* conseguidos.

resultan especialmente idóneos para evaluar sectores donde existe dificultad para encontrar una definición aceptable de precios correspondientes a los *outputs* producidos y a los *inputs* consumidos; y cuando se busca evaluar funciones de producción que conllevan múltiples *outputs* y/o múltiples *inputs*, como es el caso de la educación.

Básicamente, esta metodología se compone de dos procesos. El primero, corresponde a la construcción de una frontera de producción empírica, configurada a través de combinaciones lineales convexas de las mejores prácticas observadas. El segundo, corresponde a la cuantificación del grado de eficiencia de cada una de las observaciones muestrales que no forman parte de la frontera eficiente, la cual se realiza a través del cálculo de la distancia radial⁴ relativa de la DMU, analizada con respecto a esta frontera de las mejores prácticas (Álvarez, ed., 2001).

Estos modelos de frontera no paramétrica, bien sean DEA, DDF u otros, como por ejemplo, los modelos aditivos, evalúan la eficiencia relativa de un conjunto de unidades de decisión homogéneas (en nuestro caso, los sistemas nacionales de educación), conocidas como DMU's⁵. En particular, para evaluar los sistemas educativos, se requiere considerar los factores sociales y económicos específicos de cada país y también los factores operativos que pueden ser controlables por los gestores del sistema educativo. La importancia de las variables socioeconómicas de la familia en los resultados académicos, se encuentra ampliamente documentada (White, 1982; Berger y Toma, 1994, y Hanushek y Taylor, 1990). De igual manera, la literatura existente ofrece diferentes alternativas para el tratamiento de los factores ambientales en la evaluación de la eficiencia, no existiendo todavía

4 Existen diversas formas de medir la distancia a la frontera. En este caso, como en la mayoría de trabajos con fronteras no paramétricas, se utilizará la distancia radial. Esto se traduce en expandir los *outputs* en la misma proporción, hasta alcanzar la frontera de las mejores prácticas observadas.

5 Del inglés *Decision Making Units*.

consenso acerca de cuál es la más adecuada (Muñiz, 2002). En este estudio, se optó por ampliar y ajustar el método utilizado previamente en Giménez, Prior y Thieme (2007) siguiendo la línea de Lozano-Vivas, Pastor y Pastor (2002).

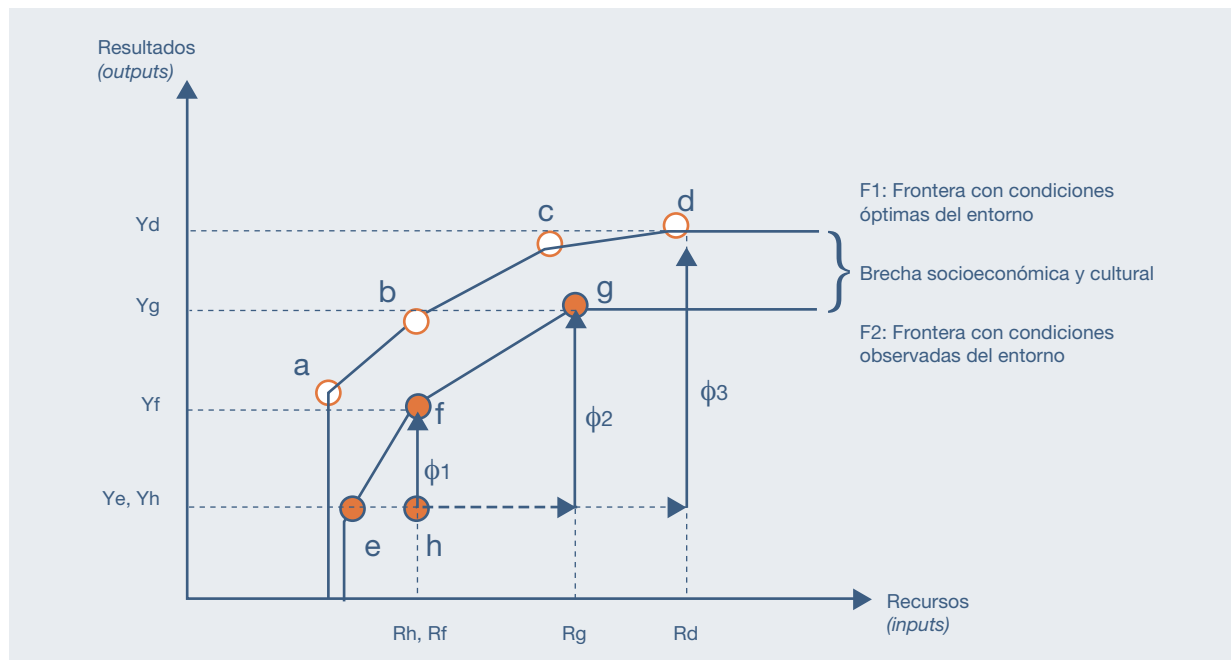
El ajuste consistió en utilizar estimaciones de DDF en vez de DEA, dada la necesidad de evaluar de manera conjunta los *outputs* deseables junto a los *outputs* no deseables; situación que relaciona a la perfección la aproximación metodológica de las DDF, con los requisitos de nuestra aplicación empírica. Por otra parte, también se determinó el máximo *output* potencial, es decir, aquel que se podría obtener si se diesen a la vez tres condiciones: 1) ser totalmente eficiente, 2) gozar de una disposición óptima de recursos, y 3) disponer de las variables del entorno más favorables. De esta forma, se pudo determinar el efecto limitativo que

tienen los recursos y el entorno, sobre los resultados, cuando su disposición no es óptima.

La medida de la eficiencia técnica de gestión, la determinación de la importancia de los factores contextuales y la determinación del máximo *output* potencial, se realizó a partir de la resolución de diversos programas lineales, construidos sobre un conjunto de ecuaciones.

En la Figura 2 se observa de manera simplificada⁶ el método propuesto para la evaluación de la eficiencia de gestión, del máximo *output* potencial de mediano y largo plazo y del máximo *output* potencial de muy largo plazo. En ella se representan los resultados obtenidos por distintos sistemas educativos, acorde a la dotación de recursos que destinan a educación.

Figura 2.
Recursos, resultados y factores del entorno



Fuente: Elaboración propia, a partir de Giménez, Prior y Thieme (2007)

6 Es simplificado, ya que la figura sólo considera la existencia de un *output*, un *input* y dos niveles posibles de dotación de variables del entorno, y contempla un modelo tradicional, donde se busca expandir los resultados deseados.

En la evaluación de la eficiencia técnica de gestión, el primer paso es determinar la frontera de las mejores prácticas observadas, considerando el nivel observado de las variables del entorno. En la Figura 2, por ejemplo, los sistemas educativos que presentan una situación de eficiencia técnica de gestión (ϕ_1) son a, b, c, d, e, f y g, por cuanto, en ese conjunto, no se observan otros sistemas educativos que, con la misma cantidad de recursos y afectados por similares condiciones del entorno, obtengan mejores resultados. De esos siete sistemas, los cuatro primeros (a, b, c, d) conforman la frontera asociada a un nivel óptimo de variables del entorno, mientras que los últimos tres (e, f, g), conforman la frontera con niveles que, aunque observados, no son óptimos.

El coeficiente de eficiencia técnica de gestión (cuyo índice se denominará ϕ_1), indica el máximo nivel de mejora posible del rendimiento académico, acompañado de la máxima reducción factible en el nivel de inequidad, dado los recursos disponibles y las condiciones socioeconómicas y culturales de cada país. Este coeficiente toma un valor igual a cero en caso de eficiencia, y superior a cero, en caso de ineficiencia.

Así por ejemplo, si el coeficiente de eficiencia técnica de gestión fuera 0,12 (lo que podría corresponder a un sistema educativo como “h”), ese sistema educativo debería ser capaz de aumentar en un 12% el puntaje de logro académico de sus estudiantes y de disminuir en el mismo porcentaje su nivel de inequidad, sin modificar su actual dotación de recursos ni las características socioeconómicas de su población; ello, por cuanto existen en ese conjunto, países similares que son capaces de lograrlo (en particular, el sistema educativo “f”).

Como el coeficiente de eficiencia técnica de gestión indica el mejoramiento que podría experimentar un país, mediante la mejora en la gestión de sus procesos dentro del sistema educativo, entonces se puede afirmar que

el grado de ineficiencia del sistema educativo de un país, es de absoluta responsabilidad de los gestores de ese sistema en su conjunto.

Ahora bien, aumentar el grado de eficiencia técnica de gestión, debería ser el primer gran objetivo de corto y mediano plazo; sin embargo, no puede ser el único objetivo: junto con ser eficientes, los gestores de sistemas educativos deben aspirar a obtener resultados (logro académico y equidad) similares a los de los países con mejores prácticas observadas (como los que obtiene en el ejemplo, el sistema educativo “g”, con similares condiciones del entorno y dotación óptima de recursos, o el sistema educativo “d” cuando se trata de condiciones óptimas del entorno y de dotación de recursos).

La no obtención de resultados similares a los de los países con máximos resultados, usualmente responde a dos problemas de distinta naturaleza y cuya solución no se ajusta a los mismos plazos. Estos problemas son: a) dotación insuficiente de recursos (cuya solución es de mediano y largo plazo), y b) características sociales, económicas y culturales de la población (cuya solución, es de muy largo plazo).

Para fijar esta meta aspiracional, en primer lugar, se debe determinar el máximo mejoramiento al que un sistema educativo podría llegar, en el mediano y largo plazo, si dispusiera de los recursos necesarios para ello. En otras palabras, obtener resultados similares a “g”. A este indicador se le denominará máximo *output* potencial de mediano y largo plazo, y su índice se representará con ϕ_2 . Para determinar este índice, primero se debe calcular el *output* máximo (Y_g), esto es, el máximo incremento porcentual de todos los *outputs* deseados (y el máximo decremento de los no deseados), considerando el nivel observado de las variables del entorno y dejando libre la dotación de recursos. En una segunda etapa, una vez que el *output* máximo (Y_g) ha sido calculado, el programa lineal determina el nivel más bajo de *inputs* que están asociados con el *output* máximo (R_g). De esta forma se

puede determinar la dotación óptima de recursos. En síntesis, el índice ϕ_1 recoge exclusivamente la problemática de gestión del sistema educativo, mientras que ϕ_2 añade a lo anterior, el efecto de la insuficiente disposición de recursos del sistema educativo. De allí que $\lambda_1 = (1+\phi_2)/(1+\phi_1)$ representa el eventual impacto negativo de la insuficiente dotación de recursos, en los logros del sistema educativo (a mayor valor de λ_1 , más fuerte es el impacto negativo derivado de la carencia de recursos).

Ahora bien, junto con determinar la eficiencia técnica de gestión (ϕ_1) y el máximo mejoramiento potencial, dadas condiciones óptimas de recursos (ϕ_2), también resulta de interés calcular el máximo mejoramiento posible, dadas óptimas condiciones del entorno (ϕ_3).

En el cálculo de este coeficiente, todos los sistemas educativos son comparados considerando sólo sus resultados de logro académico e inequidad. Esto es, se calcula la distancia que separa los resultados de un país con la frontera de la mejor práctica, sin considerar que puedan estar operando bajo condiciones negativas del entorno o con una provisión de recursos menor a la óptima. Por tanto, el índice ϕ_3 representa el máximo crecimiento potencial que un país debiera experimentar para alcanzar a aquellos que obtienen los mejores resultados (en la Figura 2, esta situación la representa el sistema educativo “d”).

De esta forma, el máximo mejoramiento potencial de muy largo plazo (ϕ_3) incorpora los efectos conjuntos de: gestión, provisión de recursos y entorno. Por tanto, representa un objetivo de muy largo plazo, al requerir, al menos, de un recambio generacional.

Considerando lo expuesto, $\lambda_2 = (1+\phi_3)/(1+\phi_2)$ representa el efecto negativo que tiene el entorno sobre los resultados del sistema educativo; pudiendo descomponerse el mejoramiento máximo de resultados, como $(1+\phi_3) = \lambda_2 * \lambda_1 * (1+\phi_1)$. Lo anterior equivale a decir que el máximo mejoramiento de los resultados se

explica por el impacto negativo de las características sociales, económicas y culturales de la población (λ_2), la dotación no óptima de recursos (λ_1), y las ineficiencias de gestión del sistema educativo ($1+\phi_1$).

Una vez evaluados los sistemas educativos, resulta interesante determinar las causas a nivel de proceso que originan estas diferencias de desempeño. Para ello se procedió a realizar un análisis de regresión⁷, considerando como variable dependiente el indicador de máximo mejoramiento potencial de mediano y largo plazo ($1+\phi_2$), y como variables independientes o explicativas, indicadores de proceso de cada sistema educativo.

1.2 Descripción de la muestra y selección de las variables

La metodología descrita anteriormente se utilizó para la evaluación del desempeño de 54 países participantes en la prueba PISA 2006, para los cuales se contó con información completa. Tal como se señaló previamente, un buen sistema educativo es aquel que obtiene altos logros académicos de sus estudiantes y, de manera conjunta, bajos niveles de inequidad educativa.

Para el caso de logro académico, se consideró de manera separada, el logro académico de Lectura y el logro promedio de Matemática y Ciencias, para recoger la eventualidad de que algunos países pudiesen hacer énfasis en uno u otro sentido. Para resguardar la validez de los datos, al promediar los puntajes de Matemática y Ciencias que se encuentran en escalas distintas, se optó por realizar los análisis con escalas estandarizadas y con los resultados originales publicados. Lo anterior permitió también determinar la robustez de los resultados.

⁷ En particular, se utilizó una regresión con variables truncadas, dadas las características del indicador de máximo output potencial, que toma un valor mínimo de 1.

La variable “inequidad educativa”, se definió en este trabajo como el grado en que los logros académicos obtenidos por los estudiantes, dependen de su nivel socioeconómico.

El informe PISA entrega una muy buena aproximación a este indicador, al determinar, para cada prueba y país, una recta de regresión de los puntajes obtenidos por los estudiantes acorde con su nivel socioeconómico⁸. A partir de estas rectas de regresión fue posible determinar la pendiente y proporcionar la diferencia de puntaje que obtienen los alumnos al variar en una desviación estándar el índice PISA de status económico, social y cultural (ESCS). En este contexto, una pendiente más pronunciada implica una mayor inequidad y por ende un mayor impacto sobre el puntaje que obtienen los alumnos, asociado a su nivel socioeconómico. Un sistema con máxima equidad será aquel cuyos resultados no estén influidos por el nivel socioeconómico de su población y, por ende, la pendiente de la recta que lo represente será nula.

Ahora bien, una variación de 30 puntos por unidad de desviación estándar, no es proporcionalmente lo mismo en un país que obtiene en promedio 400 puntos, que en otro que obtiene 500 puntos. Por tal motivo, en este estudio se ajustó este indicador, calculando una estimación del punto medio de la recta de regresión. Asimismo, al no tener certeza de la distribución normal de los resultados, se optó por calcular el promedio de puntajes obtenidos en cada prueba y país, con los percentiles 5 y 95. De esta forma, el *output* no deseado de inequidad del país “i” corresponde a un *proxy* de la “elasticidad media de logro académico y nivel socioeconómico de cada país”.

Este *proxy* de la elasticidad media se calcula como el promedio (entre las tres pruebas evaluadas en PISA) del cociente de la diferencia de puntaje obtenido por un país, al variar en media desviación estándar el

8 Cabe señalar que algunas rectas de regresión no son estadísticamente significativas.

indicador de nivel socioeconómico, sobre el promedio de puntaje obtenido en los percentiles 5 y 95.

$$\text{Inequidad Educativa}_i = \sum_{j=1}^3 ((\Delta P_{ij}) / ((\overline{P95}_{ij} + \overline{P5}_{ij}) / 2)) / 3$$

Siendo “j” cada prueba evaluada en PISA, e “i” cada uno de los 54 países considerados en la evaluación.

Para la determinación de los inputs controlables se consideraron en principio 5 variables:

- número de profesores por cada 100 alumnos
- índice de percepción de la calidad de los recursos educativos
- porcentaje promedio de estudiantes con dos o más horas de enseñanza en las tres materias
- número de computadores por cada 100 alumnos
- índice de disponibilidad de recursos humanos competentes⁹

Los modelos de frontera empírica pierden capacidad de discriminación cuanto mayor es el número de *inputs* y de *outputs* analizados. Por lo tanto, y para mitigar en lo posible los efectos de este problema, se realizó un análisis factorial con el objetivo de reducir su número. Se aplicó el método de rotación Quartimax, ya que posibilita reducir el número de variables con pesos altos en un factor, lo que facilita la interpretación de los factores (García, Gil y Rodríguez, 2000). Al examinar la conformación de los dos factores¹⁰ se ve que el primero está compuesto por “número de computadores por 100 alumnos”, “porcentaje promedio de estudiantes con 2 o más horas de enseñanza en las tres materias” e “índice de calidad de los recursos educativos”, por lo que se lo denominó “Disponibilidad y calidad de los recursos educativos”. El segundo factor concentra las variables de “número de profesores por cada 100 alumnos” e “índice de disponibilidad de recursos

9 Corresponde a 100% de carencia de recursos humanos competentes.

10 La medida de adecuación muestral de KMO arrojó un valor de 0,623, lo cual indica un buen ajuste y se hace aconsejable la aplicación del análisis factorial (García, Gil y Rodríguez, 2000). El primer factor explica un 49,2% de la varianza y el segundo explica un 22,2%. En total, con estos dos factores, se explica un 71,46% de la variabilidad total de los *inputs* controlables originales.

humanos competentes”; por lo que se lo denominó “Disponibilidad y calidad de los RRHH”.

Por tanto, cuando se describen las variables controlables por el gestor de un sistema educativo, el estudio considera los siguientes factores y variables que los conforman¹¹:

- Disponibilidad y calidad de los recursos educativos:
 - Número de computadores por cada 100 alumnos.
 - Porcentaje promedio de estudiantes con 2 o más horas de enseñanza en las tres materias.
 - Índice de calidad de los recursos educativos.
- Disponibilidad y calidad de los recursos humanos para la enseñanza:
 - Número de profesores por cada 100 alumnos.
 - Índice de disponibilidad de los recursos humanos para la enseñanza.

Como variable no controlable por el gestor (variables contextuales), se consideró el índice ESCS de PISA 2006, que incorpora antecedentes económicos, sociales y culturales de los estudiantes, sus familias y el entorno.

Como se señaló al comienzo, el presente trabajo tiene por objeto: evaluar la eficiencia técnica de gestión, calcular el máximo mejoramiento potencial alcanzable por los sistemas educativos de los 54 países estudiados, y medir el efecto negativo que tienen

sobre los resultados, los factores socioeconómicos y la dotación de recursos de cada país.

La metodología de evaluación de desempeño involucra los resultados del proceso educativo (*outputs*), los recursos controlables por el gestor (*inputs* controlables) y las variables del entorno (*inputs* no controlables o del entorno). Las variables que se consideraron en cada una de estas categorías fueron las siguientes:

- Resultados del proceso educativo (*outputs*)¹²:
 - y_1 : logro académico en Lectura.
 - y_2 : promedio de logro académico en Ciencias y Matemática.
 - y_3 : promedio de inequidad en la consecución del logro en Ciencias, Lectura y Matemática.
- Recursos controlables por el gestor (*inputs* controlables):
 - x_1 : índice de disponibilidad y calidad de los recursos educativos.
 - x_2 : índice de disponibilidad y calidad de los recursos humanos.
- Variables del entorno (*inputs* no controlables o del entorno):
 - e_1 : índice socioeconómico y cultural de PISA 2006.

Los valores descriptivos de las variables y los correspondientes a cada país considerado en la evaluación, se describen en la Tabla 1.

11 Los valores para cada país de estos dos factores, corresponden a las estimaciones de puntuaciones factoriales para cada DMU. Las puntuaciones factoriales incluyen valores positivos y negativos. Los valores positivos indican que la unidad se encuentra sobre la media en el factor evaluado. Por el contrario, valores negativos indican que dicho país presenta niveles inferiores al promedio de países considerados en el estudio. La existencia de valores negativos en los factores, nos obliga a trasladar el eje original de los *inputs* para cumplir con la condición de no negatividad de las variables (al tratarse de variables relativas a los *inputs*, esta traslación no tendrá ningún efecto en los valores estimados). Para un mayor detalle sobre los problemas de *translation invariance*, véase Cooper, Seiford y Tone (2006).

12 Como se explicó anteriormente, se realizan dos mediciones, utilizando distintos valores para y_1 e y_2 . La primera, utilizando los puntajes obtenidos e informados; y la segunda, con los puntajes estandarizados entre los 54 países considerados en la muestra.

Tabla 1.
Estadísticas descriptivas de las variables de recursos, resultados y entorno del modelo

PAÍS	Outputs					Inputs		Entorno
	Outputs deseados				Outputs no deseados			
	Evaluación 1		Evaluación 2					
	Logro Lectura	Logro Promedio Matemática y Ciencias	Logro Lectura normalizado	Logro Promedio normalizado Matemática y Ciencias	Inequidad Promedio	Factor Recursos Educativos	Factor Recursos Humanos	Indice NSEC
ARGENTINA	373,7	386,2	-1,575	-1,569	0,103	-1,526	1,269	-0,636
AUSTRALIA	512,9	523,4	0,941	0,935	0,079	1,789	-0,488	0,207
AUSTRIA	490,2	508,2	0,530	0,656	0,092	0,828	1,015	0,197
AZERBAIYÁN	352,9	429,2	-1,951	-0,818	0,028	-1,542	0,252	-0,446
BÉLGICA	500,9	515,4	0,724	0,782	0,097	0,022	0,799	0,173
BRASIL	392,9	379,9	-1,229	-1,681	0,078	-1,245	-1,129	-1,119
BULGARIA	401,9	423,8	-1,065	-0,881	0,123	-1,217	1,187	-0,210
CANADÁ	527,0	530,7	1,196	1,069	0,063	1,067	-0,664	0,368
CHILE	442,1	424,8	-0,339	-0,861	0,087	-0,980	-1,166	-0,700
CHINA -TAIPEI	496,2	540,9	0,640	1,247	0,081	0,831	0,018	-0,313
COLOMBIA	385,3	379,0	-1,365	-1,699	0,069	-0,055	-1,496	-1,004
CROACIA	477,4	480,2	0,299	0,152	0,069	-0,686	0,426	-0,109
REPÚBLICA CHECA	482,7	511,4	0,395	0,714	0,105	0,368	-0,123	0,029
DINAMARCA	494,5	504,5	0,608	0,581	0,071	1,215	-0,040	0,309
ESTONIA	500,7	523,0	0,721	0,930	0,062	0,095	-0,622	0,141
FINLANDIA	546,9	555,8	1,555	1,530	0,056	0,513	0,641	0,256
ALEMANIA	494,9	509,7	0,616	0,686	0,094	0,448	-0,762	0,293
GRECIA	459,7	466,3	-0,021	-0,106	0,079	-0,296	1,769	-0,153
HONG KONG -CHINA	536,1	544,8	1,360	1,322	0,046	1,196	-0,321	-0,674
HUNGRÍA	482,4	497,4	0,389	0,462	0,093	0,393	1,098	-0,085
ISLANDIA	484,4	498,2	0,427	0,467	0,056	0,963	0,417	0,767
INDONESIA	392,9	392,2	-1,228	-1,462	0,056	-0,811	-1,603	-1,521
IRLANDA	517,3	504,9	1,021	0,597	0,074	0,035	0,299	-0,015
ISRAEL	438,7	447,9	-0,401	-0,443	0,095	-0,332	0,828	0,216
ITALIA	468,5	468,5	0,139	-0,065	0,065	0,275	1,170	-0,070
JAPÓN	498,0	527,2	0,671	1,005	0,077	1,286	0,629	-0,011
JORDANIA	400,6	403,0	-1,089	-1,254	0,069	-0,737	-1,931	-0,571
KIRGUISTÁN	284,7	316,3	-3,184	-2,846	0,103	-2,556	-1,737	-0,659
COREA	556,0	534,8	1,720	1,133	0,061	1,295	-0,211	-0,007

(Continuación en página siguiente)

(Continuación Tabla 1)

PAÍS	Outputs					Inputs		Entorno
	Outputs deseados				Outputs no deseados			
	Evaluación 1		Evaluación 2					
	Logro Lectura	Logro Promedio Matemática y Ciencias	Logro Lectura normalizado	Logro Promedio normalizado Matemática y Ciencias	Inequidad Promedio	Factor Recursos Educativos	Factor Recursos Humanos	Indice NSEC
MACAO-CHINA	492,3	517,9	0,568	0,828	0,026	0,974	-0,998	-0,909
MÉXICO	410,5	407,7	-0,910	-1,180	0,065	-0,554	-1,699	-0,990
MONTENEGRO	392,0	405,5	-1,245	-1,216	0,063	-1,672	0,417	-0,017
PAÍSES BAJOS	506,7	527,8	0,830	1,010	0,080	0,310	-0,149	0,252
NUEVA ZELANDA	521,0	526,2	1,088	0,986	0,092	1,504	-0,874	0,104
NORUEGA	484,3	488,2	0,424	0,288	0,075	0,864	-0,020	0,421
POLONIA	507,6	496,6	0,846	0,444	0,080	-0,149	1,474	-0,301
PORTUGAL	472,3	470,2	0,207	-0,036	0,064	-0,662	2,256	-0,617
RUMANIA	395,9	416,6	-1,173	-1,017	0,084	-1,093	0,557	-0,372
RUSIA	439,9	477,6	-0,379	0,097	0,071	-1,379	-0,240	-0,101
SERBIA	401,0	435,5	-1,081	-0,673	0,082	-0,759	0,824	-0,137
ESLOVAQUIA	466,3	490,3	0,099	0,326	0,095	-0,690	0,094	-0,149
ESLOVENIA	494,4	511,6	0,607	0,722	0,084	0,128	0,907	0,129
ESPAÑA	460,8	484,2	-0,000	0,219	0,062	0,234	0,954	-0,311
SUECIA	507,3	502,8	0,840	0,557	0,075	0,593	0,574	0,237
SUIZA	499,3	520,6	0,695	0,876	0,082	0,757	0,676	0,087
TAILANDIA	416,8	419,0	-0,797	-0,972	0,067	0,551	-1,942	-1,434
TÚNEZ	380,3	375,5	-1,455	-1,762	0,059	-1,240	-0,071	-1,204
TURQUÍA	447,1	423,9	-0,248	-0,885	0,073	-0,404	-2,074	-1,277
REINO UNIDO	495,1	505,1	0,619	0,604	0,088	2,089	-0,677	0,191
ESTADOS UNIDOS	475,83*	481,6	0,27*	0,174	0,094	0,819	-0,035	0,135
URUGUAY	412,5	427,5	-0,874	-0,819	0,086	-0,971	-0,237	-0,510
Media	460,8	472,3	0,0	0,0	0,08	0,00	0,00	-0,22
Mínimo	284,7	316,3	-3,2	-2,8	0,03	-2,56	-2,07	-1,52
Máximo	556,0	555,8	1,7	1,5	0,12	2,09	2,26	0,77
Desviación estándar	55,3	54,1	1,0	1,0	0,02	1,00	1,00	0,51

* valor proyectado

En una segunda etapa del estudio se buscó determinar los factores de proceso que explican las diferencias de desempeño encontradas. Para ello, se realizó el análisis de regresión (explicitado anteriormente en el apartado metodológico) y se consideraron 17 variables de proceso provistas en el informe PISA. Una descripción de cada una de ellas, y los estadísticos descriptivos, se muestra en la Tabla 2.

Con la finalidad de clarificar el alcance de estas variables, se procedió a realizar un análisis factorial. La conformación de los factores y las variables que los componen se presentan a continuación:

Factor 1: “Capacidad de los establecimientos para la toma de decisiones autónoma o de manera conjunta con la autoridad, en materia académica”:

- Contenidos de los cursos.
- Cursos ofrecidos.
- Políticas de evaluación.
- Textos escolares.
- Políticas disciplinarias.
- Admisión de estudiantes.

Factor 2: “Uso de los mecanismos de rendición de cuentas”:

- Desempeño del director.
- Desempeño de los profesores.
- Seguimiento de la autoridad administrativa.
- Publicidad.
- Asignación de recursos.

Factor 3: “Capacidad de los establecimientos para la toma de decisiones autónoma o de manera conjunta con la autoridad, en la contratación o despido de profesores”:

- Despido de profesores.
- Contratación de profesores.



Factor 4: “Capacidad de los establecimiento para la toma de decisiones autónomas o de manera conjunta con la autoridad, en la formulación y asignación de presupuestos”:

- Formulación de presupuesto.
- Asignación de presupuesto.

Factor 5: “Política de admisión orientada al sector geográfico”:

- Residencia del alumno.
- Competencia.

Tabla 2.
Estadísticas descriptivas de las variables explicativas

Variable	Descripción de la variable	Media	Mínimo	Máximo	Desviación	CHILE
Residencia del alumno	Porcentaje de estudiantes de los establecimientos donde el director reportó la residencia del alumno como prerrequisito o alta prioridad para admitirlo en el establecimiento.	39,854	1,583	93,669	27,264	6,812
Competencia	Porcentaje de estudiantes de los establecimientos donde el director reportó dos o más otros establecimientos que competían por estudiantes en la misma área.	59,064	21,754	89,992	18,733	63,907
Publicidad	Porcentaje de estudiantes de las escuelas donde el director reportó que la información de logro era utilizada con fines publicitarios.	38,132	4,469	92,721	24,335	37,818
Desempeño del director	Porcentaje de estudiantes de las escuelas donde el director reportó que la información de logro era usada para evaluar al director.	40,017	2,959	91,093	27,208	39,412
Desempeño de los profesores	Porcentaje de estudiantes de las escuelas donde el director reportó que la información de logro era usada para evaluar el desempeño de los profesores.	57,296	8,264	100,000	28,824	56,198
Asignación de recursos	Porcentaje de estudiantes de las escuelas donde el director reportó que la información de logro era utilizada para tomar decisiones de asignación de recursos al interior del establecimiento.	39,205	1,476	86,515	25,951	86,515
Seguimiento autoridad administrativa	Porcentaje de estudiantes de las escuelas donde el director reportó que la información de logro era utilizada por la autoridad administrativa como mecanismo de seguimiento.	68,103	15,684	100,000	20,280	78,920
Contratación de profesores	Porcentaje de estudiantes de los establecimientos donde el director reportó responsabilidad para la contratación de profesores de manera autónoma o en conjunto con la autoridad gubernamental.	68,521	4,321	100,000	34,009	61,908
Despido de profesores	Porcentaje de estudiantes de los establecimientos donde el director reportó responsabilidad para el despido de profesores de manera autónoma o en conjunto con la autoridad gubernamental.	62,692	3,343	100,000	34,693	60,259
Formulación de presupuesto	Porcentaje de estudiantes de los establecimientos donde el director reportó responsabilidad para la formulación del presupuesto del establecimiento de manera autónoma o en conjunto con la autoridad gubernamental.	70,797	11,082	98,689	23,581	70,197
Asignación de presupuesto	Porcentaje de estudiantes de los establecimientos donde el director reportó responsabilidad para la decisión sobre la asignación de presupuesto al interior del establecimiento de manera autónoma o en conjunto con la autoridad gubernamental.	88,373	39,943	100,000	14,743	89,120

(Continuación en página siguiente)

(Continuación Tabla 2)

Variable	Descripción de la variable	Media	Mínimo	Máximo	Desviación	CHILE
Políticas disciplinarias	Porcentaje de estudiantes de las escuelas cuyo director reportó responsabilidad para la definición de las políticas disciplinarias del establecimiento de manera autónoma o en conjunto con la autoridad gubernamental.	94,510	41,883	100,000	12,107	99,157
Políticas de evaluación	Porcentaje de estudiantes de las escuelas cuyo director reportó responsabilidad para la definición de las políticas de evaluación de manera autónoma o en conjunto con la autoridad gubernamental.	81,933	23,650	100,000	20,615	92,430
Admisión de estudiantes	Porcentaje de estudiantes de las escuelas cuyo director reportó responsabilidad para la aprobación de ingreso de nuevos estudiantes de manera autónoma o en conjunto con la autoridad gubernamental.	86,734	27,521	100,000	15,399	98,293
Textos escolares	Porcentaje de estudiantes de las escuelas cuyo director reportó responsabilidad para la elección de los textos escolares a utilizar de manera autónoma o en conjunto con la autoridad gubernamental.	86,454	0,000	100,000	24,900	96,894
Contenidos de los cursos	Porcentaje de estudiantes de las escuelas cuyo director reportó responsabilidad para la definición de los contenidos de los cursos de manera autónoma o en conjunto con la autoridad gubernamental.	67,216	1,547	100,000	28,861	68,184
Cursos ofrecidos	Porcentaje de estudiantes de las escuelas cuyo director reportó responsabilidad para la definición de los cursos a impartir de manera autónoma o en conjunto con la autoridad gubernamental.	73,711	0,599	100,000	29,519	93,013

2. Resultados

Antes de iniciar la evaluación del desempeño de los sistemas educativos conviene realizar un análisis de los datos; en particular de los *outputs* que se utilizaron en el estudio. De este análisis se pudo apreciar que Azerbaiyán, Macao y Hong-Kong (China) son los países que muestran menores niveles de inequidad. Por el contrario, la República Checa, Bulgaria y Argentina son los países con mayores niveles de inequidad en educación. Chile por su parte, muestra un índice de inequidad educativa del 8,69%; esto es, en el punto medio de logro académico entre los resultados de los percentiles 5 y 95, el puntaje se incrementa un 8,69% al aumentar en media desviación estándar el

nivel socioeconómico y cultural de sus estudiantes. Esta cifra es significativamente mayor que la media de los países considerados (7,61%), superior a otros países de la Región como Brasil (7,82%), Colombia (6,86%) y México (6,45%), similar a Uruguay (8,6%) y sólo inferior a Argentina (10,32%). De igual manera, es considerablemente más inequitativo que sistemas educativos líderes en esta materia como Macao (2,5%), Hong-Kong (4,6%), Finlandia (5,59%) e Islandia (5,56%), aunque de características similares a países como Estados Unidos (9,43%) y el Reino Unido (8,85%).

En la Figura 3 se muestra la relación entre las variables de inequidad y logro académico en la prueba de

Lenguaje, lo que deja en evidencia que es posible obtener altos niveles de logro académico conjugados con bajos niveles de inequidad. Por tanto, obtener desempeños similares entre estudiantes de distinto nivel socioeconómico es una meta compatible con objetivos de altos desempeños de los estudiantes. Al respecto, países como Finlandia, Hong-Kong, Canadá, Macao y Corea son ejemplos de este buen desempeño. Por el contrario, países latinoamericanos como Chile, Argentina y Uruguay, y otros, como Bulgaria, Israel y Kirguistán, son la cara opuesta: con logros

académicos y equidad bajo la media de la muestra de países considerados en la evaluación.

Por su parte, de la evaluación del desempeño de los sistemas educativos, los resultados más relevantes de la aplicación de los modelos explicitados en el apartado metodológico que dan cuenta de la eficiencia técnica de gestión (ϕ_1), del máximo *output* potencial de mediano y largo plazo (ϕ_2), y del máximo *output* potencial de muy largo plazo (ϕ_3) de cada país, se muestran en la Tabla 3.

Figura 3.
Logro Prueba de Lenguaje versus niveles de inequidad

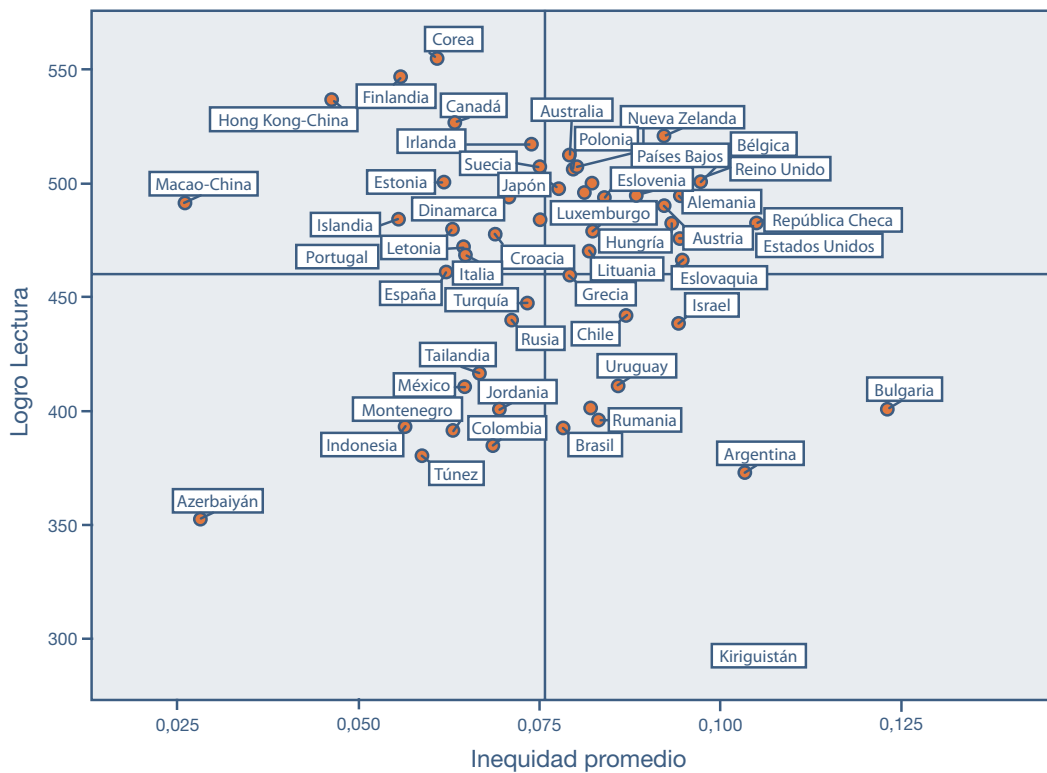


Tabla 3.
Resultados del análisis de la eficiencia y del máximo *output* potencial por país

PAÍS	Eficiencia técnica de gestión	Máximo output medio y largo plazo	Máximo output muy largo plazo	Efecto Recursos	Efecto Entorno	Efecto Gestión
	$(1+\phi_1)$	$(1+\phi_2)$	$(1+\phi_3)$	$\lambda_1 = \frac{\lambda_1}{(1+\phi_2)/(1+\phi_1)}$	$\lambda_2 = \frac{\lambda_2}{(1+\phi_3)/(1+\phi_2)}$	$(1+\phi_1)$
ARGENTINA	1,081	1,439	1,439	1,331	1	1,081
AUSTRALIA	1,062	1,062	1,062	1	1	1,062
AUSTRIA	1,094	1,094	1,094	1	1	1,094
AZERBAIYÁN	1,000	1,130	1,130	1,130	1	1
BÉLGICA	1,078	1,078	1,078	1	1	1,078
BRASIL	1,000	1,236	1,415	1,236	1,145	1
BULGARIA	1,175	1,312	1,312	1,117	1	1,175
CANADÁ	1,047	1,047	1,047	1	1	1,047
CHILE	1,017	1,258	1,258	1,237	1	1,017
CHINA - TAIPEI	1,027	1,027	1,027	1	1	1,027
COLOMBIA	1,191	1,292	1,443	1,084	1,117	1,191
CROACIA	1,066	1,158	1,158	1,086	1	1,066
REPÚBLICA CHECA	1,087	1,087	1,087	1	1	1,087
DINAMARCA	1,101	1,101	1,101	1	1	1,101
ESTONIA	1,045	1,063	1,063	1,018	1	1,045
FINLANDIA	1,000	1,000	1,000	1	1	1
ALEMANIA	1,090	1,090	1,090	1	1	1,090
GRECIA	1,188	1,192	1,192	1,004	1	1,188
HONG-KONG, CHINA	1,020	1,020	1,020	1	1	1,020
HUNGRÍA	1,117	1,117	1,117	1	1	1,117
ISLANDIA	1,115	1,115	1,115	1	1	1,115
INDONESIA	1,000	1,000	1,415	1	1,415	1
IRLANDA	1,075	1,075	1,075	1	1	1,075
ISRAEL	1,241	1,241	1,241	1	1	1,241
ITALIA	1,186	1,186	1,186	1	1	1,186
JAPÓN	1,054	1,054	1,054	1	1	1,054
JORDANIA	1,000	1,380	1,380	1,380	1	1
KIRGUISTÁN	1,000	1,731	1,731	1,731	1	1
COREA	1,000	1,000	1,000	1	1	1

(Continuación en página siguiente)

(Continuación Tabla 3)

PAÍS	Eficiencia técnica de gestión	Máximo output medio y largo plazo	Máximo output muy largo plazo	Efecto Recursos	Efecto Entorno	Efecto Gestión
	$(1+\phi_1)$	$(1+\phi_2)$	$(1+\phi_3)$	$\lambda_1 = (1+\phi_2)/(1+\phi_1)$	$\lambda_2 = (1+\phi_3)/(1+\phi_2)$	$(1+\phi_1)$
LITUANIA	1,131	1,141	1,141	1,009	1	1,131
LUXEMBURGO	1,138	1,138	1,138	1	1	1,138
MACAO-CHINA	1,000	1,000	1,000	1	1	1
MÉXICO	1,032	1,224	1,355	1,186	1,107	1,032
MONTENEGRO	1,000	1,371	1,371	1,370	1	1
PAÍSES BAJOS	1,053	1,053	1,053	1	1	1,053
NUEVA ZELANDA	1,056	1,056	1,056	1	1	1,056
NORUEGA	1,138	1,138	1,138	1	1	1,138
POLONIA	1,095	1,095	1,095	1	1	1,095
PORTUGAL	1,016	1,177	1,177	1,159	1	1,016
RUMANIA	1,156	1,334	1,334	1,154	1	1,156
RUSIA	1,000	1,164	1,164	1,164	1	1
SERBIA	1,246	1,276	1,276	1,024	1	1,246
ESLOVAQUIA	1,123	1,133	1,133	1,009	1	1,123
ESLOVENIA	1,086	1,086	1,086	1	1	1,086
ESPAÑA	1,148	1,148	1,148	1	1	1,148
SUECIA	1,096	1,096	1,096	1	1	1,096
SUIZA	1,067	1,067	1,067	1	1	1,067
TAILANDIA	1,000	1,000	1,326	1	1,326	1
TÚNEZ	1,000	1,190	1,462	1,190	1,229	1
TURQUÍA	1,000	1,025	1,244	1,025	1,214	1
REINO UNIDO	1,101	1,101	1,101	1	1	1,101
ESTADOS UNIDOS	1,154	1,154	1,154	1	1	1,154
URUGUAY	1,124	1,300	1,300	1,157	1	1,124
Media	1,078	1,152	1,183	1,071	1,029	1,078
Mínimo	1,000	1,000	1,000	1,000	1,000	1,000
Máximo	1,246	1,731	1,731	1,731	1,415	1,246
Desviación estándar	0,067	0,133	0,150	0,136	0,084	0,067

2.1 Eficiencia técnica de gestión

En la Tabla 3, la primera columna indica el coeficiente de eficiencia técnica de gestión de los sistemas educativos de cada país ($1+\phi_1$), esto es, la calidad de la gestión de los recursos, considerando las variables del entorno.

Valores igual a la unidad, indican que la gestión del sistema educativo del país es eficiente (porque se obtiene el máximo rendimiento de los recursos destinados al sistema educativo, acorde a las condiciones del entorno). En esta condición se encuentran 13 países, entre ellos: Brasil, Finlandia, Corea, Rusia y Macao. El nivel promedio de eficiencia técnica de gestión de los países considerados en el estudio es de 1.078. Ello implica que, dadas las condiciones del entorno, estos países, en promedio, deberían ser capaces de aumentar en un 7,8% el logro académico de sus estudiantes y reducir la inequidad en similar proporción, manteniendo su actual provisión de recursos y variables observadas del entorno.

En el caso de Chile (que ostenta un índice de eficiencia técnica de 1.017), se puede afirmar que, dados los recursos asignados al sistema educativo y el nivel socioeconómico de los estudiantes, sus resultados están muy próximos a su mejor nivel posible, pudiendo aspirar a aumentar sus resultados de logro académico y equidad con una mejor gestión, solo en un 1,7%. Los países que muestran peores niveles de gestión de sus sistemas educativos son Israel y Serbia (con índices de ineficiencia de gestión superiores al 20%) y en un segundo lugar, se ubican países como Bulgaria, Colombia, Grecia, Italia, Rumania y Estados Unidos (con índices de ineficiencia técnica de gestión superiores al 15%).

2.2 Máximo mejoramiento potencial de mediano y largo plazo

Una vez analizados los resultados relativos a la eficiencia técnica de gestión, la atención se centró en el máximo *output* potencial de mediano y largo plazo. En este sentido, en la segunda columna de la Tabla 3 ($1+\phi_2$), se aprecia que sólo 5 países realizan una gestión y dotación óptima de recursos, dadas sus condiciones de entorno (Finlandia, Indonesia, Corea, Macao y Tailandia). El índice de ineficiencia en la consecución del máximo *output* potencial de mediano y largo plazo promedio de la muestra de 54 países, es de 1.152. Esto implica que, en promedio, los países de la muestra podrían aumentar en un 15,2% sus resultados de logro académico y disminuir la inequidad en el mismo porcentaje, si se dispusiera de los recursos en un nivel óptimo y éstos se gestionaran apropiadamente, sin variar las condiciones del entorno.

Chile, contrario a lo que ocurre con el índice de eficiencia de gestión, muestra en este indicador (de 1.258), que si tuviera una dotación óptima de recursos y una gestión eficiente, el país podría aumentar en un 25,8% sus resultados, sin necesidad de que ocurrieran cambios en sus variables del entorno.

En cuanto al efecto negativo que los recursos ejercen sobre la eficiencia (que se muestra en la columna “efecto recursos”), se puede señalar que en Chile, la deficiente dotación de recursos, tiene un alto impacto sobre los resultados de logro académico e inequidad. De mediar una mejora en este ámbito, Chile podría aumentar sus resultados en un 23,7%.

2.3 Máximo mejoramiento potencial de muy largo plazo

La columna “máximo *output* muy largo plazo” ($1+\phi_3$) da cuenta del máximo resultado que un país podría lograr, si contara con la provisión adecuada de recursos y no se viera impactado negativamente por las condiciones

del entorno¹³. En este ámbito, como se puede apreciar, Chile no se ve especialmente perjudicado por sus condiciones del entorno ni por la provisión de recursos asociada al nivel socioeconómico. Esto ocurre porque Chile mantiene su índice de máximo *output* potencial y por ende, comparativamente, las variables socioeconómicas de la población provocan un nulo efecto sobre el máximo resultado que podría alcanzar el país. Esta situación no es igual para otros países de la Región, como Brasil, Colombia y México; para los cuales la obtención de los máximos resultados potenciales está mediada de manera importante por las variables socioculturales de la población, más allá de su provisión de recursos.

Ahora bien, en términos generales, el promedio de impacto de las condiciones del entorno sobre los niveles de máximo *output* es el más bajo de entre las tres posibles causas, llegando en promedio al 2,9%. Ello se explica por la interrelación entre los recursos de las escuelas y el nivel socioeconómico del país, que hace que las variables socioeconómicas de cada país afecten de manera marginal a la provisión de recursos para educación que el país provee.

La considerable mayor importancia de la ineficiencia de gestión de los sistemas educativos (7,8%) por sobre la atribuible a los factores ambientales (2,9%) no coincide con lo reportado en estudios a nivel de establecimientos educativos, como los de Mancebón y Mar-Molinero (2000) y Silva-Portela y Thanassoulis (2001). De hecho, estos estudios concluyen que los factores más importantes son el entorno familiar, el nivel económico (o capacidad económica) y las cualidades innatas de los alumnos.

Estas diferencias en los resultados de los análisis, probablemente se deben a la estrecha relación que existe entre las características socioeconómicas y la provisión de recursos. En efecto, las características socioeconómicas del país determinan la cantidad y calidad de los recursos educativos puestos a disposición del sistema, y por ende, la eficiencia técnica de gestión se ve afectada.

2.4 Clasificación de los sistemas educativos a partir de su problemática

Las tres últimas columnas de la Tabla 3 dan cuenta de las tres principales causas que explican la no obtención de los máximos resultados potenciales que podría alcanzar un país:

- La columna $(1+\phi_1)$ señala cuánto podrían mejorar los resultados, si mejorase la gestión de los sistemas educativos.
- La columna λ_1 , indica cuánto podrían mejorar los resultados, si existiese una óptima provisión de recursos.
- La columna λ_2 , muestra cuánto podrían mejorar los resultados, si las condiciones socioeconómicas y culturales del país fuesen las óptimas.

De estas columnas, se puede apreciar que en Chile, los problemas de gestión son bajos (1,7%) y su principal debilidad es la dotación de recursos (23,7%), no siendo el entorno un elemento que condicione sus resultados. Lo anterior contradice el estudio de Beyer (2001) —quien señala que el bajo desempeño de los alumnos chilenos, obedece a deficiencias de la gestión de su sistema educativo— pero es consistente con un estudio similar realizado con los datos de la Prueba Timss de 1999 (Giménez, Prior y Thieme, 2007).

En la Figura 4 se esquematizan las causas que explican la no consecución de los máximos resultados potenciales por parte de los distintos países participantes en PISA 2006, las cuales pueden ser agrupadas en

13 Este indicador es similar a $(1+\phi_2)$, salvo que se ha eliminado la restricción de las variables del entorno, y por tanto se comparan con una situación socioeconómica óptima. El aporte de este coeficiente se obtiene al relacionar ambos coeficientes de máximo *output* ($\lambda_2 = (1+\phi_3) / (1+\phi_2)$) que corresponde al impacto negativo de los factores del entorno, en los resultados máximos que puede obtener un sistema.



7 categorías, según la naturaleza de su problemática, la que, a su vez, se asocia con tres factores (entorno, recursos y gestión).

En los países más distantes de aquellos con óptimos resultados, la principal problemática es el entorno o una combinación de entorno y recursos. Luego, acercándose a los países con óptimos resultados, les siguen aquellos cuyo principal problema son los recursos (a veces acompañados de problemas de gestión). Finalmente, los países más cercanos a aquellos con óptimos resultados, son aquellos cuya única problemática es la gestión.

Como puede apreciarse en la Figura 4, sólo tres países (Grupo 1) tienen sistemas educativos óptimos (gestión, dotación de recursos e impacto de los factores del entorno): Finlandia, Corea y Macao-China. Estos países

son eficientes en la gestión de sus recursos, disponen de una dotación óptima de los mismos y cuentan con condiciones óptimas del entorno.

El Grupo 2 es el más numeroso (26 países) e incorpora los sistemas educativos que en promedio más se acercan a los máximos resultados. En este grupo se encuentra la gran mayoría de los países, principalmente europeos, norteamericanos y asiáticos. En promedio, este grupo muestra un índice de máximo *output* potencial de largo plazo de 1.097.

En el Grupo 3 (donde se encuentra Chile), la media del indicador de máximo *output* potencial es de 1.235. Las principales problemáticas de estos países son la gestión y la dotación de sus recursos. Este grupo, además de Chile, está compuesto por otros 13 países, entre ellos, Argentina y Uruguay.

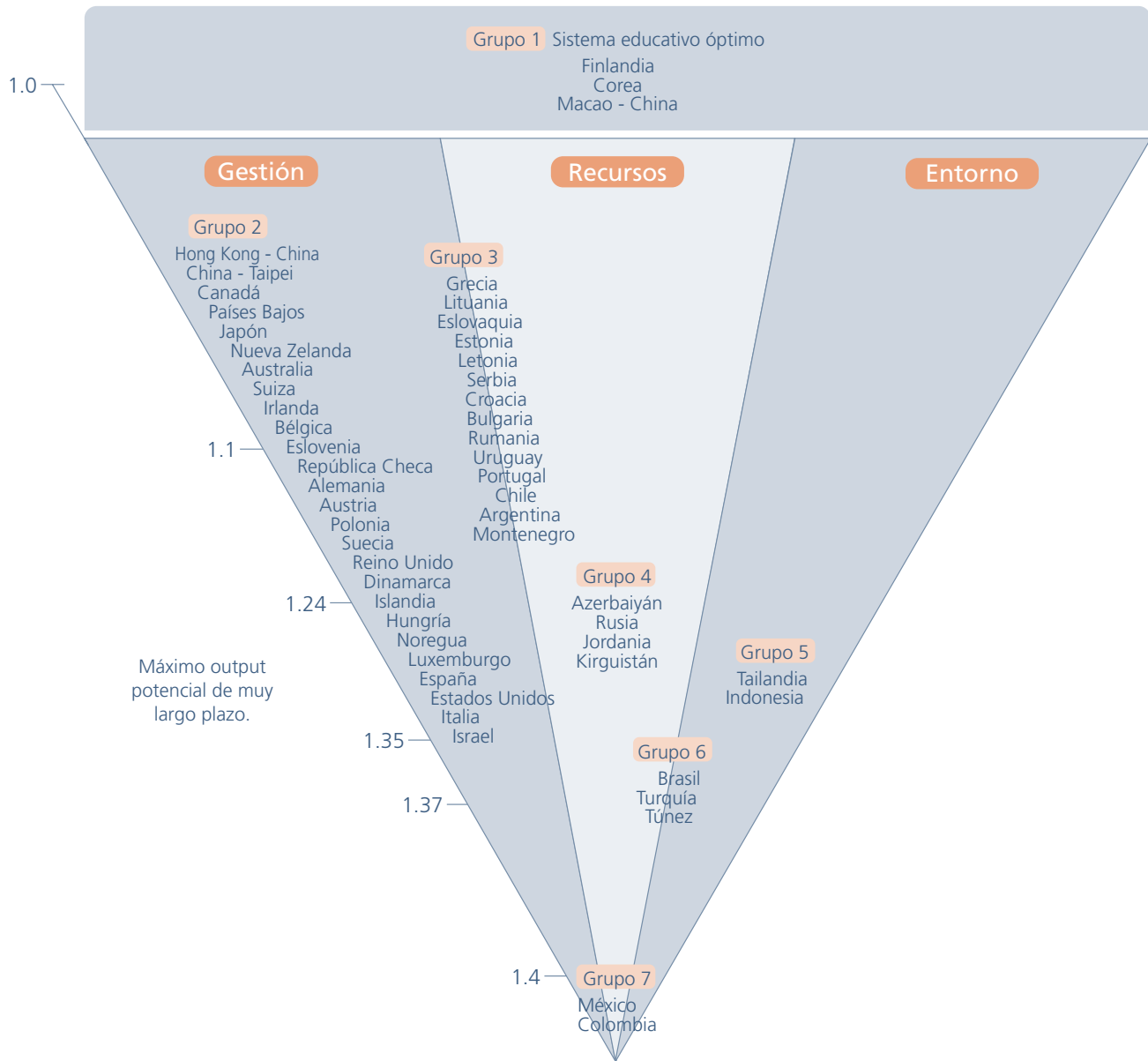
En el Grupo 4 se encuentran países como Azerbaiyán, Rusia, Jordania y Kirguistán, que tienen en común su eficiencia en gestión, pero que deben mejorar la dotación de recursos al sistema educativo. Este grupo tiene una media de crecimiento potencial de 1.351.

El Grupo 5 corresponde a los países cuya única vía de mejora son los factores del entorno. En este grupo se encuentran sólo dos países: Tailandia e Indonesia.

El Grupo 6 corresponde a países que, junto con la necesidad de mejorar sus condiciones del entorno, requieren modificar la dotación de recursos. En este grupo se encuentran Brasil, Turquía y Túnez.

Por último, en el Grupo 7, se encuentran dos países de la región (México y Colombia), que presentan problemas en todos los ámbitos: gestión, dotación de sus recursos y condiciones del entorno.

Figura 4.
Clasificación de los países de acuerdo con su problemática en la obtención de máximo *output* potencial de muy largo plazo



2.5 Variables explicativas del desempeño de los sistemas educativos

La Tabla 4 da cuenta de los resultados de la regresión con variables truncadas, considerando como variable dependiente el indicador de máximo mejoramiento potencial de mediano y largo plazo, y como variables independientes, cada uno de los cinco factores expresados anteriormente.

Los resultados mostraron que la “capacidad de los establecimientos para la toma de decisiones autónoma o de manera conjunta con la autoridad, en materia académica” (factor 1) y la “capacidad de los establecimiento para la toma de decisiones

autónomas o de manera conjunta con la autoridad, en la formulación y asignación de presupuestos” (factor 4), resultan estadísticamente significativos para explicar las diferencias de desempeño en la consecución del máximo mejoramiento potencial de mediano y largo plazo. El signo negativo del coeficiente de regresión es acorde a lo esperado, representando que a mayor capacidad del establecimiento para la toma de decisiones en materia académica y presupuestaria, es menor la distancia que lo separa del máximo *output* potencial que ese sistema educativo podría acceder con niveles óptimos de gestión y recursos, dada su actual situación de variables socioeconómicas del entorno.

Tabla 4.
Análisis de regresión con variables truncadas

$1+\phi_2$	Coef.	Error estándar	z	P>z	95% [Intervalo Conf.]	
EQ1						
Factor 1	-0,071	0,031	-2,3	0,021	-0,131	-0,010
Factor 2	0,062	0,034	1,85	0,064	-0,004	0,129
Factor 3	-0,046	0,030	-1,5	0,133	-0,105	0,014
Factor 4	-0,075	0,030	-2,47	0,013	-0,135	-0,016
Factor 5	-0,040	0,038	-1,06	0,291	-0,114	0,341
Const		0,061	17,56	0	0,95	1,189
SIGMA						
Const	0,146	0,029	4,98	0	0,884	0,020

Regresión truncada

Límite inferior = 1

Superior = +inf

Log likelihood = 46.639435

Número de casos = 45

Wald $\chi^2(5) = 12.14$

Prob > $\chi^2 = 0.0329$

3. Conclusiones

Asegurar una educación de calidad y con equidad para toda la población, tiene implicancias, no sólo en el ámbito social, sino también en el político y económico. Por tal motivo, lograr una correcta gestión de los sistemas educativos y una adecuada provisión de los recursos educativos y humanos, resultan ser objetivos generalmente aceptados.

A partir del cruce de variables de resultados deseados (logro académico) y no deseados (inequidad educativa) se pudo apreciar que es posible obtener altos logros de aprendizaje de los estudiantes, acompañados de bajos niveles de inequidad (como ocurre con Finlandia, Hong-Kong, Corea y Canadá), pero que, en la mayoría de los países, se requiere mejorar en una o en ambas direcciones. Sin embargo, un análisis de estas características, que sólo considera variables de resultados, es sólo un primer referente para evaluar el desempeño de los sistemas educativos nacionales. Una correcta comparación debe considerar los recursos controlables y del entorno que cada país tuvo a su disposición para obtener esos resultados.

De la evaluación de los 54 sistemas educativos considerados en este estudio (realizada desde el punto de vista de la eficiencia técnica, de la provisión de los recursos y del impacto de los factores socioeconómicos y culturales sobre los indicadores de desempeño), se puede concluir que el mayor problema de los sistemas educativos, tanto por el número de países afectado, como por el impacto en la obtención de máximos resultados potenciales, es la gestión de los recursos. En efecto, si el conjunto de países que componen la muestra tuviera una adecuada gestión de sus sistemas educativos, podría mejorar sus resultados de aprendizaje, en un promedio de 7,8% y disminuir en el mismo porcentaje la inequidad educativa.

Una segunda problemática que queda relevada, por su honda repercusión en los resultados de aprendizaje,

es la de la dotación de los recursos, tanto humanos como educativos. Su impacto promedio (7,1%) es muy similar al de la gestión, aunque afecta a un menor número de países.

El análisis también permite determinar que, pese a lo que podría creerse, la influencia de los factores del entorno sobre los resultados potenciales que podría alcanzar un país es muy menor: corresponde, en promedio, a sólo el 2,9%. Lo anterior no implica que ello no afecte el desempeño de los sistemas educativos, sino que su impacto ya ha sido absorbido en gran medida por la provisión de los recursos que cada país destina a educación.

Ahora bien, la distancia que separa a la mayoría de los países de aquellos que ostentan sistemas óptimos, como Finlandia, Corea y Macao, no es la misma (hay algunos más cercanos al óptimo y otros muy alejados) ni responde a problemas asociados a un mismo factor. Así, en los países con desempeños cercanos a las mejores prácticas (26 países), sólo se evidencian problemas de gestión. En los que se encuentran inmediatamente después de ese grupo (14 países, entre los cuales está Chile, Argentina y Uruguay), se observan, problemas de gestión y de dotación de recursos. Su índice promedio de máximo *output* potencial de muy largo plazo (de 1,23) da cuenta de la profundidad del impacto que tienen ambas problemáticas sobre los resultados que obtienen estos sistemas educativos.

En el caso de Chile, los problemas de gestión son bajos, su principal debilidad es la dotación de recursos, que tiene un impacto negativo del 23,7% en el máximo *output* potencial de muy largo plazo, no siendo el entorno un elemento que condicione sus resultados. Lo anterior pone de manifiesto la necesidad imperiosa de aumentar significativamente la dotación y la calidad de los recursos puestos a disposición de la educación chilena. Un aumento de tales recursos (acompañado de una mejora en la gestión) podría traducirse en un considerable aumento del logro

académico y de la equidad. La dotación óptima de recursos que le permitiría a Chile alcanzar el máximo *output* de mediano y largo plazo, corresponde a un aumento del 51,8% en el factor de disponibilidad y calidad de los recursos educativos, y del 87,8%, en el factor de disponibilidad y calidad de los recursos humanos para la enseñanza.

Un hecho que llama la atención del caso chileno, y de otros países de relativo bajo nivel socioeconómico y cultural, es el nulo impacto de los factores socioeconómicos de la población, lo que podría aparecer como contra intuitivo. Ello se comprende a cabalidad, cuando se explicitan los países que sirven de referencia. En efecto, cuando se compara el desempeño de Chile considerando la frontera con variables observadas, el conjunto de países que son la unidad de referencia, se construye a partir de una combinación lineal entre Hong-Kong, Macao (China) y Turquía. Todos ellos, de similar nivel socioeconómico y cultural, pero que, sin embargo, obtienen resultados de logro y equidad significativamente superiores al nuestro. Por tal motivo, para Chile, el entorno socioeconómico no es una excusa para no obtener resultados superiores, ya que otros países de similares características, son capaces de lograrlos.

En cuanto a las variables asociadas al proceso educativo nacional, los resultados arrojaron dos factores explicativos de las distancias que separan a los distintos países en la consecución de sus máximos resultados potenciales: “Toma de decisiones en la formulación de presupuestos” y “Toma de decisiones en aspectos académicos”. En efecto, el análisis empírico mostró que los mejores resultados se asocian a la mayor autonomía de los sistemas educativos (o a la existencia de toma de decisiones compartidas), en materia académica y de formulación y asignación de presupuestos. Esa mayor autonomía, consiste en que la gestión de los centros educativos es desarrollada por equipos directivos que, en el proceso de toma de decisiones, saben asumir prácticas y comportamientos

gerenciales. Sin embargo, se debe ser cauto en asignar a este factor una causalidad única, pues el hecho de que exista esta mayor autonomía, puede responder a los niveles de desempeño de los establecimientos.

Para finalizar, cabe señalar las limitaciones que tiene el presente estudio. En primer lugar, se debe reiterar que la comparación entre sistemas educativos supone incorporar las peculiaridades de cada uno, en un modelo que siempre resultará insuficiente. En segundo lugar, los resultados y conclusiones son válidos para una muestra que no es estadísticamente representativa de la realidad mundial, por cuanto la inclusión de cada país en la prueba PISA no es aleatoria. Finalmente, las principales variables consideradas en el estudio, fueron construidas a partir de la información disponible; mucha de la cual es autoreportada.

Referencias

- Álvarez, A. (Ed.) (2001). *La Medición de la Eficiencia y la Productividad*. Madrid, España: Ediciones Pirámide.
- Berger, M. y Toma, E. (1994). “Variation in state education policies and effects on student performance”. *Journal of Policy and Management*, 13(3).
- Bessent, A. y Bessent, E. (1980). “Determining the Comparative Efficiency of Schools through Data Envelopment Analysis”. *Educational Administration Quarterly*, 16. Austin, Texas: University Council for Educational Administration, University of Texas.
- Bessent, A. y otros (1982). “An Application of Mathematical Programming to Assess Productivity in the Houston Independent School District”. *Management Science*, 28. Hanover, Maryland: Institute for Operations Research and the Management Sciences.
- Beyer, H. (2001). “Falencias institucionales en educación: reflexiones a propósito de los resultados del Timss”. *Estudios Públicos*, 82. Santiago, Chile: Centro de Estudios Públicos (CEP).
- Briec, W. (2000). “An extended Färe – Lovell technical efficiency measure”. *International Journal of Production Economics*, 65. Linköping, Suecia: Elsevier.
- Cooper, W., Seiford, L. y Tone, K. (2006). *Introduction to Data Envelopment Analysis and Its Uses: With DEA Solver Software and References*. Nueva York: Springer.

- Färe, R. y Grosskopf, S. (2000). "Theory and Application of Directional Distance Functions". *Journal of Productivity Analysis*, 13. Nueva York: Springer.
- Färe, R., Grosskopf, S. y Weber, W. (1989). "Measuring School District Performance", *Public Finance Quarterly*, 17. Thousand Oaks, California: Sage.
- García, E., Gil, J. y Rodríguez, G. (2000). *Análisis Factorial*. Madrid, España: La Muralla.
- Giménez, V., Prior, D. y Thieme, C. (2007). "Technical efficiency, managerial efficiency and objective-setting in the educational system". *Journal of the Operational Research Society*, 58. Hampshire, Reino Unido: Palgrave Macmillan.
- Gross, M. (2003). "Educational Systems and perceived social inequality". *European Societies*, 5. Viena, Austria: European Sociological Association.
- Hanushek, E. y Taylor, L. (1990). "Alternative assessments of the performance of schools: measurement of state variations in achievement". *Journal of Human Resources*, 25(2). Madison, Wisconsin: The University of Wisconsin Press.
- Lozano-Vivas, A., Pastor, J.T. y Pastor, J.M. (2002). "An efficiency comparison of European banking systems operating under different environmental conditions". *Journal of Productivity Analysis*, 18(1). Nueva York: Springer.
- Mancebón, M.J. (2003). *La evaluación de la eficiencia de los centros educativos públicos*. Aragón, España: Consejería de Educación y Ciencia, Diputación General de Aragón.
- Mancebón, M.J. y Mar Molinero, C. (2000). "Performance in Primary Schools", *Journal of the Operational Research Society*, 51. Hampshire, Reino Unido: Palgrave Macmillan.
- Muñiz, M.A. (2002). "Separating managerial inefficiency and external conditions in data envelopment analysis". *European Journal of Operational Research*, 143. Oxford, Reino Unido: Elsevier B.V.
- Pedraja, F. y Salinas, J. (1996). "Eficiencia del Gasto Público en Educación Secundaria: Una Aplicación de la Técnica Envolvente de Datos", *Hacienda Pública Española*, 138. Madrid, España: Instituto de Estudios Fiscales, Ministerio de Hacienda.
- Peña, C. (2004). "Igualdad educativa y sociedad democrática", en García Huidobro (ed.), *Políticas Educativas y Calidad. Reflexiones del Seminario Internacional*. Santiago, Chile: Fundación Ford/Universidad Alberto Hurtado.
- Prior, D. (1992). "Los modelos frontera en la evaluación de la productividad". *ESICMarket Economic and Business Journal*. Octubre – Diciembre. Madrid, España: ESIC Business & Marketing School.
- Prior, D., Verges, J. y Vilardell, I. (1993). *La evaluación de la eficiencia en los sectores privado y público*. Madrid, España: Instituto de Estudios Fiscales.
- Puryear, J. (2003). "La educación en América Latina: Problemas y desafíos", en Gajardo, M. y Puryear, J. (eds.), *Formas y reformas de la educación en América Latina*. Santiago, Chile: Lom Ediciones.
- Silva-Portela, M. y Thanassoulis, E. (2001). "Decomposing School and School-Type Efficiency". *European Journal of Operational Research*, 132. Oxford, Reino Unido: Elsevier B.V.
- Smith, P. y Mayston, D. (1987). "Measuring Efficiency in the Public Sector". *OMEGA International Journal of Management Science*, 15. Dearborn, Michigan: School of Management, University of Michigan.
- Teddlie, Ch. y Reynolds, D. (2000). *The international handbook of school effectiveness research*. Oxford, Reino Unido: Routledge.
- Thanassoulis, E. y Dunstan, P. (1994). "Guiding Schools to Improved Performance Using Data Envelopment Analysis: An Illustration with Data from a Local Education Authority". *Journal of the Operational Research Society*, 45. Hampshire, Reino Unido: Palgrave Macmillan.
- Thieme, C. (2005). "Eficiencia Técnica y Eficacia Potencial en los establecimientos chilenos de educación básica". *Oikos*, Año 8, 18 (2). Santiago, Chile: Escuela de Administración y Economía, Universidad Católica Silva Henríquez.
- White, K. (1982). "The relation between socio-economic status and academic achievement". *Psychological Bulletin*, 91. Washington: American Psychological Association (APA).

Distribución de puntajes SIMCE en PISA y SIMCE 2006, en perspectiva comparada

Por
Catalina Fernández¹

Resumen

El presente estudio tuvo por objeto determinar la calidad predictiva de la muestra de estudiantes de 2° Medio que rindió la prueba PISA 2006. Para ello, se compararon las distribuciones de los puntajes obtenidos por dichos alumnos en la prueba SIMCE 2006, con la de todos los estudiantes de 2° Medio a nivel nacional. Los hallazgos sugieren que la calidad predictiva de la muestra está dentro de rangos de calidad diseñados, y por lo tanto, es posible establecer equivalencias de puntajes entre PISA y SIMCE, pese a las diferencias en sus escalas. Ello abre la posibilidad de que, en posteriores estudios, se puedan generar escalas de equivalencia entre estos instrumentos, siendo estas escalas, ajustadas por un factor de corrección (denominado sesgo) en agregaciones generales como la de género o nivel socioeconómico, a nivel nacional.

¹ Catalina Fernández integra el Equipo de Análisis y Comunicación de Resultados SIMCE, del Ministerio de Educación.

Introducción

En el año 2006, en Chile, se realizaron dos evaluaciones: la evaluación nacional SIMCE, que, como es habitual, evaluó a todos los estudiantes que cursaban 2° Año de Educación Media en el país, y la evaluación internacional PISA, que tradicionalmente evalúa a una muestra de estudiantes de 15 años de edad, y que no considera en su composición, el nivel escolar que cursan los estudiantes, sino variables asociadas al establecimiento. Durante ese año, se ofreció ampliar la muestra, en función de los requerimientos de cada país, razón por la cual la prueba PISA también se aplicó a una muestra de estudiantes de 2° Medio, representativa de los estudiantes que cursan ese grado a nivel nacional, a la que, en adelante, denominaremos “muestra ampliada”.

Las pruebas SIMCE y PISA poseen diversas características distintivas. Así, mientras las pruebas SIMCE miden los logros de aprendizaje de los estudiantes respecto del Marco Curricular² (Objetivos Fundamentales y Contenidos Mínimos Obligatorios); PISA busca medir el grado en que los jóvenes de 15 años se encuentran preparados para enfrentarse a los retos de la vida adulta. Asimismo, aunque ambas evaluaciones están sometidas a un proceso de escalamiento según la metodología de Teoría Respuesta al Ítem³ (TRI); en PISA se utiliza un modelo que emplea un parámetro único (Modelo *Rasch*) en el cual las curvas características de los ítems solo dependen de la dificultad de éstos, y en SIMCE, en cambio, se utiliza un modelo de tres parámetros, por lo que la curva característica de las preguntas depende de tres factores: (a) el parámetro de dificultad del ítem; (b) el parámetro de

discriminación del ítem; y (c) el parámetro de “acierto por adivinación” (o “azar”). Este último, cubre el hecho de que, en una prueba de selección múltiple, todos los estudiantes tienen alguna oportunidad de contestar el ítem correctamente por azar, independientemente de la dificultad del mismo. Otro aspecto distintivo es la escala: mientras la calificación global de PISA⁴ se expresa en unidades de una escala con media de 500 puntos⁵ y desviación típica de 100; la calificación global como punto de partida, en SIMCE, es de 250 puntos, con desviación típica de 50. Este promedio es fijado una sola vez para cada nivel y subsector⁶, lo que permite realizar comparaciones entre aplicaciones.

Considerando que ambas evaluaciones tienen distintas escalas y miden competencias y habilidades distintas, surge la pregunta sobre el grado de correspondencia entre ellas. Si estas resultasen correspondientes, a partir de una escala de equivalencias horizontal entre SIMCE-PISA, se podría inferir el comportamiento de los resultados nacionales a nivel del establecimiento, en la prueba PISA, considerando agregaciones reportadas por SIMCE y no contempladas explícitamente en PISA.

Para abordar este problema, en este estudio se realizó una comparación de distribuciones entre los resultados obtenidos por los estudiantes en la prueba censal SIMCE y la muestra ampliada PISA, en términos de los puntajes SIMCE, y se examinó la calidad predictiva de la muestra ampliada, respecto de la población nacional que rinde SIMCE.

En concreto, se plantearon los siguientes objetivos generales:

2 Para mayor información acerca de los antecedentes generales de las pruebas SIMCE, véase www.simce.cl.

3 Según esta teoría, las pruebas se analizan considerando que las preguntas tienen comportamientos independientes y regidos sólo por la habilidad en el constructo medido por la prueba. El comportamiento de las preguntas se resume en su curva característica, que se modela a partir de distintas funciones. Para mayor información, véase OCDE, 2003b.

4 El año 2006 los resultados para Chile en PISA fueron: 438 puntos en Ciencias, 442 puntos en Lectura y 411 puntos en Matemática. Para mayor detalle de los principales resultados de Chile en Pisa 2006, véase www.simce.cl.

5 Para el conjunto de países pertenecientes a la OCDE.

6 El año 2006 el puntaje promedio SIMCE en Lengua Castellana y Comunicación fue de 254 puntos y en Matemática, de 252 puntos.

- Comparar la distribución de los resultados SIMCE de los estudiantes que pertenecen a la muestra ampliada, con los del total de estudiantes de 2° Medio evaluados el año 2006.
- Evaluar la representatividad de la muestra ampliada de estudiantes de 2° Medio, como predictora de resultados de la prueba SIMCE.

En función de lo anterior, se establecieron los siguientes objetivos específicos:

1. Identificar la distribución de resultados SIMCE 2006 de la muestra ampliada para Lenguaje⁷ y Matemática, globalmente y según grupos poblacionales (género, población nacional y otras características sociales).
2. Identificar la distribución de resultados de todos los estudiantes que rindieron SIMCE 2006 (Matemática y Lenguaje), globalmente y según grupos poblacionales (sexo, población nacional y otras características sociales).
3. Examinar la calidad predictiva de la muestra ampliada, respecto de los resultados generales de la prueba SIMCE 2006 de 2° Medio⁸.

A continuación se especifica la metodología de trabajo utilizada para dar respuesta a los objetivos planteados. Luego, se presentan los resultados obtenidos, para finalizar con las principales conclusiones y recomendaciones que se desprenden de la investigación.

1. Metodología

El presente estudio se realizó sobre la base de los resultados obtenidos por tres subconjuntos de estudiantes: (a) los 5.233 estudiantes de 15 años pertenecientes a 173 establecimientos educacionales que rindieron PISA 2006; (b) la muestra representativa de estudiantes de 2° Medio (muestra ampliada) que rindió PISA 2006, la cual estaba conformada por un total de 4.772 estudiantes⁸, y (c) los 148.139 estudiantes pertenecientes a 2.454 establecimientos, que rindieron SIMCE 2006.

Existen técnicas descriptivas e inferenciales para la comparación de distribuciones de puntajes entre dos grupos. Como cada una de estas técnicas proporciona distintos tipos de información, en este estudio se usaron ambos tipos. En efecto, las técnicas descriptivas son útiles para examinar la factibilidad de que dos distribuciones pertenezcan a la misma familia⁹; en tanto que, las técnicas inferenciales permiten establecer la verosimilitud, en términos probabilísticos, de que dos distribuciones sean equivalentes. Por otra parte, las técnicas descriptivas examinan gráficamente los valores resumen¹⁰ de las distribuciones involucradas o la frecuencia de los distintos valores, usando histogramas. Para aplicar con éxito¹¹ las técnicas inferenciales, es recomendable haber determinado el tipo de distribución que se quiere testear, lo que sólo es factible de hacer una vez que se han examinado descriptivamente las distribuciones. A continuación se describen las técnicas utilizadas en este estudio.

7 Por economía se utiliza el nombre de "Lenguaje" para la prueba SIMCE de "Lengua Castellana y Comunicación".

8 Se consideran en el estudio sólo los estudiantes de 2° Medio que tenían datos para PISA y SIMCE, por lo tanto, el número de estudiantes de este grado analizados es menor que el número de estudiantes de 2° Medio que rindió PISA.

9 Es inútil aplicar una técnica inferencial cuando las técnicas descriptivas han arrojado la no factibilidad de que dos distribuciones provengan de la misma familia.

10 Como los percentiles, que se pueden representar gráficamente.

11 Es cierto que el Test de Kolmogorov-Smirnov se puede aplicar sin conocer las distribuciones subyacentes, pero en ese caso el error de decidir que las distribuciones no son distintas cuando en realidad lo son, es más alto.

1.1 Técnicas descriptivas¹²

1.1.1 Comparación de distribuciones (metodología gráfica): percentiles (BOX-PLOT)

Esta técnica se utilizó para obtener una aproximación gráfica (denominada diagrama de *Box Plot*) de la simetría o asimetría de los puntajes SIMCE, basada en la población nacional que rindió SIMCE y en la muestra ampliada. Estos gráficos permitieron resumir la información utilizando 5 medidas estadísticas: el valor mínimo, el primer cuartil, la mediana, el tercer cuartil y el valor máximo. A partir de esta técnica, se pudo examinar si las medidas estadísticas de ambos grupos tenían similar comportamiento relativo. En caso de que así no fuera, inmediatamente se rechazaría la posibilidad de que ambas distribuciones pertenecieran a la misma familia si, por ejemplo, se advirtiera que existen comparativamente muchos más valores pequeños en una distribución que en la otra.

1.1.2 Distribución (histograma): género

Para poder identificar el tipo o familia¹³ de distribución a la que pertenecen las distribuciones estudiadas, se recurrió al histograma, que consiste en graficar la frecuencia que presenta cada rango de observación. Esta metodología se utilizó para determinar si los datos provenían de una distribución conocida que, en este caso, se supuso como “normal”.

12 Dixon y Massey, 1983. Para mayor información véase el Anexo.

13 Las distribuciones pueden ser del tipo normal, chi-cuadrado, gamma, entre otros.

1.2 Técnicas inferenciales. Comparación estadística de distribuciones: Kolmogorov-Smirnov (K-S)

Esta técnica se utilizó para determinar el grado de coincidencia entre la distribución basada en una muestra y la distribución teórica¹⁴. Esta prueba determinó si era razonable plantear la hipótesis de que los datos provenientes de la muestra, procedían de la distribución teórica planteada.

Generalmente los softwares estadísticos tienen en sus salidas los estimadores, incorporando el *valor-p*, donde el criterio de decisión asociado al estadístico de contraste D observado, queda definido por¹⁵:

$$\text{valor-p} = P(D > D_{\text{observado}} / H_0 \text{ cierta})$$

Para el caso práctico, si el *valor-p* es mayor a 0.05 no es posible rechazar la hipótesis de que los datos provenientes de la muestra tienen una distribución normal.

En la elección de las técnicas descriptivas e inferenciales primaron varios criterios. El primero fue el rango recorrido de la variable que genera la distribución, en este caso los puntajes. Para las variables del tipo continuo o intervalar (que pueden tomar cualquier valor en la recta real o dentro de un intervalo) se pueden generar las medidas estadísticas mencionadas, las cuales no son apropiadas para las variables de tipo categórico. El segundo criterio fue el tamaño de muestra disponible. En este caso, los percentiles pueden ser estimados confiablemente, ya que las muestras disponibles son bastante grandes (mayores a 100 observaciones).

14 En este estudio, la distribución teórica se centró específicamente en una distribución normal.

15 Para obtener mayores antecedentes de la construcción del test K-S, véase la descripción del Anexo.

A continuación se presentan los resultados de este estudio. Primero los relativos a las comparaciones de distribuciones de puntaje según género (con las técnicas descriptivas e inferenciales, por separado) y luego los derivados de la aplicación de los factores de corrección.

2. Resultados

2.1 Comparación de distribuciones de puntaje según género, usando técnicas descriptivas

En Lenguaje, al comparar la distribución de los puntajes SIMCE de la muestra ampliada con los de la población, según género, se observa que son similares. En efecto, como se aprecia en la Figura 1, en ambos grupos analizados, el rango de puntajes de las mujeres se concentra entre 170 y 350 puntos, y el de los hombres, alrededor del mismo rango, pero con un tercer cuartil inferior al de las mujeres.

Ahora bien, para determinar si los datos pueden ser asociados a alguna distribución conocida, se realizó el recuento sobre las frecuencias, y se comparó la distribución de los datos de la muestra ampliada, con los de la población, en términos de puntajes SIMCE. Así, como se observa en la Figura 2, para las mujeres, los puntajes de la muestra ampliada se encuentran levemente desplazados hacia la derecha respecto de los puntajes poblacionales, tanto en Lenguaje como en Matemática. En otras palabras, la esencia de la distribución se captura muy bien con la muestra ampliada, desviándose sólo levemente respecto del promedio poblacional SIMCE. Ahora bien, aunque en Matemática, se presenta una diferencia más ostensible que la observada en Lenguaje, su distribución tampoco se ve alterada. En la Figura 3, se observa que, para los hombres, los puntajes de la muestra ampliada también se encuentran desplazados levemente hacia la derecha, sin embargo, en la prueba de Lenguaje esta conclusión es menos evidente, pues ambas distribuciones se sobreponen casi por completo.

Figura 1.
Distribución por género, en SIMCE Lenguaje

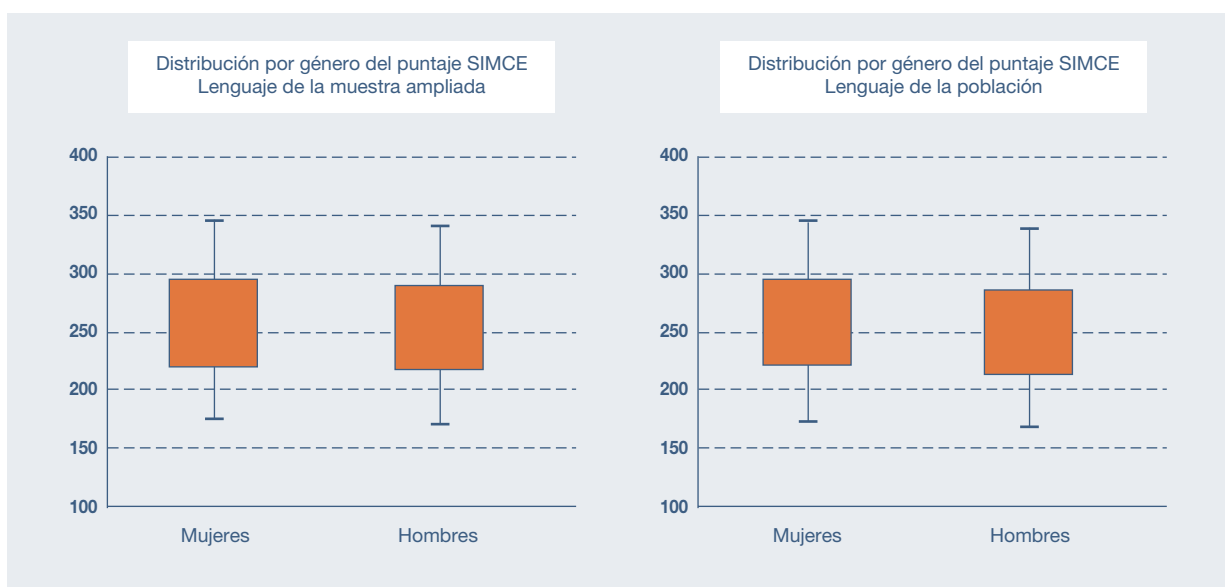


Figura 2.
Distribución del puntaje SIMCE obtenido por las mujeres, en Lenguaje y Matemática

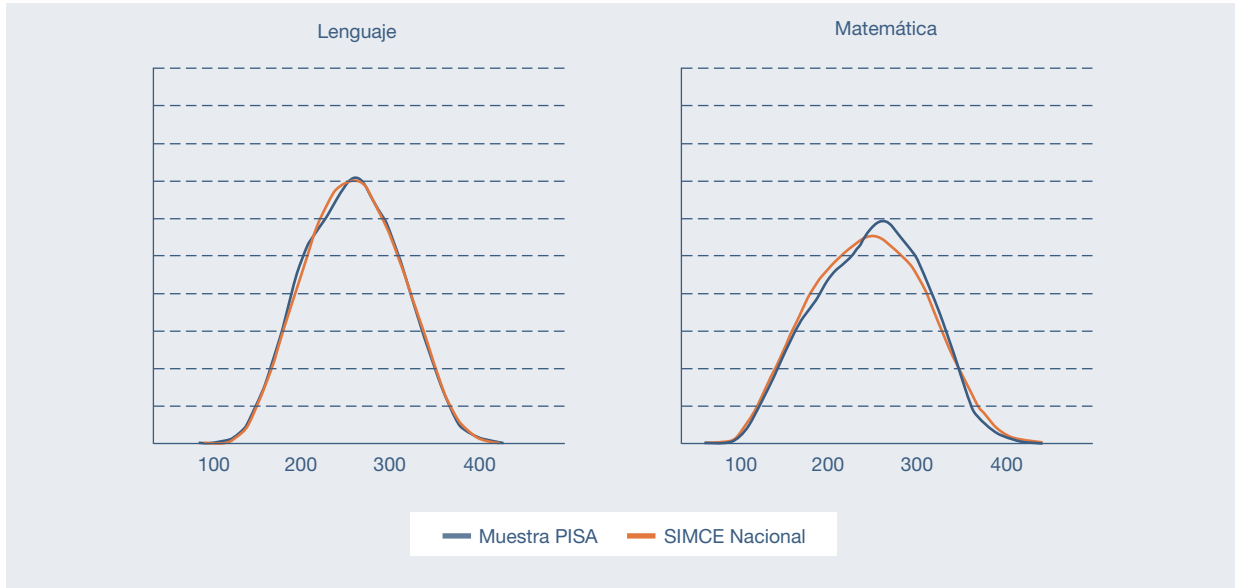
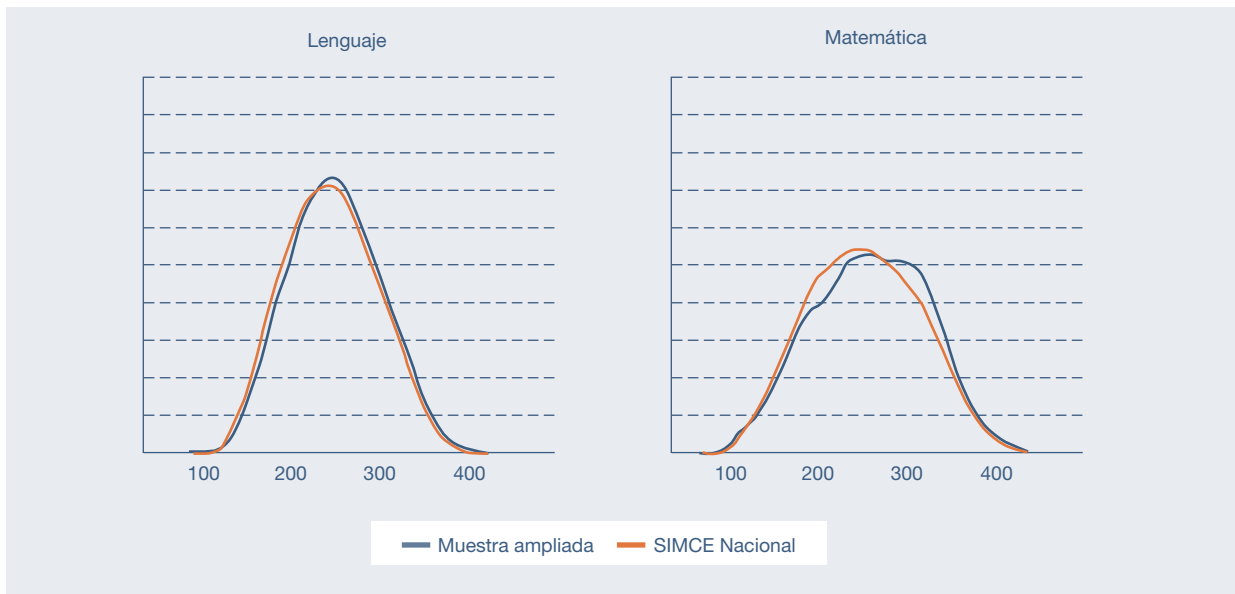


Figura 3.
Distribución del puntaje SIMCE obtenido por los hombres, en Lenguaje y Matemática

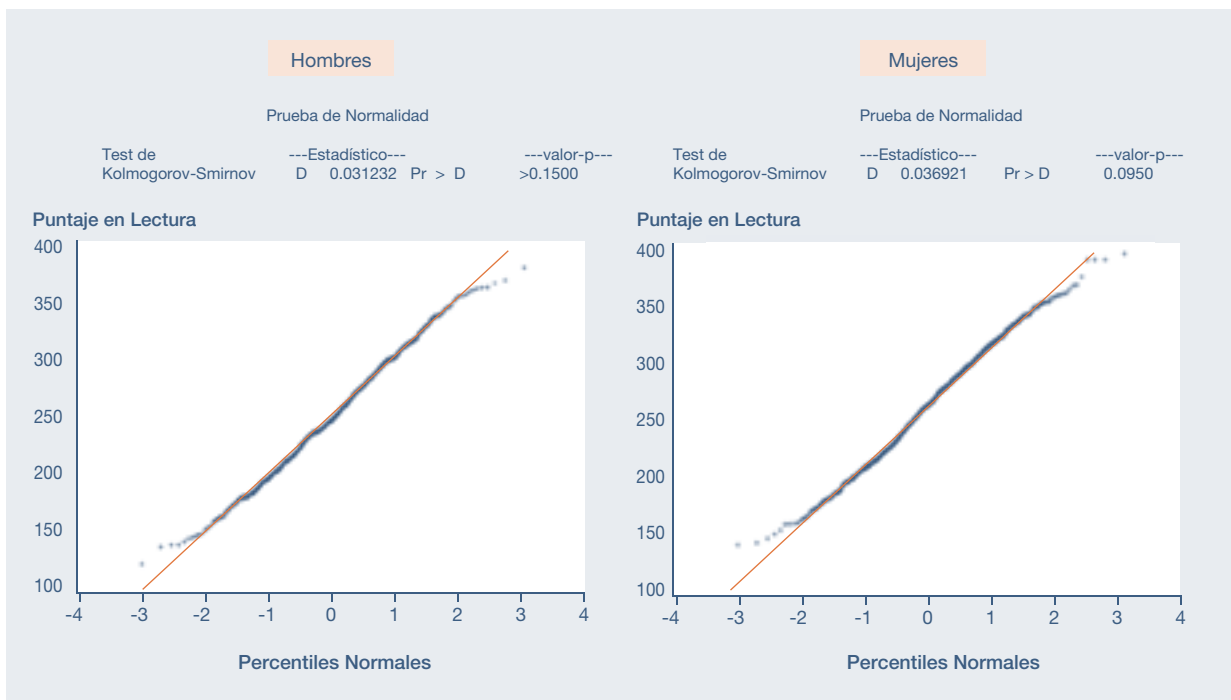


2.2 Comparación de distribuciones de puntaje según género, usando técnicas inferenciales¹⁶

Como se señaló en la sección metodológica, un tercer análisis consistió en determinar si la distribución basada en la muestra ampliada es factible de ser ajustada por medio de una distribución conocida, e incluir el valor del estadístico para crear intervalos de confianza en la estimación de características de interés. Para lo anterior se utilizó el test estadístico

Kolmogorov-Smirnov (K-S). Como se observa en la Figura 4, basado en la distribución de puntajes de los hombres en Lenguaje, los puntos se ajustan casi perfectamente a la línea recta. A partir de los gráficos, se deduce que a nivel de significancia del 5%, no se rechaza la hipótesis nula de normalidad de los datos. En el caso de las mujeres, la conclusión fue la misma: los datos, con una confianza del 95%, provienen de una distribución normal¹⁷, y lo mismo ocurre con la prueba de Matemática¹⁸.

Figura 4.
Gráficos de probabilidad normal, basados en la muestra ampliada, según género, en Lenguaje



¹⁶ Der y Everitt, 2002. Para mayor detalle véase el Anexo.

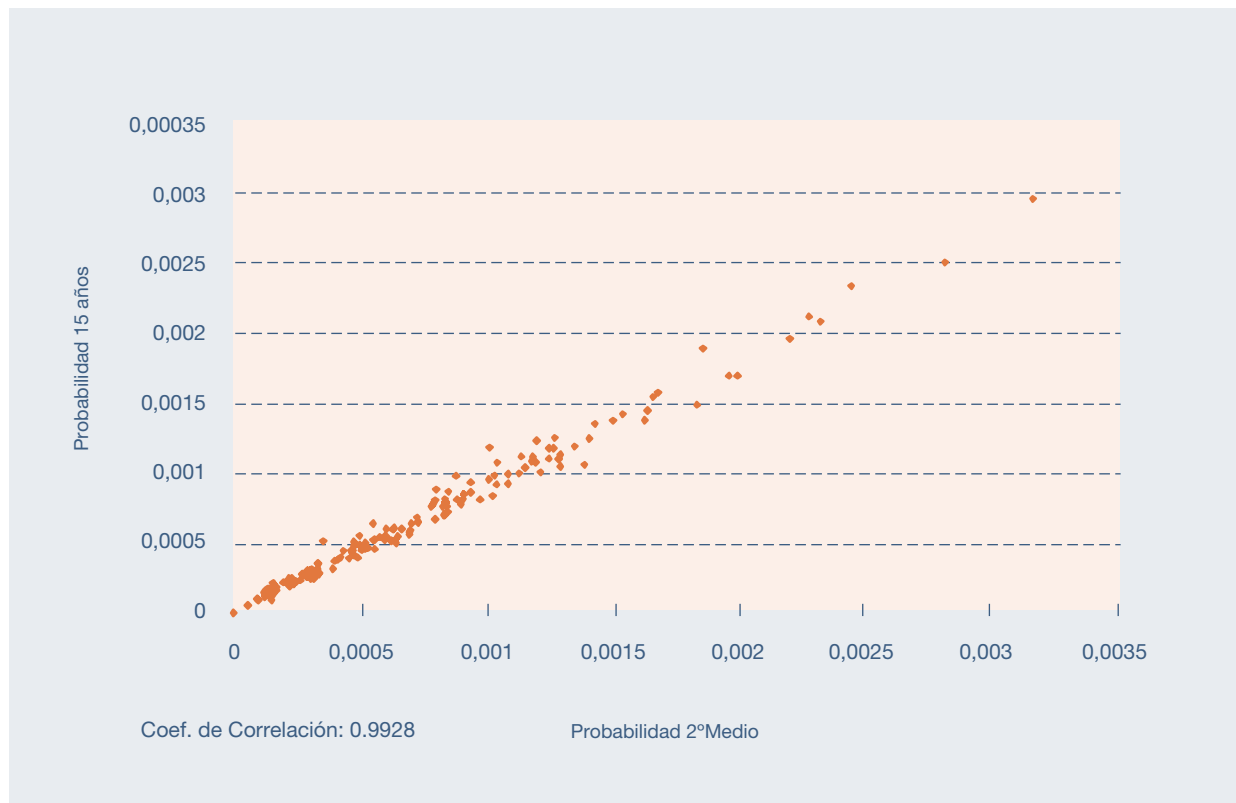
¹⁷ Al establecer que los datos provienen de una distribución normal, se hace más fácil el cálculo de intervalos de confianza para estadísticos de interés. Ello, por cuanto sus valores se encuentran tabulados y disponibles en la mayoría de los softwares estadísticos.

¹⁸ La distribución según género en Matemática, se encuentra en el Anexo, Figura A.1.

Para la selección de los establecimientos que participaron en la prueba PISA (para ambas muestras: la de 15 años y la ampliada), se utilizó un diseño que determinaba la probabilidad de selección de cada establecimiento, de acuerdo con el número de estudiantes de 15 años que estuvieran matriculados. Ahora bien, con el propósito de responder al problema de este estudio, relativo a la capacidad predictiva de la muestra ampliada, se efectuó un análisis de correlación entre la probabilidad

de selección de un establecimiento para la prueba PISA y la probabilidad que tiene un establecimiento de ser seleccionado para participar en la prueba SIMCE de 2° Medio. El resultado de este análisis se muestra en el Gráfico 1, en el que se aprecia que la relación entre selección basada en probabilidades, es directamente proporcional (casi perfecta asociación lineal con correlación 0.99).

Gráfico1: Correlación entre la probabilidad de selección de un establecimiento en la muestra PISA de 15 años y la probabilidad de selección de un establecimiento en la prueba SIMCE 2° Medio 2006



De lo anterior, se puede afirmar que, si bien los establecimientos fueron seleccionados con los estudiantes de 15 años como población objetivo, su probabilidad de selección hubiera sido casi exactamente la misma, si la población objetivo hubieran sido los estudiantes de 2° Medio.

2.3 Factores de corrección

Como se mencionó en la introducción de este artículo, este estudio pretende ser un insumo para establecer la equivalencia de puntajes¹⁹ entre las mediciones PISA y SIMCE. A partir de los datos presentados anteriormente, es posible afirmar que la muestra ampliada es una buena predictora de los resultados SIMCE, sin embargo, es necesario contar con estimaciones de la prueba SIMCE basadas en la muestra ampliada, corrigiendo por el factor de corrección que se obtiene a partir de la muestra ampliada. Al establecer la equivalencia de puntajes y calcular el estimador poblacional, es necesario sumar la discrepancia que se obtiene al calcular parámetros poblacionales, a partir de la muestra ampliada.

Para el cálculo de esta discrepancia se requiere obtener el puntaje promedio según la agregación de interés, tanto con los datos de la muestra ampliada como con los valores poblacionales de los datos SIMCE 2006 de 2° Medio. Luego se calcula la diferencia directa entre estos dos valores, obteniéndose el factor de discrepancia para corregir los datos en términos de equivalencia de puntajes.

A modo de ejemplo, en la Tabla 1, se presenta el factor de discrepancia basado en género, según subsector de aprendizaje²⁰. Como se observa, el factor no supera los 3 puntos en ningún subsector, en ambos géneros. Si se considera una estimación en Lenguaje para los hombres, en base a la muestra ampliada, su puntuación promedio, en términos de equivalencia, debe ser corregida en 1.66 puntos.

Ahora bien, si se desea estimar una puntuación promedio en la prueba nacional SIMCE, por ejemplo, en base a la muestra ampliada, los estadísticos deben ser corregidos por el factor denotado anteriormente.

Tabla 1.
Puntajes promedio según muestra y población, para los subsectores de Lenguaje y Matemática, por género

		Género						
		Muestra PISA 2° Medio (incluido los pesos)			Censal SIMCE 2° Medio			Sesgo
		Promedio	STD	N	Promedio	STD	N	
Hombre	Lenguaje	252.18	51.24	2589	250.516	51.711	118044	-1.664
	Matemática	260.10	66.43	2557	257.229	65.649	118045	-2.871
Mujer	Lenguaje	257.42	51.74	2278	257.867	51.783	124772	0.447
	Matemática	245.91	62.05	2275	247.145	64.019	124693	1.235

19 Utilizando técnicas no-paramétricas, fundadas en métodos de equiparación basados en correspondencia de distribuciones.

20 Para ver los sesgos a nivel de grupo socioeconómico (GSE) y a nivel nacional, véanse las figuras A.2 y A.3 del Anexo.

3. Conclusiones

Aludiendo a los objetivos planteados al inicio de este documento, la principal conclusión que se desprende del estudio realizado es que la muestra ampliada de estudiantes de 2° Medio, es representativa, tanto para las agregaciones a nivel nacional como para las de género, en lo que concierne a los datos de las pruebas SIMCE de 2° Medio 2006, de Lenguaje y Matemática.

Adicionalmente, en base a la metodología K-S, se puede establecer que la distribución de resultados de la muestra ampliada es normal: al comparar las distribuciones de la muestra ampliada y la población nacional que rinde SIMCE, se aprecia una similitud de las distribuciones (tanto en Lenguaje como en Matemática), con un leve desplazamiento hacia la derecha de la muestra ampliada, lo cual indica que, en distribución, tiene un promedio levemente mayor.

En términos de proyección, cabe señalar que los resultados de este estudio, permiten establecer criterios para la equivalencia de escalas entre las pruebas PISA y SIMCE, pudiendo inferir el resultado de los establecimientos del país en la prueba SIMCE, y con ello, examinar la situación de grupos y/o subpoblaciones de estudiantes no contempladas explícitamente en PISA, como por ejemplo, los grupos socioeconómicos SIMCE. Adicionalmente, la calidad

predictiva de la muestra ampliada, permite afirmar que es posible calcular tendencias en base a esta muestra, para compararlas con mediciones futuras SIMCE²¹. En efecto, por constituir una muestra representativa de la población, es viable realizar comparaciones a través del tiempo, aun cuando no sean mediciones consecutivas.

El presente estudio inicia el camino de estudios posteriores para determinar la equivalencia de puntajes entre pruebas de distinta naturaleza y escalas, como las de SIMCE y PISA.

Referencias

- Der, G. y Everitt, B. (2002). *A Handbook of Statistical Analyses Using SAS*. Boca Raton, Florida: Chapman & Hall/CRC Press.
- Dixon, W. y Massey, F. (1983). *Introduction to Statistical Analysis*. Nueva York: McGraw-Hill.
- OCDE (Organización para la Cooperación y el Desarrollo Económico) (2003a). *Pisa 2003 data Analysis Manual SAS Users*. París, Francia.
- _____ (2003b). *PISA 2003 Data Analysis Manual SPSS® USERS; the Rasch Model*. París, Francia.
- SIMCE (Sistema de Medición de la Calidad de la Educación) (2008). *Resultados Nacionales SIMCE*. Santiago, Chile. Extraído desde www.simce.cl

21 En la prueba SIMCE es frecuente que el número y característica de los establecimientos medidos difiera entre un año y otro, por lo cual los promedios generales no son estrictamente comparables.

Anexo

Figura A.1.
Gráficos de probabilidad normal, basados en la muestra ampliada, según género, en Matemática

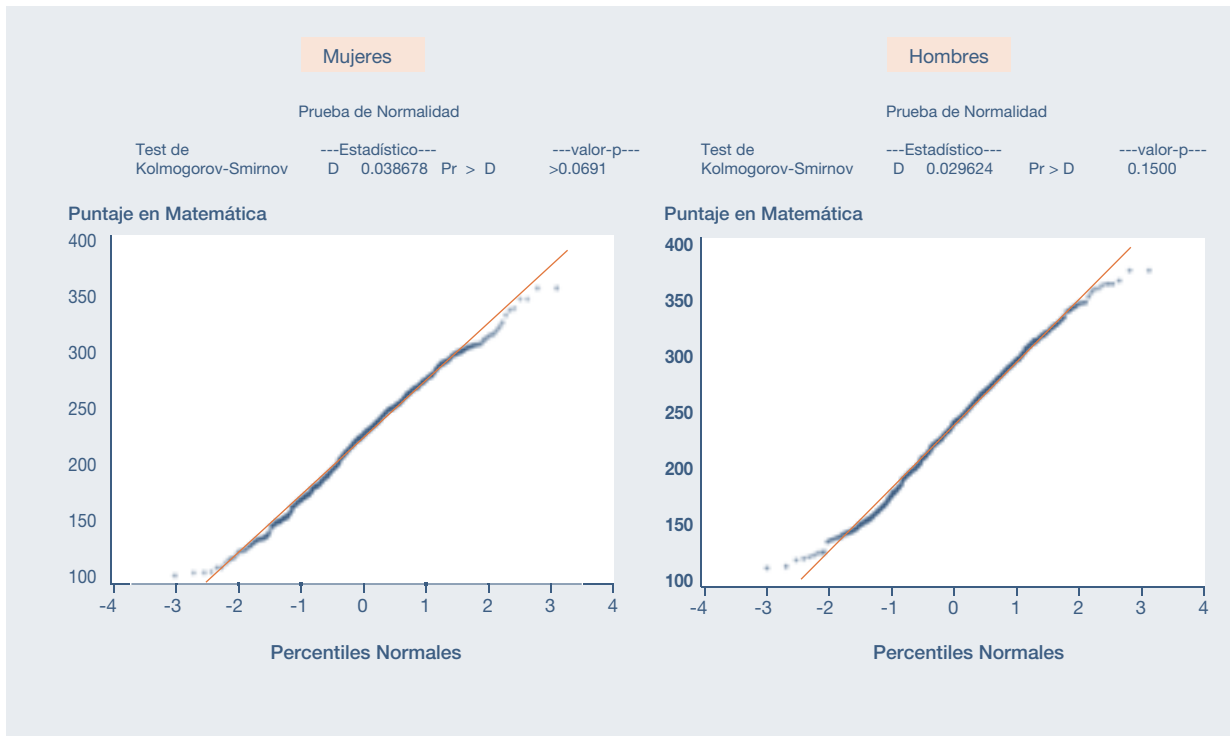


Figura A.1.
Puntajes promedio de la muestra y la población, obtenidos en los subsectores de Lengua Castellana y Comunicación y Matemática, según grupo socioeconómico

GSE								
		Muestra PISA 2° Medio (incluidos los pesos)			Censal SIMCE 2° Medio			Sesgo
		Promedio	STD	N	Promedio	STD	N	
Bajo	Lenguaje	228.16	44.90	789	228.476	44.170	45947	0.316
	Matemática	217.64	51.70	790	218.129	52.843	45994	0.489
Medio Bajo	Lenguaje	239.08	45.21	1917	240.814	46.670	97722	1.734
	Matemática	232.70	56.54	1919	233.505	57.399	97717	0.805
Medio	Lenguaje	271.28	45.50	1298	266.985	47.472	58962	-4.295
	Matemática	274.32	56.86	1296	266.936	58.725	58874	-7.384
Medio Alto	Lenguaje	292.31	46.16	576	290.293	46.586	24936	-2.017
	Matemática	312.15	51.30	574	302.435	56.790	24944	-9.715
Alto	Lenguaje	307.01	40.63	288	306.466	44.226	16348	-0.544
	Matemática	323.75	51.97	254	327.488	50.026	16306	3.738

Figura A.2.
Puntajes promedio de la muestra y la población obtenidos en los subsectores de Lengua Castellana y Comunicación y Matemática, a nivel nacional

Nacional								
		Muestra PISA 2° Medio (incluidos los pesos)			Censal SIMCE 2° Medio 2006			Sesgo
		Promedio	STD	N	Promedio	STD	N	
Lenguaje		254.61	51.54	4868	254.275	51.866	243915	-0.335
Matemática		253.38	64.81	4833	252.013	65.005	243835	-1.367

Definición Prueba de Contraste Kolmogorov-Smirnov

La hipótesis que se contrasta es:

H_0 : Los datos provenientes de la muestra tienen una distribución normal.

H_1 : Los datos provenientes de la muestra no tienen una distribución normal.

donde el estadístico de contraste es definido como:

$$D = \text{Sup}_{1 \leq i \leq n} |\hat{F}_n(x_i) - F_0(x_i)|; \text{ donde}$$

x_i : Se define como el i -ésimo valor observado en la muestra (en donde en primera instancia se definen los estadísticos de orden).

$\hat{F}_n(x_i)$: Se define como el estimador basado en la probabilidad de observar valores menores o iguales a x .

$F_0(x_i)$: Se define como la probabilidad de observar valores menores o iguales que cuando H_0 es cierta.

Al observar el planteamiento anterior se desprende que el estadístico de contraste D representa la mayor diferencia absoluta observada entre la frecuencia

acumulada observada $\hat{F}_n(x)$ y la frecuencia acumulada proveniente de una distribución normal $F_0(x)$.

Para determinar si los valores observados $\hat{F}_n(x)$ son similares a los esperados $F_0(x)$, es necesario prestar atención al estadístico de contraste D . Este será pequeño en caso de que las distribuciones sean similares. Mientras mayor sea la discrepancia entre la distribución empírica $\hat{F}_n(x)$ y la distribución teórica $F_0(x)$, el estadístico de contraste será mayor.

Para decidir, con cierto grado de confianza, cuál de las hipótesis planteadas puede ser considerada como verdadera, se plantea el siguiente criterio:

$\text{Si } D \leq D_\alpha \Rightarrow \text{Aceptar } H_0$

$\text{Si } D > D_\alpha \Rightarrow \text{Rechazar } H_0$

En donde D_α queda definido por:

$$\alpha = P(\text{Rechazar } H_0 / H_0 \text{ cierta}) = P(D > D_\alpha)$$

Los datos siguen una distribución normal.

α : nivel de significancia definido a un nivel arbitrario de 5%.

Rendimiento en la prueba PISA: ¿es posible entender los alcances y límites de las comparaciones entre países?

Por
Jorge González
Ernesto San Martín¹

Resumen

El tipo de comparaciones más común en informes PISA, es aquel que utiliza cantidades estadísticas, como promedio, desviación estándar y cuantiles, que describen la distribución de puntajes obtenidos por los estudiantes. Sin embargo, este tipo de comparaciones es limitada.

El objetivo de este estudio fue mostrar los alcances que tienen estas comparaciones entre países y las limitaciones de las mismas, introduciendo una forma visual que va más allá del reporte y la comparación de estadísticos que describen la distribución de puntajes.

Para ello, se aplicó una metodología complementaria que se basa en estimadores no paramétricos de la densidad de puntajes, la que permitió entender los alcances y los límites de las comparaciones entre países realizadas mediante modelos lineales jerárquicos.

Entre los principales resultados, se encontró que la eficiencia que presenta un país respecto de otro, no está explicada necesariamente por las covariables utilizadas y que, utilizando modelos lineales jerárquicos, es posible realizar comparaciones bajo iguales condiciones.

¹ Jorge González y Ernesto San Martín son profesionales del Centro de Medición MIDE UC de la Pontificia Universidad Católica de Chile. San Martín además es Profesor Asociado del Departamento de Estadística de la Pontificia Universidad Católica de Chile.

Introducción

Además de la información de rendimiento de cada país, uno de los contenidos de los informes PISA que se espera con mucho interés, es el que compara el rendimiento de los diferentes países participantes. Ello, por cuanto la realidad nacional de un país puede ser enriquecida por medio de análisis comparativos internacionales, dado el amplio contexto que los mismos ofrecen para interpretar y comparar los resultados de cada país (OCDE, 2007a).

Las comparaciones que se presentan en los informes PISA (por ejemplo, OCDE, 2007a, b y c), se basan principalmente en descripciones de la distribución de los puntajes obtenidos por los estudiantes. A nivel de país, frases como “Finlandia es el mejor país evaluado en Ciencias, obteniendo un puntaje promedio de 563 puntos”, constituyen un ejemplo de reporte y comparación basado en el promedio de los puntajes (OCDE, 2007c). A nivel de Región (por ejemplo, América Latina) generalmente se reportan, además de comparaciones del puntaje promedio, la variabilidad asociada a los puntajes dentro de cada país. Ambas informaciones (el promedio y la variabilidad asociada) son medidas útiles para la descripción de la distribución de puntajes, pero son limitadas: una comparación más acabada, requeriría, de, por ejemplo, una estimación completa de la distribución de los puntajes.

Otro tipo de análisis que también se presentan en los informes PISA (OCDE, 2007a) es el que permite establecer, por ejemplo, cómo el rendimiento académico está determinado por factores sociodemográficos y socioculturales (tales como el número de libros en el hogar, el acceso a la tecnología, y las políticas y prácticas educativas); con ello, es posible comprender y analizar la influencia de estos factores en el desarrollo del conocimiento y de las destrezas en el entorno familiar y escolar, además de las implicaciones que tienen a la hora de desarrollar una política educativa. Frases como “Menos del 10% de la variación en

el rendimiento, es explicada por el *background* del estudiante, en 5 de los 7 países con el puntaje promedio más alto en Ciencias, o sobre 530 puntos” (OCDE, 2007b), sirven de ejemplo de este tipo de análisis, en los cuales se compara el rendimiento por país, corregido por las covariables.

En otras palabras, se compara el efecto que tienen los diferentes factores en la explicación de la variabilidad, al interior de cada país, ya sea individual o por establecimientos educacionales. La herramienta de análisis utilizada para este tipo de comparaciones son los modelos lineales jerárquicos.

Ahora bien, la comparación basada en el efecto de ciertos factores, y la estimación de la distribución de puntajes, no constituyen enfoques excluyentes, sino complementarios, en el sentido de que siempre es posible describir la distribución de puntajes predichos, una vez que se ha sido corregido por ciertos factores de interés.

En otro ámbito, cabe señalar que, un aspecto muy importante de considerar en grandes estudios de evaluación educacional, como PISA, es el carácter observacional², de los mismos. Cuando los gobiernos y las agencias educacionales solicitan recomendaciones de políticas educativas, estas recomendaciones se realizan a partir del comportamiento de los factores en otro país. Por ejemplo, para Chile, se compara el logro académico de los alumnos de un país, donde la variabilidad socioeconómica sea, en promedio, menor a la chilena, y se establece que el logro educacional chileno *mejoraría* si las diferencias económicas disminuyeran. Este tipo de afirmaciones, que se justifican en el hecho de que existe al menos un caso donde esto ocurre, siempre se plantean como una eventualidad posible: “podrían mejorar”, “podrían aumentar”. Este planteamiento condicional de las recomendaciones,

2 Un estudio observacional es una investigación empírica de tratamientos, políticas o exposiciones, y de los efectos causados por ellos, pero difiere de un estudio experimental, en el hecho de que un investigador no puede controlar la asignación de tratamientos a los sujetos (Rosenbaum, 2002).

tiene su origen en el carácter observacional de estos estudios, porque de ellos, solo se puede “suponer” que ocurrirá un evento: nunca se tendrá la posibilidad de probarlo, ya que ello implicaría, asignar a diferentes sujetos (países, en este caso), diferentes niveles socioeconómicos.

Teniendo en cuenta que toda recomendación de política educacional basada en la evaluación PISA, está limitada por el carácter observacional de los datos y por la información de las comparaciones (que utilizan sólo algunas descripciones de la distribución de puntajes), es importante entender el alcance que tienen las comparaciones entre países.

El objetivo de este estudio fue precisamente mostrar los alcances que tienen las comparaciones entre países, y las limitaciones de las mismas, así como introducir una forma visual que va más allá del reporte y comparación de estadísticos que describen la distribución de puntajes. En otras palabras, dado que la comparación que se presenta en los resultados PISA (basada en promedio y variabilidad asociados), es limitada, uno de los objetivos de este estudio consistió en aplicar un modelo de estimación de la distribución de puntajes que complementara el análisis. Asimismo, considerando que las recomendaciones que actualmente emanan de los resultados PISA, tienen un carácter condicional, en este estudio se analizaron las distribuciones de puntajes predichos, cuando se corrige por factores de interés. De esta manera, fue posible entender los alcances que tienen las comparaciones entre países y las limitaciones de las mismas.

El presente artículo está organizado en tres secciones. En la primera, se incluye la metodología utilizada, haciendo una breve descripción de los estimadores no paramétricos de densidad, que son los que en este estudio se propone utilizar, como complemento a los datos de promedio y variabilidad que se entregan en los resultados PISA³. Además, se describen los modelos

lineales jerárquicos y se interpretan en el contexto de la *comparación de países*. En la segunda sección, se presentan los resultados obtenidos después de aplicar la metodología a los puntajes de la prueba de Ciencias PISA 2006. Finalmente, se ofrece una discusión de la metodología y de cómo utilizar los resultados obtenidos para la elaboración de políticas educacionales.

1. Metodología

En este estudio, se hizo una comparación entre los países latinoamericanos que aplicaron PISA 2006, usando estimaciones de las distribuciones de los puntajes predichos en la prueba de Ciencias, que se obtienen después de corregir por factores de interés. La estimación de la distribución de puntajes se realizó utilizando estimadores no paramétricos de la densidad, mientras que los puntajes predichos se obtuvieron mediante el uso de modelos lineales jerárquicos.

La estrategia de análisis que se siguió se puede resumir en los siguientes puntos:

1. Para entender cómo los factores explican los puntajes, se ajustaron varios modelos que progresivamente fueron incluyendo diferentes covariables. Posteriormente, se estimaron las distribuciones de puntajes predichos, considerando el efecto particular del país (efecto aleatorio). Como resultado de lo anterior, se obtuvieron y compararon tantas distribuciones diferentes como modelos ajustados se aplicaron.
2. Para visualizar en qué consiste la explicación de los puntajes por los factores (incluyendo el efecto país), se contrastaron las distribuciones de puntajes predichos, con la distribución de puntajes observados. Esto no es otra cosa que un “análisis visual de residuos”.
3. Para ofrecer una alternativa a la comparación señalada en 1) se comparó a los países mediante

³ Estos estimadores serán una manera visual de describir la distribución de puntajes, sean estos observados o predichos.

los puntajes predichos, considerando un efecto país promediado o “efecto país nulo”. Esta última comparación resultó pertinente, por cuanto el interés reside en predecir puntajes de estudiantes por país como si cada uno de ellos perteneciera al país de efecto país igual a cero. En otras palabras, la comparación se hace bajo condiciones similares, no a nivel de las covariables, sino a nivel de lo que representa a cada país y que no está incorporado en las covariables.

A continuación, se presenta el detalle de la metodología utilizada en la estrategia de análisis.

1.1 Estimadores no paramétricos de la densidad

1.1.1 Histograma

El histograma es el estimador de densidad más ampliamente utilizado. Sirve para representar gráficamente, mediante barras, la distribución de una variable de interés. Es una herramienta sumamente útil ya que permite representar la información de manera resumida y fácil de interpretar.

Para su construcción se requiere determinar un punto de origen, (x_0) , un ancho de banda (h) y, sobre la base de ambos, se construyen intervalos en los cuales el valor del histograma es proporcional a la cantidad de observaciones que yacen en dicho intervalo⁴ (Silverman, 1986).

Si bien el histograma es sumamente útil para la presentación y exploración de los datos, su construcción es sensible a la elección del punto de origen y del ancho de banda. En otras palabras, la forma del histograma puede variar considerablemente, dependiendo de la elección de estos parámetros, y ello distorsiona una adecuada interpretación de la distribución de los datos.

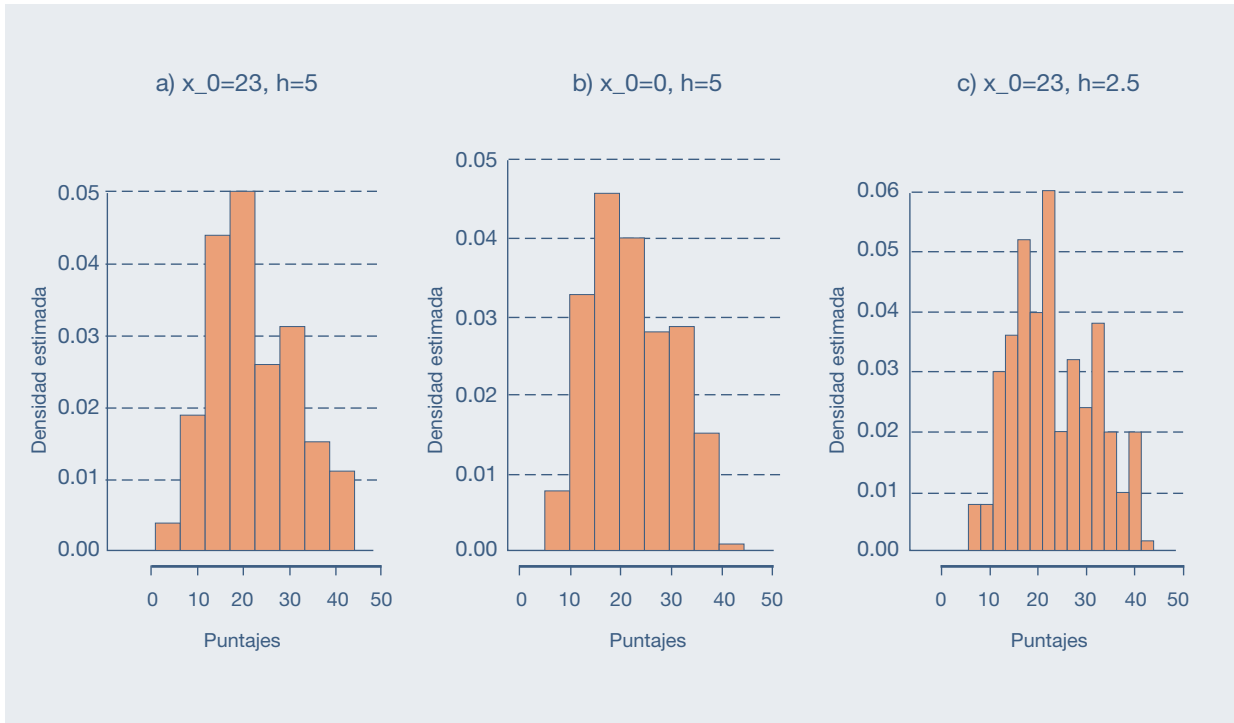
En la Figura 1 se muestran tres histogramas distintos, contruidos a partir de los mismos datos, pero variando los parámetros involucrados en su construcción. En los histogramas (a) y (b), se mantuvo fijo el ancho de banda (h) y se modificó el punto de origen (x_0) . En el histograma (c), se mantuvo el mismo punto de origen que en el histograma (a), y se redujo el ancho de banda a la mitad.

Al observar los tres histogramas, se pueden apreciar algunas diferencias entre los estimadores. El histograma (c) parece mostrar una densidad bimodal que no es clara en los otros dos. Si se comparan los primeros dos histogramas, se observa que el (a) posee una cola derecha con mucha más masa que el (b). En resumen, el análisis en base a histogramas presenta el riesgo de llevar a conclusiones sesgadas por las decisiones tomadas en la construcción de los gráficos. Así, aunque el histograma es una forma práctica de representar la información, es conveniente explorar otras alternativas para la estimación de la densidad.



⁴ Para detalles técnicos de la construcción del histograma, ver Anexo.

Figura 1.
Histogramas de puntajes, que consideran distintos anchos de banda y puntos de origen, para el mismo conjunto de datos



1.1.2 Estimación no-paramétrica de densidades basada en Kernel

De manera similar al histograma, el estimador *Kernel*⁵ considera dos parámetros en su construcción: el de suavizamiento (h) y la función de *Kernel* (K). Para construirlo, se escoge K como una función simétrica de densidad continua (por ejemplo una distribución

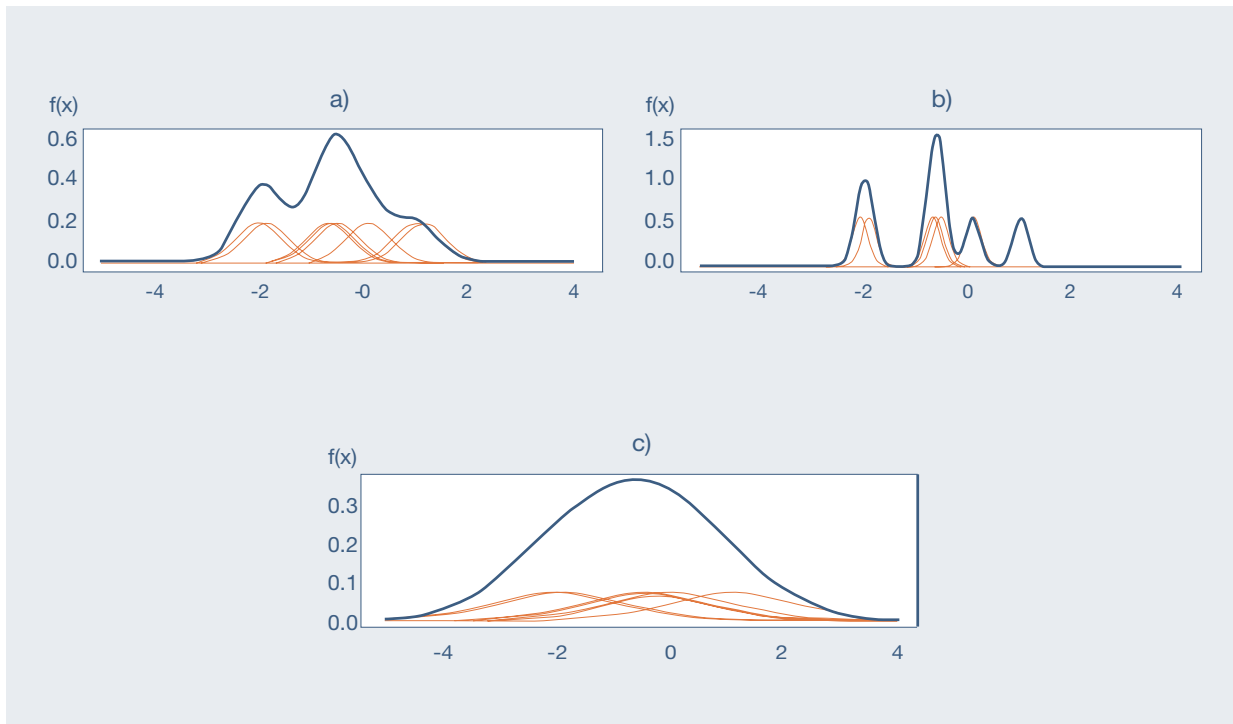
normal de media 0 y varianza 1), y se suman n densidades con forma K , centradas en cada observación, con una varianza igual a h^2 .

En la Figura 2 se muestran ejemplos de estimaciones basadas en *Kernel*, construidas con distintos valores de los parámetros⁶.

⁵ Para detalles técnicos de la construcción del estimador *Kernel*, ver Anexo.

⁶ Estas estimaciones solo tienen un fin ilustrativo, por lo que están basadas en sólo 7 observaciones, aun cuando una estimación con tan pocos datos posiblemente no sea muy realista.

Figura 2.
Estimación de Kernel: (a) con ancho de banda 0.3, (b) con ancho de banda 0.1 y (c) con ancho de banda 0.8



En cada gráfico de la Figura 2, se aprecian siete curvas pequeñas y una curva envolvente que es la suma de las siete funciones pequeñas y que corresponde a la estimación *Kernel*. Los gráficos (b) y (c) se incorporaron para visualizar el efecto del parámetro de suavizado que se observa en (a). En efecto, los valores muy pequeños de h generan una estimación menos suave (gráfico b), mientras que los valores mayores de h (gráfico c), suavizan demasiado la estimación, no dejando observar fenómenos locales.

Como ya se mencionó, para la estimación *Kernel* es necesario escoger el valor de h y la función K . El criterio para esta elección es el mismo que se utiliza para cualquier estimador, a saber, se intenta que la estimación de la densidad esté lo más cerca posible de la verdadera densidad. Una medida de la discrepancia entre estas funciones de densidad en cada punto es el

Error Cuadrático Medio (ECM). Utilizando esta medida, se demuestra que la elección óptima del parámetro de suavizado implica un compromiso entre el sesgo del estimador y su varianza⁷.

El *Kernel* más comúnmente utilizado, es la densidad normal estándar, y por ello es el que se utilizó en este estudio⁸ para la estimación de la distribución de puntajes observados, y puntajes predichos corregidos por covariables. La idea fue comparar dichas distribuciones visualmente y, además, calcular una distancia entre ellas, para decidir si una es significativamente diferente de la otra.

⁷ Detalles técnicos de estos resultados pueden ser consultados en el Anexo.

⁸ Cabe aclarar que si se escoge otro Kernel, los resultados son prácticamente iguales. Para detalles, ver Silverman (1986).

1.2 Comparación de resultados entre países

Uno de los objetivos que se persigue al comparar países, es tener algún criterio cuantitativo que permita decidir cuándo dos países son similares y, consecuentemente, cuándo no lo son. Para este fin, una herramienta muy utilizada en la comparación entre países en la prueba PISA, son los modelos lineales jerárquicos (Mizala y Romaguera, 2000 y 2004), que se utilizan para analizar y comparar cómo determinados factores sociodemográficos influyen en los resultados, y así realizar recomendaciones de política educacional enfocadas sobre los factores modificables.

El rol de los factores explicativos sólo puede especificarse teniendo en cuenta las propiedades estructurales de un modelo lineal jerárquico, en particular, la interpretación estadística del efecto aleatorio que típicamente representa el efecto país⁹. Una vez que este aspecto es explicitado, la comparación entre países se debe realizar, comparando las distribuciones de puntajes predichos (o explicados) por determinados factores. De esta manera, será posible entender el alcance de dichos factores en las predicciones (o explicaciones) de puntajes PISA, que son el sustrato de las comparaciones entre países.

Una vez que se han obtenido puntajes predichos, corregidos por las covariables, siempre es posible obtener una estimación de la distribución de dichos puntajes y posteriormente comparar dichas distribuciones¹⁰.

9 También puede denotar el *efecto-escuela* u otro tipo de efecto. Para decidir cuál efecto incluir en el modelo, es necesario tener presente el significado estadístico del efecto aleatorio.

10 Además de comparar gráficamente las distribuciones de puntajes de los países, se utilizó la distancia de Kolmogorov entre las distribuciones correspondientes, a fin de saber si dicha distancia era estadísticamente diferente de cero. Para ello, fue suficiente realizar un test de Kolmogorov-Smirnov, el cual está basado en la distancia de Kolmogorov.

1.2.1 Estrategia de análisis

En este estudio se utilizaron modelos jerárquicos de efectos aleatorios, para explicar los puntajes obtenidos en la prueba PISA de Ciencias, por medio de factores, tanto a nivel individual, como a nivel del establecimiento educacional.

La aplicación de estos modelos se realizó sobre la información del cuestionario de estudiantes. Debido al tipo de comparación que interesaba, el análisis consideró al conjunto de todos los países (54 en total)¹¹.

El primer paso en la construcción de la base de datos de este trabajo fue fusionar las bases de estudiantes y de establecimientos (OCDE, 2005, p.128). Un segundo paso consistió en eliminar los valores faltantes, para lo cual se siguieron las recomendaciones del manual PISA 2003 (OCDE, 2005, p.176). El ajuste contempló todos los casos que restan, una vez que se han eliminado los datos faltantes. Todos los modelos jerárquicos con efectos aleatorios analizados en este estudio fueron ajustados, utilizando el procedimiento *MIXED* de SAS (Littell y otros, 2006). Finalmente, se realizó una estandarización de los pesos de manera que su suma fuera igual al número de estudiantes en el conjunto de datos (OCDE, 2005, p.174).

Cabe destacar que no se utilizó un solo modelo, sino que se ajustaron varios, incorporando paso a paso determinadas covariables. Los puntajes predichos que se obtuvieron de cada uno de los modelos resultantes fueron posteriormente comparados, utilizando sus respectivas estimaciones de densidad.

Para analizar la importancia de las covariables en el modelo, además de lo que tradicionalmente se reporta (coeficientes de regresión, indicando el aporte de cada factor en la explicación de los puntajes), se utilizó una

11 Sin embargo, se reportaron sólo las comparaciones entre Chile y países latinoamericanos.

forma más informativa, que permite capturar lo que explícitamente explican las covariables. En efecto, por medio de la comparación entre los estimadores de densidad de los puntajes observados y el estimador de los puntajes predichos por las covariables, se pudo visualizar con claridad en qué consiste la explicación de los datos, y cómo es la situación comparada de los distintos países.

1.2.2 ¿Por qué modelos con efectos aleatorios?¹²

Como resulta fundamental mantener el aspecto condicional en las comparaciones, en este estudio se compararon los puntajes de estudiantes pertenecientes a un cierto país, condicional a sus covariables (tanto a nivel individual, como a nivel de establecimiento) y, adicionalmente, condicional al efecto del país.

Tradicionalmente, se asume que los modelos lineales jerárquicos poseen las siguientes características:

1. Si datos individuales (en este caso, los puntajes individuales en PISA 2006) están anidados en grupos (en este caso, países), entonces se debe utilizar este tipo de modelos.
2. El efecto aleatorio representa el “efecto grupo” (en este caso, el efecto país) y al ser especificado en forma aleatoria, se asume una “súper población” de la cual se extraen al azar los grupos que se tiene en la base de datos bajo estudio. En nuestro ejemplo, esto correspondería a afirmar que los países de la base de datos de PISA, han sido seleccionados al azar de entre los países del mundo.

Cabe entonces hacerse la siguiente pregunta: ¿es razonable afirmar que los países que participan en la medición de PISA han sido seleccionados al azar del conjunto de los países del mundo? Evidentemente que

no; de hecho, hay decisiones políticas y económicas que deben satisfacerse para formar parte de la medición de PISA. Por lo tanto, si se mantiene la interpretación tradicional de los modelos lineales jerárquicos, mencionada en los párrafos anteriores, esta respuesta negativa implicaría inevitablemente que no se podrían ajustar modelos de multinivel, donde el efecto aleatorio representa el efecto país.

Otra situación en la cual la interpretación tradicional de los modelos lineales jerárquicos no se satisface, es en el análisis de datos censales. Por ejemplo, un análisis de valor agregado (que se operacionaliza por medio de un modelo lineal jerárquico o de multinivel) a nivel de todos los establecimientos educacionales de Chile, no tendría sentido, pues no existe una “súper población” de colegios chilenos.

Considerando las situaciones anteriores, cabe preguntarse si es posible proponer una interpretación de los modelos lineales jerárquicos que, por una parte, sea compatible con la estructura estadística de los mismos y, por otra, provea de criterios mínimos que deben considerarse a la hora de decidir si una determinada base de datos debe analizarse con este tipo de modelos.

En el marco de este estudio, se ha considerado que sí es posible responder satisfactoriamente a estos dos requerimientos, mientras se hagan explícitos los aspectos estructurales de los modelos lineales jerárquicos¹³.

En el contexto de los modelos lineales jerárquicos, se puede demostrar que, una vez que se ha promediado respecto de la distribución de probabilidad del efecto aleatorio (que representa el “efecto grupo”), el parámetro asociado a dicha distribución representa la covarianza que existe entre las mediciones de cada país y no, como suele interpretarse, la varianza del

12 Para un entendimiento acabado de las ideas que se presentan en esta sección, se recomienda encarecidamente leer el Anexo, en el cual se formalizan matemáticamente los conceptos presentados.

13 Los aspectos estructurales de un modelo son como el “esqueleto” sobre el cual dicho modelo se monta: el significado que el mismo tiene, depende completamente de ese “esqueleto”.

efecto aleatorio. Por lo tanto, si teóricamente se supone que la dependencia observada de las medidas individuales, se debe al grupo al cual pertenecen; entonces debe usarse un modelo con efecto aleatorio¹⁴. Así, por ejemplo, si se asume que los puntajes individuales en una prueba dependen unos de otros, porque los individuos pertenecen a colegios diferentes, entonces se utiliza un modelo con efecto aleatorio, que se especifica a nivel de establecimiento. Si, por el contrario, se supone que esos puntajes dependen unos de otros, porque los alumnos y alumnas han sido tratados por un mismo profesor, entonces se especificará un modelo de efecto aleatorio, donde el efecto aleatorio se define a nivel de profesor.

En este trabajo, se asume que los puntajes individuales obtenidos en PISA 2006, dependen unos de otros, por el hecho de que los individuos pertenecen a países diferentes, por lo que se especificó un modelo de efecto aleatorio, donde el efecto aleatorio se definió a nivel de país.

Es importante mencionar que el aspecto jerárquico de estos modelos, lo define el efecto aleatorio en cuestión, y no la estructura anidada de los datos. Más aún, es posible especificar modelos jerárquicos con jerarquías más complejas, como por ejemplo un efecto aleatorio a nivel de establecimiento y de profesor; o de establecimiento y de país. Estas estructuras complejas pueden, ciertamente, ser especificadas sin atender a consideraciones teóricas, pero un buen uso de estos modelos, necesita del entendimiento de las consideraciones teóricas, para luego especificar un modelo en particular.

Este camino de modelización, implica muchas preguntas, cuando se usan modelos complejos. Por ejemplo, si teóricamente suponemos que la dependencia entre los puntajes individuales depende de la pertenencia a diferentes establecimientos y a países específicos,

entonces usaremos un modelo jerárquico con dos efectos aleatorios. Lo que debe ser decidido teóricamente (es decir, con la intervención de expertos educacionales, por ejemplo) es el tipo de correlaciones que se pueden especificar entre los efectos aleatorios. Si este aspecto no se piensa cuidadosamente, se obtendrá un modelo jerárquico, estadísticamente ambiguo, es decir, que puede explicar situaciones observacionales completamente diferentes¹⁵.

Puede demostrarse también que el efecto aleatorio (que representa el “efecto grupo”, en este caso el efecto país) satisface una propiedad estructural (que en adelante será llamada *relación estructural*) que le da una interpretación sustantiva o que, al menos, puede traducirse en términos sustantivos o teóricos en el contexto de las aplicaciones: el efecto aleatorio de un país es igual a la diferencia entre (a) el puntaje predicho (o puntaje explicado) de los individuos por sus respectivas covariables y el efecto aleatorio de ese país, y (b) el puntaje predicho de esos mismos individuos con las mismas covariables como si estuvieran en un país promedio o de efecto aleatorio nulo (pues se supone, en la especificación del modelo, que la media del efecto aleatorio es nula).

En términos operacionales, el *país promedio* depende de la población de países que se considera en el análisis. Esto implica que la comparación de eficiencia entre países, depende de la población de países utilizada, por lo que la medida de eficiencia es relativa. En consecuencia, un país puede ser eficiente, en relación a un conjunto de países determinado (por ejemplo, países latinoamericanos), pero puede resultar ineficiente, en relación a otro conjunto de países (por ejemplo, países con índices de alfabetización lectora del 95%). Por ello, es crucial especificar en el análisis, el conjunto de países con el que se hará la comparación, pues las conclusiones dependerán de las características de dicho grupo.

14 El parámetro que representa esa dependencia es la covarianza que existe entre las mediciones de cada país.

15 Técnicamente, se trata de un problema de identificación.

La relación estructural antes mencionada, permite dar una interpretación adicional del efecto aleatorio: éste resume las características del país que no están representadas por las covariables y que a su vez permiten dar cuenta de los resultados. Por lo tanto, al considerar un país de referencia con efecto aleatorio nulo, la diferencia antes mencionada puede entenderse como la eficiencia del país: cuanto más positiva es dicha diferencia, más eficiente es dicho país, comparado con otro que tiene una diferencia negativa. Este tipo de comparaciones tiene sentido, pues se hace con un referente único.

2. Resultados

En esta sección se muestran los hallazgos obtenidos de la aplicación de las técnicas descritas, en la comparación de las distribuciones de puntajes observados y de puntajes predichos, utilizando covariables. El interés se centra en comparar a Chile con sus pares latinoamericanos que rindieron la prueba de Ciencias PISA 2006 (en adelante “PISA”).

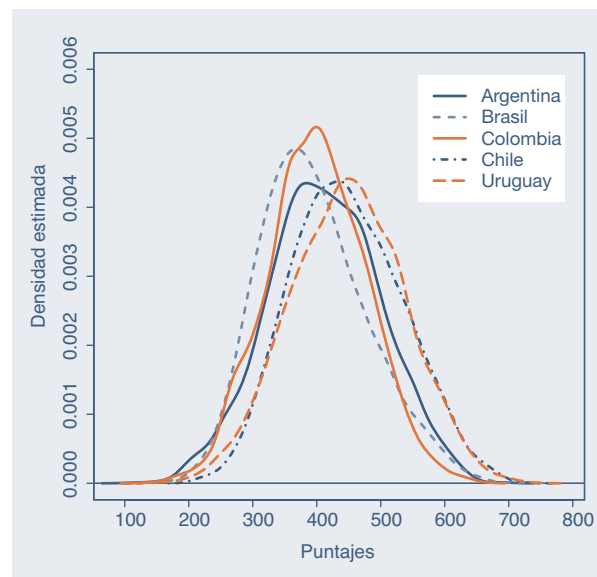
En la primera parte, se presenta un análisis gráfico de comparación de las distribuciones de puntajes observados de Argentina, Brasil, Chile, Colombia y Uruguay. Posteriormente, se ajustan modelos lineales jerárquicos, utilizando distintas covariables para la explicación de los puntajes, y finalmente, se comparan los puntajes predichos entre países y las distribuciones por país, con los modelos ajustados. Para el cálculo de los puntajes predichos se utilizan covariables, tanto a nivel de estudiantes, como de establecimientos.

2.1 Comparación de los puntajes observados

En la Figura 3, se muestran las distribuciones de puntajes observados de Chile y los países latinoamericanos que participaron en PISA, superpuestas para cada

uno de los países¹⁶. En este gráfico se puede apreciar que Brasil se ubica por debajo de los demás países, mientras que Uruguay y Chile están más a la derecha de la escala de puntajes. Nótese que para estos dos últimos países, las diferencias de puntaje no son uniformes a lo largo de la escala. Por ejemplo, entre los 200 y 300 puntos, Chile obtiene mejores resultados que Uruguay, sin embargo, en la zona media, es Uruguay quien obtiene mejores resultados. En la zona alta de puntajes (550 a 750), ambos países obtienen similares resultados. Este tipo de conclusiones no son posibles cuando sólo se reporta el promedio por país, como es habitual en esta clase de estudios.

Figura 3.
Distribuciones de puntajes observados en los países latinoamericanos participantes en PISA



¹⁶ El puntaje de un estudiante fue considerado como el promedio de los cinco valores plausibles correspondientes a dicho estudiante.

Ahora bien, para evaluar estadísticamente las diferencias entre las distribuciones de puntajes obtenidos por Chile y el resto de los países latinoamericanos, utilizamos el test de *Kolmogorov-Smirnov*. En la Tabla 1 se muestra el *valor-p* asociado a cada uno de los tests. De acuerdo con los resultados de la tabla, se puede concluir que no existirían diferencias significativas entre Chile y Uruguay, pero sí entre Chile y los otros países latinoamericanos participantes¹⁷. Se puede entonces afirmar que el rendimiento en Ciencias es mejor en Chile y Uruguay que en los restantes países de la comparación. Ello, por cuanto sus distribuciones de puntajes están a la derecha de las distribuciones de Argentina, Brasil y Colombia.

17 En el Informe Internacional *PISA 2006* (OCDE, 2007b, pp.54 y siguientes), se reportan las medias en Ciencias de cada país. En particular, en la Tabla 2.11b de dicho informe, se reportan las comparaciones entre las medias de los diferentes países. Para Chile se reporta un puntaje promedio de 438 puntos con una desviación estándar igual a 4.3, mientras que para Uruguay se reporta una media de 428 puntos con una desviación estándar de 2.7. La diferencia entre Uruguay y Chile es, según este informe, estadísticamente significativa. Esto parece contradecir los resultados que reportamos en este trabajo. Sin embargo, es importante recordar la advertencia que el mismo informe internacional hace acerca de las comparaciones múltiples. En la nota al pie número 13 (que aparece en la página 118 del mismo informe) se enfatiza que las comparaciones múltiples han sido reportadas para comparar el logro de un país respecto de *todos los restantes países*. La metodología estadística de comparación utilizada en PISA 2006, a diferencia de años anteriores, no estuvo basada en el método de Bonferroni. Sin este método, la probabilidad de afirmar equivocadamente que las diferencias son significativas aumenta con el número de países que se comparan. De hecho, si se compara Bulgaria con Chile, Bulgaria tiene un promedio de 434 puntos (con una desviación estándar de 6.1), es decir, un promedio mayor que el de Uruguay, pero que, según el informe internacional, no es significativamente diferente de Chile. La comparación entre Chile y Uruguay está basada en comparaciones de distribuciones de puntajes, y no sólo en puntajes promedio (junto a sus desviaciones estándares); aunque las distintas conclusiones pueden explicarse técnicamente, esta divergencia es una buena ilustración del mensaje central de este trabajo, a saber: basar las comparaciones en distribuciones de puntajes. A pesar de lo dicho, dejamos para discusiones futuras el ahondar en las aparentes contradicciones.

Tabla 1.
Valor-p asociado al test de Kolmogorov-Smirnov, utilizado en la comparación de distribución de puntajes observados

Países comparados	<i>D</i>	Valor- <i>p</i>
Chile vs. Uruguay	0.021	0.46
Chile vs. Argentina	0.140	0.00
Chile vs. Colombia	0.219	0.00
Chile vs. Brasil	0.210	0.00

2.2 Comparación de puntajes predichos: variables utilizadas y modelos ajustados

Antes de mostrar los resultados de esta comparación, cabe aclarar que todos los modelos que se ajustan incorporan un intercepto, denotado por *Int*. Además, se utilizan las siguientes covariables¹⁸:

- **Género:** sexo del estudiante (donde, hombre = 0 y mujer = 1)
- **CantLib:** cantidad de libros en el hogar (donde, 1 = 0-10 libros; 2 = 11-25 libros; 3 = 26-100 libros; 4 = 101-200 libros; 5 = 201-500 libros, y 6 = más de 500 libros)
- **ImpCie:** importancia que se le da a las ciencias (donde, 1 = muy importante; 2 = importante; 3 = de poca importancia, y 4 = sin importancia)
- **Nse:** un proxy del nivel socioeconómico
- **Ddcia:** tipo de establecimiento (donde, privado = 1 y público = 0)

Estas covariables fueron elegidas por su relevancia en este tipo de estudios y, aunque se pudo haber incorporado otras, se trató de utilizar aquellas que aseguraran datos en todos los países, con el objeto de no reducir demasiado el número de casos en la muestra, por datos faltantes¹⁹.

18 Los nombres originales de las variables, tal como aparecen en las bases de datos, fueron modificados por conveniencia de exposición.

19 Por ejemplo, si se hubieran incorporado variables del tipo "¿ha usado alguna vez un computador?" o "¿con qué frecuencia utiliza el computador?", la base de datos habría perdido 50% de los casos.

En la Tabla 2 se muestra el total de modelos que se ajustaron para los análisis de comparación, y en ella se puede observar que las covariables son progresivamente incorporadas en los modelos.

Tabla 2.
Modelos ajustados

Modelo 1	Int + u_j
Modelo 2	Int + Sexo + u_j
Modelo 3	Int + Sexo + CantLib + u_j
Modelo 4	Int + Sexo + CantLib + ImpCie + u_j
Modelo 5	Int + Sexo + CantLib + ImpCie + Nse + u_j
Modelo 6	Int + Sexo + CantLib + ImpCie + Nse + Ddcia + u_j

En la Tabla 3 se muestran estadísticos de bondad de ajuste de los 6 modelos considerados. Según ambos criterios *AIC* y *BIC*, el modelo 6 es preferible a los otros modelos²⁰.

Tabla 3.
Índices de ajuste para modelos considerados

	Modelo1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
-2 LogVer	3895531.11	3895337.54	3848934.74	3840488.76	3824543.21	3823996.12
AIC	3895537.11	3895345.54	3848952.74	3840512.76	3824569.21	3824024.12
BIC	3895543.08	3895353.50	3848970.64	3840536.62	3824595.07	3824051.97

2.3 Evolución de densidades de puntajes predichos por país

A continuación, para entender el rol de las covariables en la explicación del puntaje predicho, se presentan para cada país, las distribuciones de dichos puntajes. En la Figura 4 se muestran las distribuciones de puntajes predichos en cada uno de los países, considerando los modelos²¹ 3, 4, 5 y 6. Además, se muestran las distribuciones de puntajes observados.



De la Figura 4, en primer término, se puede apreciar que en Chile, los factores sólo explican los puntajes que están entre los 350 y 550 puntos, aproximadamente. Los puntajes altos (mayores a 550 puntos) y bajos (menores a 350 puntos), no son explicados por factores socioeconómicos. Por lo anterior, cualquier recomendación que se haga sobre la base de Chile, debe explicitar que responde a explicar puntajes en la escala media de habilidades de Chile.

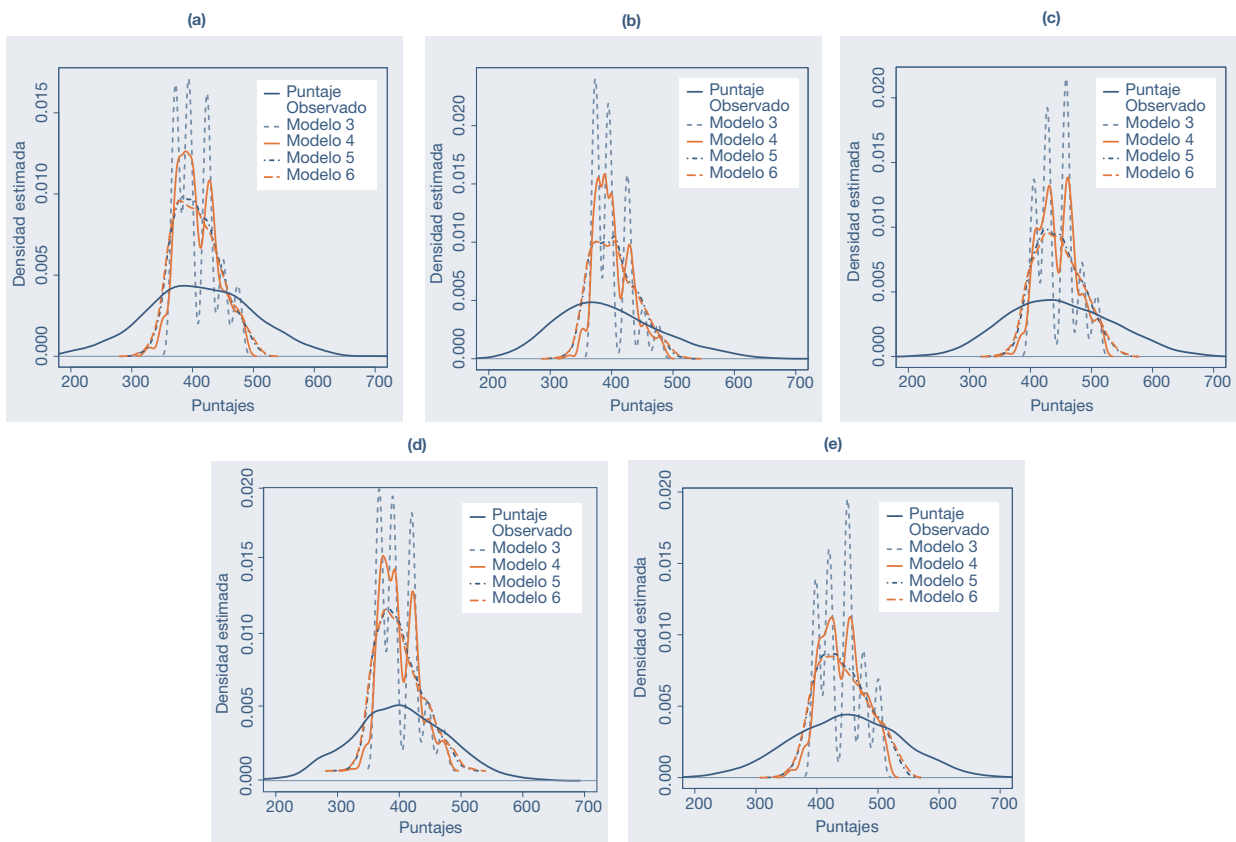
Ahora bien, en cuanto a la comparación con el resto de los países, en la Figura 4 se puede observar que:

²⁰ Valores pequeños son preferibles.

²¹ Las figuras correspondientes a los modelos 1 y 2, no se muestran por insuficiencia de espacio.

- El caso de Uruguay es similar al de Chile, mientras que, por ejemplo, en Brasil, el tramo de puntajes explicados por los factores es menor (entre los 300 y 500 puntos).
- En Brasil, se aprecia que los factores explican los puntajes de estudiantes con puntajes medios y bajos, pero no los más bajos. Es decir, no hay factores que den cuenta del comportamiento de los estudiantes con puntajes menores a 300 puntos.
- En Argentina, Colombia, Chile y Uruguay, las distribuciones de puntajes predichos son razonablemente simétricas, mientras que la distribución de Brasil es asimétrica a la izquierda, demostrando que en este país, los puntajes se concentran en la parte baja de la escala.
- En todos los países, los modelos 5 y 6 producen prácticamente la misma distribución. Esto significa que para estos países, tanto el proxy del nivel socioeconómico como el tipo de colegio, entregan prácticamente la misma información.

Figura 4. Evolución de las densidades de puntajes predichos por los factores, incluyendo el efecto país: (a) Argentina, (b) Brasil, (c) Chile, (d) Colombia y (e) Uruguay



Finalmente, cabe destacar que, paulatinamente los puntajes predichos explican la variabilidad de los puntajes. En efecto, como se sabe, a medida que se incorporan más factores, la explicación de los mismos “mejora”. Esta mejora suele caracterizarse por medio de las varianzas al interior de cada país (intra) y entre los países (inter). Como se observa en la Tabla 4, a medida que se incorporan más factores, la varianza intra disminuye, mientras que la varianza inter aumenta. Este aumento parece deberse al hecho de que sólo se están incorporando covariables a nivel individual²².

Por medio de los estimadores no paramétricos de la densidad, no sólo se entiende en qué consiste la mejora, sino además qué parte de la escala de habilidad observada se explica por dichos factores.

Tabla 4.
Varianzas inter e intra, para modelos 3, 4, 5, y 6

Varianza	Modelo 3	Modelo 4	Modelo 5	Modelo 6
Inter	1928.03	2193.74	2215.76	2173.64
Intra	6380.10	6215.36	5916.07	5906.08

2.4 Comparación de los puntajes predichos por país

Las figuras 5 y 6 muestran cómo difieren las explicaciones proporcionadas por los factores (es decir, los puntajes predichos), cuando se usan los modelos 3 y 6 en todos los países en comparación. En efecto, como se puede apreciar, pese a la incorporación de covariables, las diferencias entre países persisten. Tomando como referencia el modelo 6, es claro que Brasil está por debajo del resto de los países, mientras que Chile y Uruguay poseen distribuciones más a la derecha en la escala de puntajes.

22 Se ajustaron otros modelos con más covariables y las varianzas inter disminuyeron pero, como ya se mencionó, la base de datos disminuyó fuertemente debido a la alta cantidad de no respuesta.

Figura 5.
Comparación de los puntajes predichos por los factores, incluyendo el efecto país para el modelo 3

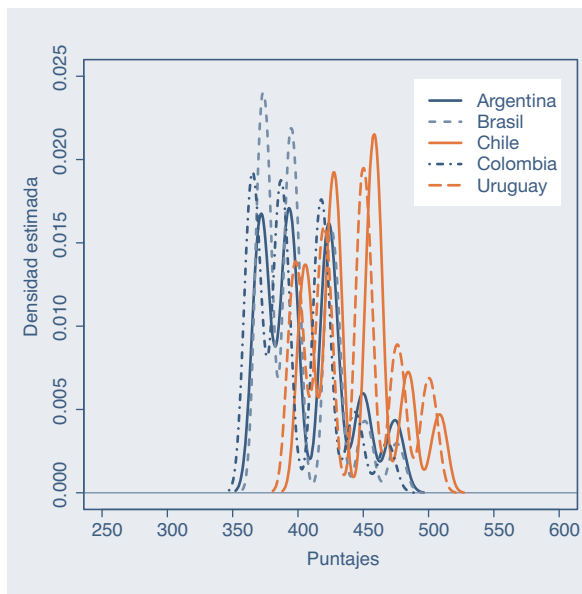
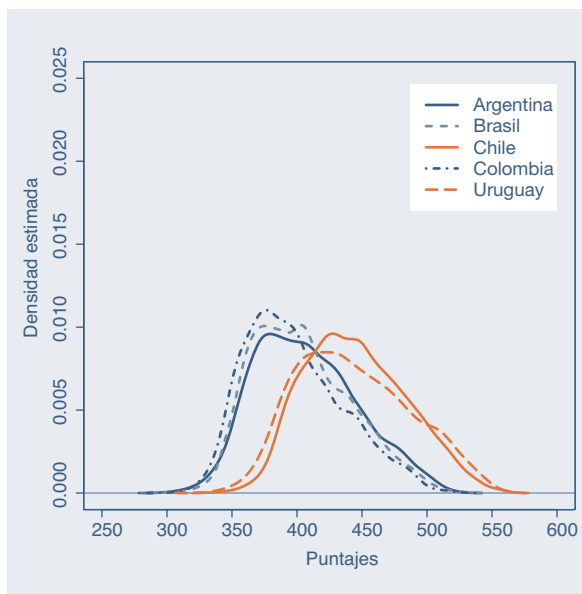


Figura 6.
Comparación de los puntajes predichos por los factores, incluyendo el efecto país para el modelo 6



La Tabla 5 muestra los *valores-p* asociados al test de KS que compara a Chile con el resto de los países latinoamericanos. Nuevamente no existiría diferencia estadísticamente significativa entre Chile y Uruguay, lo cual se refleja de forma clara en la Figura 6.

Tabla 5.
Valor-p asociado al test de Kolmogorov-Smirnov, utilizado en la comparación de distribución de puntajes predichos, utilizando el modelo 6

Países comparados		<i>Valor-p</i>
Chile vs. Uruguay	0.034	0.34
Chile vs. Argentina	0.387	0.00
Chile vs. Colombia	0.353	0.00
Chile vs. Brasil	0.266	0.00

2.5 Comparación de puntajes predichos por país, bajo iguales condiciones

Como se mencionó anteriormente, es posible predecir los puntajes por país bajo condiciones similares, esto es, como si los estudiantes perteneciesen a un país promedio.

Si se compara la Figura 7 con la Figura 4, se aprecia que los puntajes predichos por los factores con efecto país nulo, están más a la derecha que los puntajes predichos por los factores, incluyendo el efecto país. La diferencia que se genera entre estos puntajes sirve para comparar eficiencias por país.

Las figuras 8 y 9 muestran las densidades de estos puntajes predichos por las covariables incluidas en los modelos 3 y 6. De las figuras se puede apreciar que todas las distribuciones de los puntajes predichos se acercan y tienden a explicar un mismo rango de puntajes. Este no era el caso en los resultados reportados en la sección anterior (ver figuras 5 y 6).

Así, por ejemplo, la distribución de puntajes predichos de los estudiantes brasileños, quienes mostraban puntajes en la parte baja de la escala, se acerca mucho a las distribuciones de puntajes de estudiantes chilenos y uruguayos, cuando se establece como si todos estos estudiantes pertenecieran a un país promedio común. Igualmente, si se toma como referencia el modelo 6 (Figura 9), se puede concluir que en general, los estudiantes brasileños con bajo rendimiento (es decir, que están en la izquierda de la escala de puntajes), tienen más bajo rendimiento cuando se los compara con los estudiantes chilenos y uruguayos, siempre bajo el escenario de que pertenezcan a un país promedio común.

Este fenómeno ayuda a interpretar fácilmente la estimación de los efectos aleatorios por país reportados en la Tabla 6. En esta tabla, además de los 5 países antes considerados en la comparación, se agrega el resto de los países americanos participantes en PISA 2006 (Canadá, México y Estados Unidos), además de los países cuyo efecto se encuentra más cercano a 0 (en negrita). A partir de esta tabla, cabe destacar que las predicciones/estimaciones de los efectos aleatorios son bastante precisas (el error estándar es al menos 9 veces más pequeño que la estimación correspondiente); salvo para el caso de los países cuyo efecto es cercano a 0. La no significación del efecto de estos países, parece enfatizar el hecho de que el país de efecto nulo o promedio es sólo una referencia que tiene sentido en la especificación del modelo jerárquico ²³. En efecto, se observa que para los efectos aleatorios predichos, se tiene:

$$\hat{U}_{\text{Colombia}} < \hat{U}_{\text{Argentina}} < \hat{U}_{\text{Brasil}} < \hat{U}_{\text{México}} < \hat{U}_{\text{Uruguay}} < \hat{U}_{\text{Chile}}$$

²³ Técnicamente, se trata de un efecto aleatorio no observable.

Figura 7.
Evolución de las densidades de puntajes predichos por los factores, en un país de efecto nulo:
(a) Argentina, (b) Brasil, (c) Chile, (d) Colombia y (e) Uruguay

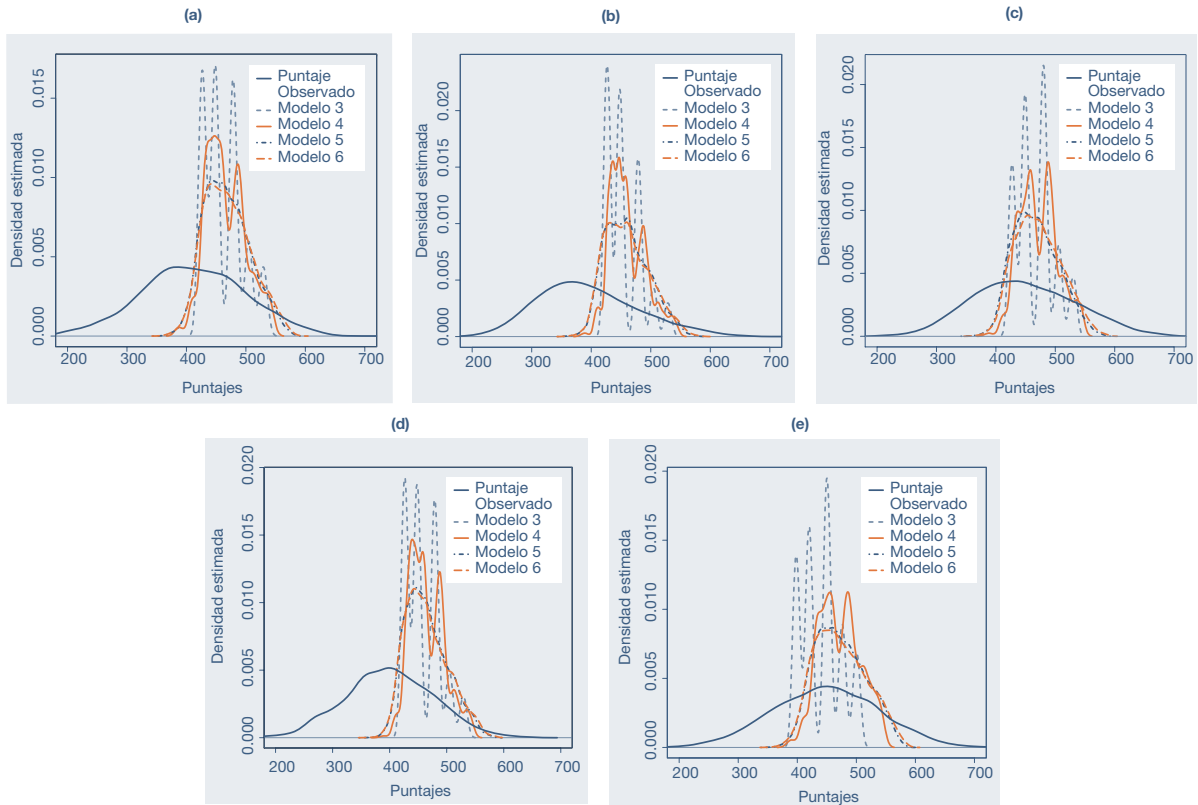


Figura 8.
Comparación de los puntajes predichos por los factores, en un país de efecto nulo para el modelo 3

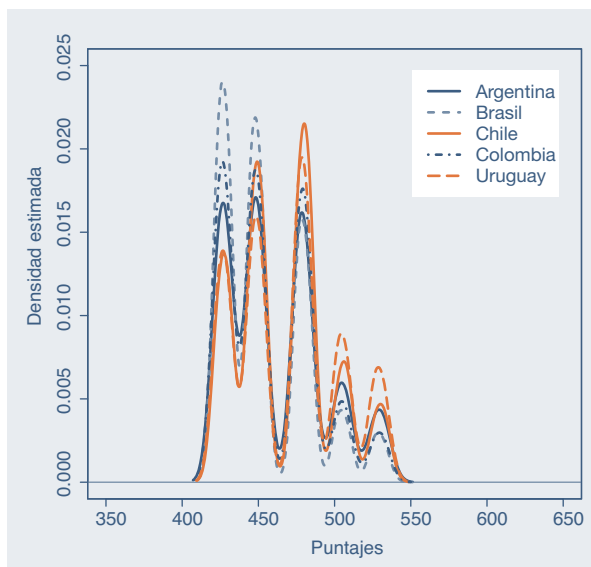


Figura 9.
Comparación de los puntajes predichos por los factores, en un país de efecto nulo para el modelo 6

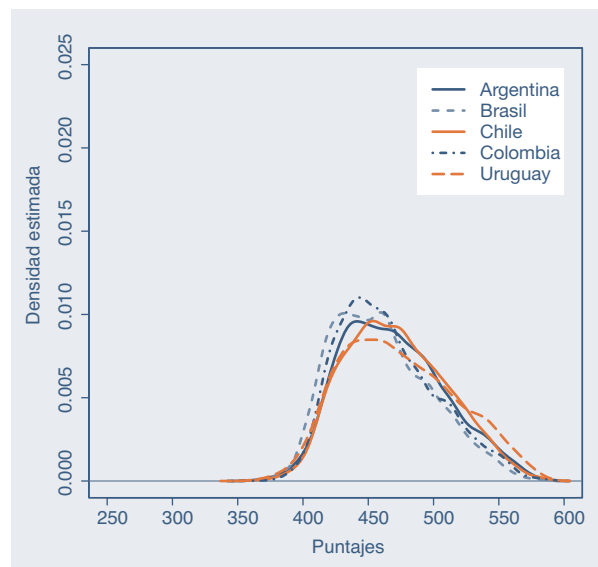


Tabla 6.
Predicción del efecto de los países americanos que participaron en PISA 2006 y país con efecto cercano a 0

Modelo	País	$\hat{\beta}_j$	$SE(\hat{\beta}_j)$	G.L.	Valor- t	Valor- p
	Argentina	-54.95	6.12	323109.00	-8.98	0.00
	Brasil	-53.11	6.04	323109.00	-8.79	0.00
	Canadá	49.50	6.00	323109.00	8.24	0.00
	Chile	-21.69	6.09	323109.00	-3.56	0.00
Modelo 3	Portugal	0.00	6.11	323109.00	0.00	1.00
	Colombia	-61.17	6.11	323109.00	-10.01	0.00
	México	-38.95	6.00	323109.00	-6.49	0.00
	Estados Unidos	13.81	6.08	323109.00	2.27	0.02
	Uruguay	-28.48	6.11	323109.00	-4.66	0.00
	Argentina	-58.57	6.50	323106.00	-9.00	0.00
	Brasil	-58.03	6.44	323106.00	-9.02	0.00
	Canadá	47.26	6.40	323106.00	7.38	0.00
	Chile	-26.90	6.48	323106.00	-4.15	0.00
Modelo 4	Luxemburgo	1.22	6.49	323106.00	0.19	0.85
	Colombia	-67.96	6.50	323106.00	-10.46	0.00
	México	-43.80	6.39	323106.00	-6.85	0.00
	Estados Unidos	11.22	6.47	323106.00	1.73	0.08
	Uruguay	-30.93	6.50	323106.00	-4.76	0.00
	Argentina	-60.18	6.53	323105.00	-9.22	0.00
	Brasil	-56.41	6.47	323105.00	-8.72	0.00
	Canadá	43.01	6.43	323105.00	6.69	0.00
	Chile	-22.16	6.50	323105.00	-3.41	0.00
Modelo 5	Luxemburgo	3.88	6.51	323105.00	0.60	0.55
	Colombia	-66.57	6.52	323105.00	-10.21	0.00
	México	-42.56	6.43	323105.00	-6.62	0.00
	Estados Unidos	6.57	6.50	323105.00	1.01	0.31
	Uruguay	-30.08	6.52	323105.00	-4.61	0.00
	Argentina	-62.19	6.47	323104.00	-9.61	0.00
	Brasil	-56.36	6.40	323104.00	-8.80	0.00
	Canadá	44.31	6.37	323104.00	6.95	0.00
	Chile	-26.33	6.45	323104.00	-4.08	0.00
Modelo 6	Islandia	-2.21	6.48	323104.00	-0.34	0.73
	Colombia	-67.01	6.46	323104.00	-10.37	0.00
	México	-42.61	6.36	323104.00	-6.70	0.00
	Estados Unidos	7.72	6.44	323104.00	1.20	0.23
	Uruguay	-30.09	6.46	323104.00	-4.66	0.00



Las desigualdades antes observadas, implican que Colombia obtendría puntajes más bajos que el resto de los países en comparación, mientras que Chile obtendría los mejores resultados. Una pregunta natural es entonces qué tan bajos son los resultados de Colombia y qué tan altos son los de Chile, respecto del resto de los países. Teniendo en cuenta la interpretación estructural del efecto país, es posible responder a estas preguntas. En efecto, los puntajes predichos de los colombianos, respecto de Colombia son mucho menores que los puntajes predichos de los mismos estudiantes, si estuvieran en el país con efecto nulo. Para Uruguay y Chile no ocurre de la misma manera. Por tanto, se puede concluir que Colombia es muy ineficiente al comparar su logro predicho con el que obtendría al estar en las condiciones de un país con efecto nulo (condiciones que no están representadas por las covariables, sino por características del país no incluidas en las covariables), mientras que Chile y

Uruguay son más eficientes, pues sus logros se parecen al que obtienen, al estar sus estudiantes en un país con efecto nulo. Conclusiones similares se pueden obtener considerando, por ejemplo, México, Uruguay y Chile o cualquier otra combinación de países.

3. Conclusiones

Uno de los objetivos que persigue la comparación de rendimiento entre países, es dar la posibilidad a un país particular, de aprender de otro con mejor rendimiento y, eventualmente, también, de no repetir o consagrar errores de países con peores rendimientos. Dado lo anterior, resulta vital que las comparaciones se realicen bajo iguales condiciones, pues de lo contrario, no es posible sugerir acciones educacionales en el marco de una política pública. En caso de estar en iguales condiciones, no hay que olvidar que las políticas que se desprendan de los análisis sólo son sugerencias, dado el carácter observacional de la información disponible; esto es, si para un país determinado, un factor específico tiene un efecto importante sobre el puntaje PISA, esto no debe interpretarse en términos causales diciendo que "si dicho factor aumenta, entonces el puntaje aumentará". Sólo podemos constatar relaciones entre factores y rendimiento, y monitorear dicha relación en aplicaciones sucesivas de la misma prueba.

En este estudio se logra (mediante el uso de modelos lineales jerárquicos), establecer comparaciones bajo iguales condiciones. El análisis se enfoca principalmente en la comparación entre Chile y los otros países latinoamericanos, permitiendo obtener las siguientes conclusiones, a considerar como insumos de eventuales políticas públicas:

- Existen países con rendimientos en PISA muy superiores a los chilenos.

- El rendimiento de Chile en PISA, es superior al de la gran mayoría de los países latinoamericanos.
- Colombia es ineficiente, comparado con Chile y Uruguay.
- Chile y Uruguay son igualmente eficientes.
- El rendimiento de estudiantes pertenecientes a distintos países, en condiciones similares, es muy parecido.

Considerando esta información, podría decirse, por ejemplo, que Chile constituiría un “ejemplo a imitar” por Colombia, si este último quisiera tener una eficiencia como la chilena. Pero ¿qué debería hacer Colombia para lograr una eficiencia como la de Chile? Por de pronto, no podría considerar las covariables que ayudan a explicar los diferentes rendimientos, pues la eficiencia se debe a lo que no está contenido en dichas covariables. Sólo si se encontrara un país tal, que su efecto país fuera muy cercano a cero, entonces se podría caracterizar la eficiencia en términos de los valores tomados por dichas covariables, pero los datos disponibles no permiten hacerlo.

Es importante destacar aquí que estas conclusiones se basan en comparaciones entre distribuciones de puntajes predichos, y no sólo entre distribuciones de puntajes observados. Las comparaciones se basan precisamente en los puntajes predichos, porque de esa manera se puede incluir en ellas los factores que ayudan a explicar la variabilidad de los rendimientos en PISA (es decir, la distribución de puntajes observados). Si sólo se realizaran comparaciones utilizando rendimientos observados, entonces sólo restaría describir las diferencias entre las distribuciones de dichos puntajes. Pero dado que el foco es proporcionar información para sugerir acciones de política pública, es fundamental usar las predicciones, es decir, utilizar lo que efectivamente se logra explicar de la variabilidad observada.

Ahora bien, en términos del aporte del método utilizado, se pueden concluir dos aspectos importantes, a ser considerados cuando se elaboren recomendaciones educativas, a partir de resultados comparados:

1. La eficiencia que presenta un país respecto de otro, no está explicada por las covariables utilizadas, sino por otras consideraciones. Por lo tanto, el desafío institucional es detectar un país muy eficiente (por ejemplo, Finlandia, que efectivamente tiene un efecto muy por encima del valor de referencia 0) y observar, no lo que las covariables informan, sino lo que se desprende de otros aspectos no considerados en estas covariables²⁴.
2. Los puntajes predichos por los factores, incluyendo el efecto país, explican sólo la parte “media” de la escala de habilidades observadas. Se requiere, por tanto, detectar covariables que permitan explicar los puntajes superiores e inferiores. Por ende, una acción de política pública es observar los medios socioculturales de los altos y bajos puntajes, a fin de generar indicadores que puedan ser aplicados sistemáticamente.

Los factores que se han utilizado en este estudio, no dan cuenta de toda la variabilidad observada en el rendimiento educacional. Esto puede considerarse como una insuficiencia en la medición de factores asociados al rendimiento, e invita a enriquecer los cuestionarios; pero por otro lado, permite afirmar que la variabilidad en el rendimiento de un país, no se agota en los factores asociados. Esta sola constatación puede ayudar a ampliar los límites de acción de la política pública, aunque no permite señalar qué otras áreas puede abarcar. De allí entonces que sea necesario insistir en la necesidad de enriquecer los indicadores

²⁴ Puesto que se trata de estudios observacionales, será imprescindible consignar dicha información en cuestionarios posteriores para así obtener evidencia empírica de la eventual importancia de dicho factor.

en los cuestionarios y encuestas educacionales, de manera de intentar medir comportamientos individuales aún no explicados. Esta es ciertamente una acción de política pública, pues requiere del acuerdo entre los países para generar comisiones de expertos que propongan nuevos indicadores. Una vez propuestos y validados estos indicadores, se requiere del acuerdo de los países para integrarlos en cuestionarios futuros, y finalmente, se requiere una discusión de los eventuales hallazgos. Se tiene la certeza de que se trata de una empresa de largo aliento, pero al menos invita a ampliar los horizontes de los análisis, mejorar las explicaciones estadísticas y eventualmente obtener mayor información para sugerir políticas públicas.

Referencias

- Basu, D. (1975). "Statistical Information and Likelihood". *Sankhya: The Indian Journal of Statistics*, 37. Calcuta, India: The Indian Statistical Institute.
- Littell, R.C. y otros (2006). *SAS System for Mixed Models*. Cary, North Carolina: SAS Institute Inc.
- Mizala, A. y Romaguera, P. (2000). "School performance and choice: The chilean experience". *The Journal of Human Resources*, 35. Madison, Wisconsin: The University of Wisconsin Press.
- _____ (2004). "School and teacher performance incentives: The latin american experience". *International Journal of Educational Development*, 24. Elsevier.
- OECD (Organización de Cooperación y Desarrollo Económicos) (2005). *PISA 2003 Data Analysis Manual: SAS® Users*. París, Francia.
- _____ (2007a). *PISA 2006: Science Competencies for Tomorrow's World: Analysis*,(1). París, Francia.
- _____ (2007b). *PISA 2006: Science Competencies for Tomorrow's World: Executive Summary*,(1). París, Francia.
- _____ (2007c). *PISA 2006: Data*,(2). París, Francia.
- Rosenbaum, P. R. (2002). *Observational Studies*. Second Edition. Nueva York: Springer.
- Silverman, B. W. (1986). *Density estimation: for statistics and data analysis*. Londres: Chapman and Hall.
- Snijders, T. and Bosker, R. (1999). *Multilevel Analysis: An introduction to basic and advanced multilevel modeling*. Londres: Sage.

Anexo

1. Construcción del histograma

Cada intervalo se construye a partir de los valores x_0 y h , además de un entero m positivo (positivo, cero o negativo) mediante la fórmula:

$$I_m = [x_0 + mh, x_0 + (m + 1)h)$$

Una vez determinados los intervalos, se construye el histograma de la siguiente forma: para $i = 1, \dots, n$, siendo n el número de observaciones disponibles:

$$\hat{f}(x) = \frac{1}{nh} (\text{nro. de } X_i \text{ en el mismo intervalo que } x) = \frac{1}{nh} \sum_{i=1}^n w(X_i)$$

donde $w(X_i) = 1$, si X_i está en el mismo intervalo que x ; y $w(X_i) = 0$ en caso contrario. Nótese que el valor del histograma en cada intervalo es proporcional a la cantidad de observaciones que se incluyen en dicho intervalo. La fracción inicial $\frac{1}{nh}$ es un ajuste necesario para que el área de todas las barras sumen 1 (Silverman, 1986).

2. Construcción del estimador Kernel

El estimador de Kernel se define como:

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x - X_i}{h}\right)$$

donde h es el parámetro de suavizamiento y K es la función de Kernel. El estimador consiste en un promedio de distribuciones normales centradas en cada observación, con varianza h^2 .

3. Detalles técnicos de la elección de los parámetros involucrados en la estimación Kernel

Midiendo el error de estimación

Para expresar más formalmente la idea de minimizar la distancia entre \hat{f} y f , supongamos que se posee una muestra X_1, \dots, X_n de observaciones independientes provenientes de una distribución continua f . La idea central consiste en medir la distancia entre la estimación Kernel \hat{f} y la verdadera densidad f y luego elegir h y K tales que minimicen dicha distancia. Una medida de la discrepancia entre estas funciones en cada punto, es el Error Cuadrático Medio (ECM). Su definición es:

$$ECM_x(\hat{f}) = \{E(\hat{f}(x)) - f(x)\}^2 + Var(\hat{f}(x)) = \text{sesgo}^2(\hat{f}(x)) + \text{varianza}(\hat{f}(x))$$

La ecuación anterior muestra que el error de estimación se puede separar en dos componentes: el sesgo elevado al cuadrado (que mide la distancia entre el valor estimado del Kernel en un punto dado y el valor verdadero de la densidad en ese punto) y la varianza del estimador en el punto. El ECM_x está definido para cada x fijo, de manera que ofrece una medida de error puntual o local. Para tener una noción global del error de estimación, se puede recurrir al Error Cuadrático Medio Integrado (ECMI), que se define como:

$$ECMI = \int ECM_x(\hat{f})dx = \int \{E(\hat{f}(x)) - f(x)\}^2 dx + \int Var(\hat{f}(x))dx$$

El $ECMI$ también se descompone en el sesgo al cuadrado integrado y la varianza integrada. Se busca que $ECMI$ sea lo más pequeño posible. Para este efecto, la elección de K no es demasiado relevante, por lo que nos concentraremos en analizar el efecto de h sobre el error.

Sesgo versus varianza

Para simplificar el análisis consideraremos K una función de densidad simétrica con esperanza cero y varianza σ^2 . En primer lugar, nótese que el sesgo no depende del tamaño muestral sino sólo de h . En efecto, se puede demostrar que:

$$sesgo_h = E(\hat{f}(x)) - f(x) = \frac{1}{h} \left\{ \int K\left(\frac{x-y}{h}\right) f(y) dy \right\} - f(x)$$

Esta conclusión es interesante pues indica que el sesgo no puede ser eliminado, aumentando el tamaño de la muestra.

A continuación se presentan aproximaciones del sesgo y la varianza integrados.

$$\begin{aligned} sesgo_h(x)^2 dx &\approx \frac{1}{4} h^4 \int f''(x)^2 dx \\ V(\hat{f}(x)) &\approx n^{-1} h^{-1} \end{aligned}$$

Estos dos elementos sumados, conforman el $ECMI$.

Mediante estas fórmulas se puede entender el principal problema de la estimación de densidades: si se intenta reducir el sesgo al cuadrado integrado, se debería escoger h muy pequeño, lo que implicaría una varianza integrada grande. En otras palabras, la elección del parámetro de suavizamiento implica un compromiso entre sesgo y varianza, y esto es cierto para cualquier método de estimación de densidades que se quiera utilizar.

Elección del parámetro de suavizamiento óptimo

El valor de h óptimo se obtiene minimizando el *ECMI*. Este procedimiento nos lleva al valor:

$$h_{opt} = k_2^{-\frac{2}{5}} \left\{ \int K(t)^2 dt \right\}^{\frac{2}{5}} \left\{ \int f''(x)^2 dx \right\}^{-\frac{1}{5}} n^{-\frac{1}{5}}$$

Sin embargo, no es posible obtener h_{opt} pues, en la práctica, f es desconocida. Por ello, es necesario explorar formas alternativas para la elección de h . El método que se utilice para elegir el parámetro de suavizamiento, depende mucho del propósito para el cual se está estimando la densidad. Si se trata de una primera aproximación a los datos para plantear posibles modelos o hipótesis, es suficiente probar distintos valores de h y elegir uno en forma arbitraria. Para el lector que esté interesado en una forma automática de elegir h , se puede emplear el enfoque propuesto por Silverman (1986). Este consiste en reemplazar $f(x)$ por la densidad normal con media cero y varianza σ^2 . Si además se usa un Kernel normal, el valor de h_{opt} queda dado por:

$$h_{opt} = 1.06\sigma n^{-\frac{1}{5}}$$

En esta fórmula el único valor desconocido es σ^2 , el cual puede ser estimado a partir de la varianza muestral.

4. Especificación de los modelos lineales jerárquicos utilizados

Denotando por Y_{ij} al puntaje del individuo i perteneciente al país j , y X_{ij} al vector que contiene las covariables individuales e institucionales, los modelos lineales jerárquicos se especifican de la siguiente manera:

$$(Y_{ij} | X_{ij}, u_j) \sim \mathcal{N}(\beta_0 + \beta' X_{ij} + u_j, \sigma^2), \quad j = 1, \dots, J, \quad i = 1, \dots, n_j$$

$$(u_j | X_{ij}) \sim \mathcal{N}(0, \tau^2), \quad j = 1, \dots, J$$

donde el símbolo i significa “condicional a”, y u_j es el efecto del país j . Se supone además que los u_j son mutuamente independientes, condicionalmente a todas las covariables, y los Y_{ij} son mutuamente independientes, condicionalmente a los efectos aleatorios y a todas las covariables. Por ejemplo, si a nivel de individuo tenemos el género como covariable (que denotamos por G_{ij}), su nivel educacional (que denotamos por NEd_{ij}) y el tipo de establecimiento educacional al cual asiste (que denotamos por TC_j), entonces

$$(Y_{ij} | G_{ij}, NEd_{ij}, TC_j, u_j) \sim \mathcal{N}(\beta_0 + \beta_1 G_{ij} + \beta_2 NEd_{ij} + \beta_3 TC_j, \sigma^2)$$

$$(u_j | G_{ij}, NEd_{ij}, TC_j) \sim \mathcal{N}(0, \tau^2)$$

donde $j = 1, \dots, J$ e $i = 1, \dots, n_j$. El primer nivel especifica la parte del puntaje Y_{ij} que puede explicarse con los factores antes mencionados, además del efecto-país u_j ; esto es,

$$E(Y_{ij} | G_{ij}, NE_{ij}, TC_j, u_j) = \beta_0 + \beta_1 G_{ij} + \beta_2 NE_{ij} + \beta_3 TC_j + u_j$$

Aquí, "la parte de Y_{ij} que es explicada por los factores y el efecto país" es lo que técnicamente se llama esperanza condicional, y es denotada por $E(Y_{ij} | G_{ij}, NE_{ij}, TC_j, u_j)$.

La "parte no explicada de Y_{ij} " es lo que técnicamente se llama error de medición, y está definido como:

$$\epsilon_{ij} = Y_{ij} - E(Y_{ij} | G_{ij}, NE_{ij}, TC_j, u_j)$$

Es decir, corresponde a la diferencia entre lo que se quiere explicar (en este caso, el puntaje Y_{ij}) y la parte que se logra explicar con los factores (incluido el efecto país).

Se puede verificar que, si todas las variables aleatorias incluidas en el modelo tienen varianza finita, el error de medición no está correlacionado con la esperanza condicional de Y_{ij} dado los factores (incluyendo el efecto país). Esta correlación nula significa que lo explicado por los factores no comparte información con el error de medición.

Aspectos estructurales de los modelos

De acuerdo con el Principio de Verosimilitud (Basu, 1975), los parámetros de un modelo deben interpretarse respecto del modelo estadístico, que es el que caracteriza las observaciones bajo estudio. Cuando se asume un modelo con efecto aleatorio, el modelo estadístico (que se obtiene después de promediar respecto de la distribución de probabilidad del efecto aleatorio) tiene una estructura determinada que busca explicar las observaciones. En particular, dado que los errores ϵ_{ij} son no correlacionados entre sí y además no correlacionados con U_j , se puede demostrar que:

$$\text{cov}(Y_{ij}, Y_{i'j} | X_{ij}) = \tau^2 \quad i \neq i'$$

Es decir, una vez que se ha promediado respecto de la distribución de probabilidad del efecto aleatorio, el parámetro τ^2 representa la covarianza que existe entre las mediciones de cada país y no, como suele interpretarse, la varianza del efecto aleatorio (Snijders y Bosker, 1999). Evidentemente, τ^2 representa la varianza de u_j en el modelo que genera u_j , pero el foco estadístico es explicar las observaciones por medio de distribuciones de muestreo, por lo que los parámetros de interés deben interpretarse respecto de esta última distribución. Esto lleva a concluir que, usar un modelo con efecto aleatorio significa modelar la correlación que se asume que existe entre las mediciones de un mismo país.

Puede demostrarse que el efecto aleatorio satisface la siguiente propiedad estructural:

$$u_j = E(Y_{ij} | X_{ij}, u_j) - E(Y_{ij} | X_{ij}), \quad \text{para todo } i = 1, \dots, n_j$$

Puesto que $E[E(Y_{ij} | X_{ij}, u_j) | X_{ij}] = E(Y_{ij} | X_{ij})$, entonces $E(Y_{ij} | X_{ij})$ corresponde al “promedio” de los puntajes predichos de los individuos por sus respectivas covariables y el efecto aleatorio del país j . Este “promedio” es respecto de la distribución del efecto aleatorio u_j . Por lo tanto, la relación estructural significa que el efecto aleatorio del país j es igual a la diferencia entre: (a) el puntaje predicho de los individuos por sus respectivas covariables y el efecto aleatorio de ese país, y (b) el puntaje predicho de esos mismos individuos con las mismas covariables, como si estuvieran en un país promedio.

A fin de enfatizar el sentido de esta relación estructural, consideremos dos países, denotados por $j = 1$ y $j = 2$. Entonces:

$$u_1 = E(Y_{i1} | X_{i1}, u_1) - E(Y_{i1} | X_{i1}), \quad u_2 = E(Y_{i2} | X_{i2}, u_2) - E(Y_{i2} | X_{i2})$$

Para el país 1, u_1 corresponde a la diferencia que hay entre: (a) los puntajes predichos de los individuos del país 1 por sus covariables X_{i1} y el efecto u_1 , y (b) los puntajes predichos de esos mismos individuos por sus mismas covariables. Lo mismo para el país 2. Así, esta diferencia respeta las particularidades de cada país, al mismo tiempo de realizar la comparación respecto de una referencia común, a saber, un país promedio caracterizado por un efecto aleatorio 0.

Finalmente, es importante mencionar que las covariables no están correlacionadas con el efecto aleatorio. Así,

$$\text{cov}(u_j, X_{ij}) = \text{cov}[E(u_j | X_{i1}), X_{i1}] = \text{cov}(0, X_{i1}) = 0$$

Esto significa que las covariables incluidas en el modelo **no** comparten información común. Dicho de otra manera, todo lo capturado en los factores o covariables, no está capturado por el efecto aleatorio, y viceversa.

Conclusiones 5



Conclusiones

Esta publicación es producto del esfuerzo por seguir investigando los datos de PISA en la búsqueda de soluciones y formas de mejorar la educación.

Las buenas noticias de PISA 2006 para Chile fueron que los estudiantes chilenos, formados completamente en el currículo reformado, alcanzaron en general mejores resultados que sus pares latinoamericanos, y que en Lectura obtuvieron mejores puntajes que los estudiantes chilenos que rindieron PISA cinco años antes. Sin embargo, los primeros análisis también revelaron dos aspectos que evidencian inequidad: la segregación socioeconómica y académica¹ del sistema educativo y la disparidad en el rendimiento de hombres y mujeres.

En este capítulo se pretende llamar la atención sobre aspectos específicos que se destacan en el conjunto de artículos. No es una síntesis ni pretende agotar las posibilidades de proyecciones que se plantean a partir de los diversos estudios: solo se revisarán algunos temas que parecen ser relevantes y fértiles insumos de política pública.

1 El término segregación se utiliza para indicar el carácter internamente homogéneo de los establecimientos y muy diferenciado de los otros. Esto puede ser producto de la auto-selección o de la selección, por parte del establecimiento, en términos académicos o por características sociales, económicas y culturales de la familia. PISA permite medir la segregación socioeconómica (composición), por una parte, y también cuenta con una variable que indica si el establecimiento realiza selección académica para incorporar a los estudiantes, por otra.

Incidencia de las características socioeconómicas individuales y del establecimiento

En varios artículos de esta publicación se hace referencia a variables relativas al nivel socioeconómico de los estudiantes (individual o del grupo curso). Ello coincide con lo que habitualmente plantea la literatura sobre el tema, donde se reitera la importancia de este factor en el rendimiento de los estudiantes. En los estudios aquí presentados, en general, se utiliza el índice socioeconómico y cultural elaborado por PISA, con la excepción del artículo de Kluttig y otros, donde se usa la clasificación de grupos socioeconómicos construida por SIMCE.

Más que constatar los hallazgos de que a mayor nivel socioeconómico mayor rendimiento, estos trabajos tratan de avanzar, ya sea determinando qué otras variables pueden mejorar el rendimiento, independientemente del entorno sociocultural, o preguntándose qué sucedería si mejoraran las condiciones socioeconómicas a nivel nacional. La escuela no puede suplir las carencias sociales provocadas por el insuficiente desarrollo del país o por las deficiencias en la distribución de los ingresos, pero sí puede (y debe) desarrollar estrategias que incrementen el aprendizaje de todos los estudiantes, sin importar el origen socioeconómico de las familias. La forma de desarrollar dichas estrategias es la pregunta importante a responder.

Como se ha visto, la publicación incluye tres análisis multinivel² (Treviño et al.; Valenzuela et al., y Bellei et al.) que buscan encontrar los factores asociados al

logro de los estudiantes³, uno respecto de cada área de aprendizaje (Ciencias, Matemática y Lectura), los cuales resultan muy coincidentes. Estos tres estudios coinciden en destacar que el nivel socioeconómico del curso o “efecto par” tiene mayor incidencia (controlando otros factores) que el origen socioeconómico individual. Esto significa que al estar juntos los estudiantes con mejores condiciones socioeconómicas y culturales, se potencian mutuamente para un mayor aprendizaje. A la inversa, al estar juntos los que tienen condiciones más desfavorables, también se potencian mutuamente, en desmedro de los aprendizajes. Este es un tema reiterado en las investigaciones basadas en resultados de los estudiantes chilenos y parece claramente recomendable que la política educacional ponga atención en él, en términos de buscar asemejar la composición estudiantil entre los establecimientos. La evidencia indicaría que lo más relevante parece ser la composición de los establecimientos y no el NSE de los alumnos en sí mismo.

Sobre este ámbito, aunque se han hecho esfuerzos en el sentido correcto —como la subvención de educación preferencial, la exigencia de un 15% de estudiantes de escasos recursos en todos los establecimientos subvencionados y la prohibición de selección hasta 6° Básico por razones académicas o socioeconómicas, establecida en la recientemente aprobada Ley General de Educación— no parecen ser suficientes. Por ejemplo, la subvención preferencial está operando solamente en Enseñanza Básica y su objetivo es compensar el mayor costo de educar a estudiantes más vulnerables. La capacidad de esta medida de reducir la segregación del sistema, deberá comprobarse en el tiempo. De hecho, hay quienes sostienen que generará una mayor segregación (o

2 Se refiere a una metodología que pretende determinar cuánto cambia una variable dependiente por cada unidad de variación de la independiente, suponiendo que los restantes factores en juego no se alteran. Las variables pueden corresponder al individuo o a agrupaciones mayores.

3 El libro contiene otro artículo que utiliza la metodología de análisis jerárquico -González y San Martín- pero con otro objetivo. A partir de los residuos de los modelos ajustados, y por ende haber controlado todos los factores dentro del modelo, estiman el “efecto país” que estaría dado por otros aspectos que no son medidos en los cuestionarios. Además, consideran para su análisis todo el rango de distribución de puntajes y no solo los promedios.

“descreme” del estudiantado), que la actual, atrayendo hacia determinados establecimientos a los mejores estudiantes del nivel socioeconómico bajo.

El desafío político es significativo y pasa por un cambio cultural de una sociedad en la que oponerse radicalmente a todo tipo de segregación, resulta excepcional.

En efecto, los países emblemáticos por su alto rendimiento, como Finlandia, Hong-Kong, Canadá, Corea y China Taipei, tienen menor variación entre escuelas que Chile y, más importante, todos tienen menor variación entre escuelas que se explique por el nivel socioeconómico de las escuelas y de los estudiantes⁴. Por otra parte, estos países también tienen mayor variación al interior de las escuelas que las que presenta Chile. De hecho, como destacan Valenzuela y otros, Chile es uno de los países que tiene mayor segregación socioeconómica en el sistema escolar y donde las escuelas más seleccionan por razones académicas.

Sin intención de entrar en el debate, con aquellos que sostienen que si no hubiera segregación habría una tendencia a la mediocridad desincentivando la excelencia, baste con decir que los países con una menor segregación (o con menor variación entre escuelas, que al interior de cada escuela) no solo tienen resultados más equitativos, sino también mejores resultados a nivel de toda la población⁵. En este sentido, como se destaca en el trabajo de Treviño, las escuelas con nivel socioeconómico promedio muy bajo muestran la mayor dispersión de puntajes promedio por escuela y también la mayor proporción de rendimiento bajo

el nivel esperado⁶. De esto se pueden desprender al menos dos conclusiones: (a) todo lo que eleve el nivel de desarrollo económico del país, especialmente de los más desfavorecidos, favorecerá un mejor rendimiento general y (b) si se logra subir el promedio del nivel socioeconómico de las escuelas con mayor pobreza (sea por mayores ingresos de la economía, mejor distribución de ellos o por redistribución de los estudiantes más desfavorecidos entre otras escuelas), necesariamente debería aumentar el rendimiento promedio por escuela.

Relacionado con lo anterior, aunque desde otra perspectiva, cabe mencionar que la distribución de puntajes tiene una forma semejante a la distribución de las escuelas según su nivel socioeconómico promedio. Sin embargo, se requiere más investigación en relación a la posibilidad de distribuir a los alumnos en las escuelas de una manera más homogénea y para determinar qué sucede con los estudiantes de distintos estratos socioeconómicos, en situación de mayor o menor segregación socioeconómica.

En síntesis, a este respecto, cabe señalar que uno de los objetivos claves de la política educativa en los próximos años debería ser reducir la señalada segregación del sistema educativo; ello permitiría no solo tener un sistema más equitativo e integrado, sino también mejorar los promedios nacionales de los resultados de aprendizaje. Así también, para la investigación, es clave la pregunta sobre los factores más determinantes que estarían incidiendo en la segregación en el caso de Chile, para desde allí abordar modos de corregirla.

4 OCDE, 2007 op. Cit., Figuras 4.6 y 4.10

5 Ver: Gerard Ferrer Esteban, José Luis Castel Baldellou y Ferran Ferrer Julià: Las desigualdades del sistema educativo a través del estudio PISA 2003, Revista de Educación, extraordinario 2006, pp. 399-428. Xavier Dumay* & Vincent Dupriez: Does the School Composition Effect Matter? *Some Methodological and Conceptual Considerations*, LES CAHIERS DE RECHERCHE EN EDUCATION ET FORMATION, Université Catholique de Louvain (UCL) N° 60, Juin 2007.

6 Aquí el nivel esperado se entiende como el rendimiento pronosticado por el modelo para los estudiantes (o los establecimientos), de acuerdo a su nivel socioeconómico.

Otros factores que se asocian al rendimiento

Los tres estudios que realizan análisis multinivel no utilizan el mismo modelo y sus variables dependientes corresponden al rendimiento en distintas áreas de aprendizaje. De esta manera, los distintos modelos aplicados por los investigadores muestran que algunos aspectos que se relacionan con el logro en algunas áreas, no lo hacen o lo hacen en menor medida en otras. A continuación se hará referencia a aquellos factores que se destacan por coincidentes, así como a aquellos que levantan dudas.

Dependencia político-administrativa de los establecimientos

La dependencia político-administrativa de los establecimientos, muestra diferencias significativas a favor de los establecimientos públicos en el rendimiento de Matemática, pero no en el de Lectura (aunque para ambas áreas se usó el mismo modelo). En Ciencias, en cambio, los establecimientos particulares pagados y particulares subvencionados muestran un mayor rendimiento que los municipales, después de haber controlado otros once factores a nivel escuela (nivel socioeconómico promedio, rendición de cuentas, profesores calificados y otros), y a nivel de estudiantes (género, oportunidades de investigación científica, afinidad con las ciencias, etc.).

Respecto de la dependencia de los establecimientos, se puede señalar, por tanto, que no siempre constituye un factor explicativo de los resultados y además convendría explorar cuáles son las variables que intervienen en los distintos tipos de administración. La evidencia del SIMCE, por ejemplo, muestra que las distintas dependencias tienen rendimientos diferentes, en función del nivel socioeconómico de las escuelas⁷.

⁷ Ministerio de Educación de Chile: *Resultados Nacionales SIMCE, 2008*. Tablas 1.10 y 2.6

Aspectos actitudinales de los estudiantes

La comparación de Matemática y Lectura, utilizando los mismos factores, confirma las hipótesis que señalan que la escuela es mucho más determinante para la primera que para la segunda de estas disciplinas.

En efecto, en Matemática, 53% de la varianza es explicada por los factores considerados en el estudio, mientras que en Lectura, solo es el 40%. En el caso de Matemática, los factores significativos son, a nivel de estudiante, género, grado, nivel socioeconómico y cultural, expectativas laborales y valoración del rendimiento; a nivel del establecimiento, el nivel socioeconómico y cultural de los compañeros y las horas de instrucción en Matemática; y, a nivel institucional, la selección académica, la expectativa de los padres por alto rendimiento académico y la dependencia pública de los establecimientos. En Lectura, probablemente hay otros factores, no considerados, tales como aspectos del entorno familiar que, de ser medidos, darían mejor cuenta de los resultados. Para esta área de aprendizaje, se explica una menor proporción de la varianza y los factores explicativos significativos son menos que para Matemática. A nivel del establecimiento, para Lectura resultan significativos los mismos factores que en Matemática; a nivel individual también hay coincidencia, exceptuando la valoración por el rendimiento (que no es significativa en Lectura) y a nivel institucional, solo resulta significativa la selección académica.

Por su parte, en Ciencias, los factores del modelo resultan explicativos de 45% de la variación, aunque solo un factor actitudinal resulta ser explicativo: la afinidad con las ciencias.

Aspectos escolares de organización y gestión

Uno de los resultados más notables de estos estudios, de acuerdo con Valenzuela y Bellei, es la variación del rendimiento según los grados cursados por los

estudiantes, pese a que en Chile el aumento es mucho menor que en los tres países utilizados en la comparación (España, Uruguay y Polonia). En efecto, en Chile, los estudiantes progresan muy poco entre séptimo, octavo y primero medio, lo que podría explicarse, al menos, por dos factores: el hecho de que en estos grados es donde se incorporan curricularmente los ejes en que los estudiantes muestran la mayor deficiencia: Álgebra y Datos y Azar (según señalan Alfaro y Gormaz) y el hecho de que en esos grados se cuente con profesores que no son especialistas en las materias (como se plantea en los dos estudios de factores asociados; el de Lectura y el de Matemática), lo que podría modificarse con las propuestas de cambio que en esa línea plantea la nueva Ley General de Educación. Como se sabe, esta cambia la estructura del sistema educacional dejando seis años de Educación Básica y seis de Educación Media. La hipótesis en juego es que la medida favorecería una mayor eficiencia, dado que los profesores tendrán mayor formación y posiblemente a los estudiantes en grados 7 y 8 se les exija mayores responsabilidades al estar en un medio de estudiantes con más edad y no menos, como en Básica.

Por su parte, las horas de instrucción en Matemática y Lectura es un factor que se asocia positivamente y de manera significativa al rendimiento: una hora más de instrucción en una de estas asignaturas podría significar un aumento promedio de seis puntos en la evaluación de la misma, descontando el efecto de otros factores. En este tema, Chile ha avanzado en forma importante: la reforma curricular de fines de los años noventa aumentó las horas de clases anuales en 8% si no se tenía jornada completa y en 37% si además se aumentaba a jornada completa. Esto para la educación básica y 1° y 2° Medio⁸. No obstante, se debe tener en cuenta que medidas de

este tipo deben ir acompañadas de un uso eficiente del tiempo en el aula para que produzcan los cambios deseados. Probablemente, en este momento, el foco deberá ponerse en aumentar la eficiencia de esta mayor cantidad de horas de las que ya se dispone.

Un último aspecto de la organización y la gestión que resulta relevante destacar es la selección académica de los estudiantes, por parte de la escuela, la cual también se asocia positivamente al rendimiento. Chile es el país con una mayor proporción de escuelas que seleccionan a los estudiantes en términos académicos, comparado con los otros países participantes en PISA. Esto contribuye a la segregación del sistema educacional, al cual ya se ha hecho referencia, y si bien se asocia positivamente al rendimiento promedio de las escuelas, no favorece el rendimiento global de los estudiantes. Su efecto es lo que se conoce internacionalmente como “descreme” de estudiantes, que acentúa la segregación del sistema educacional, perjudicando el rendimiento de estudiantes en escuelas menos selectivas.

Género

Algo que está perfectamente claro en los modelos multiniveles comentados, es la variación en el rendimiento entre hombres y mujeres en Chile. En Matemática y Ciencias, las mujeres rinden claramente menos que los hombres. En Chile esta asociación es especialmente fuerte en Ciencias. Así, pese a que en la mitad de los países las mujeres obtienen mejores resultados y, aunque en muchos no hay diferencias significativas por género, Gubler y Williamson demuestran que en Chile, las diferencias por género se dan en todas las subescalas de Ciencias y que las mujeres chilenas se distancian más respecto de los promedios de las mujeres de los países de la OCDE, que los resultados de los hombres respecto de sus homólogos OCDE.

8 Ver: Gysling, J.: “Reforma Curricular: Itinerario de una transformación Cultural”; y Cox, C.: “Las Políticas Educativas de Chile”. En: Cox, C. (Editor): Políticas educativas en el cambio de siglo. Editorial Universitaria, 2003 y los OF-CM de básica 2002 y media 2005. Págs. 238 y 76.

En Lectura, se da la asociación contraria, ya que las mujeres internacionalmente tienen mejor puntaje promedio que los hombres y, en nuestro país, los hombres aumentaron más sus puntajes, disminuyendo así la brecha con las mujeres.

Este tema de la disparidad de rendimiento por género deberá ser abordado con urgencia, pues además de afectar la equidad, afecta en forma importante el tipo de desarrollo que se quiere lograr en el país. En Chile, a diferencia de muchos otros países, la brecha entre hombres y mujeres tiende a aumentar con el tiempo y en el transcurso de la vida escolar.

Las diferencias de género muy probablemente se generan al interior de la sala de clase, como un reflejo de lo que sucede en toda la sociedad. Esto demanda un esfuerzo decidido por parte de los profesores para combatir la disparidad. Esto se debería enfatizar desde el inicio de la formación docente. Lamentablemente, como indica el estudio de Cares, la formación de los futuros docentes no está orientada a analizar y a poner atención sobre esta situación.

Gestión o recursos, ¿dónde poner el énfasis?

Este es un tema clave de política educacional, sobre el cual mucho se ha discutido. Durante los últimos 20 años se han hecho grandes esfuerzos por aumentar la dotación de recursos de los establecimientos educacionales, como lo avalan los datos de gasto en educación del periodo. A este respecto, el artículo de Kluttig y otros, señala que, pese a que la dotación de computadores en las escuelas ha aumentado considerablemente, no ha ocurrido lo mismo con la frecuencia de uso de los mismos, y, por tanto, no se ha evidenciado el impacto de estos a nivel de rendimiento.

Por su parte, en dos de los estudios sobre factores

asociados (el de Lectura y el de Matemática), se concluye que en Chile hay un serio problema de productividad o eficiencia del sistema, por cuanto cada grado escolar adicional en Chile, no impacta en los resultados de manera tan positiva como en otros países. En efecto, aunque los estudiantes chilenos tienen más años de escolaridad (recursos), no logran compensar las diferencias de puntaje con los españoles y polacos (que superan a los chilenos), lo que se atribuye a esta falta de productividad. Por ello, son las características de los establecimientos las que, de modificarse, ayudarían a estrechar la distancia en el rendimiento con otros países. Ahora bien, de entre estas características, especialmente importante es el “efecto par” (medido como el nivel socioeconómico promedio de la escuela) y, en menor medida, las horas de instrucción en el área del caso.

No era objetivo de estos artículos profundizar respecto a cuál sería la razón de esta falta de productividad o ineficiencia. Sin embargo, al comparar las características individuales de los estudiantes chilenos, con las de estudiantes de España y Polonia queda en claro que ellas explican muy poco de la diferencia de puntajes promedio en Lectura y en Matemática, entre Chile y los otros dos países. Es decir, si los estudiantes chilenos tuvieran las características de los españoles o polacos, por ejemplo, tendrían menos grados de escolaridad, y probablemente aumentaría la brecha con esos países.

A lo anterior se suma la eficiencia de los docentes. Por ejemplo, en el análisis de los factores asociados al rendimiento en Ciencias, se señala que si las escuelas chilenas contaran con profesores más calificados, los puntajes aumentarían en 17 puntos. Este hallazgo parece contradictorio con lo observado en el artículo de Thieme, quien, con otra metodología, concluye que en Chile se necesitan más recursos y, en menor medida, mejorar la gestión. Sin embargo, al mirar el detalle del análisis realizado por Thieme, se observa que él incluye como recurso controlable por el gestor,

el índice de disponibilidad y calidad de los recursos humanos⁹; sin considerar que en Chile, el gestor municipal tiene muy poca capacidad de controlar sus recursos humanos, dadas las restricciones impuestas por el estatuto docente.

¿Qué nos dice la prueba PISA en relación con la enseñanza y aprendizaje de los estudiantes chilenos?

Es clara, y ha sido explicitada en varios artículos de este libro, la carencia de PISA para develar lo que realmente sucede al interior de la sala de clase y cuáles son sus posibilidades para explicar los aprendizajes. En efecto, PISA solo dispone de las percepciones de directores y estudiantes que responden un cuestionario. No se tienen declaraciones ni respuestas de los propios docentes y, además, no todos los estudiantes que rinden la prueba están en el mismo curso ni han tenido necesariamente la misma experiencia y trayectoria escolar.

Todo lo anterior es una limitante que obedece a posturas metodológicas, y que se compensa con la descripción detallada de lo que aprenden los estudiantes, la que permite inferir aspectos del currículo implementado, de la interpretación que se hace respecto del currículo nacional en las evaluaciones SIMCE, y de aspectos presentes o ausentes en el currículo explícito. Además, PISA permite analizar, de qué depende la dificultad de los ítems en esta prueba teniendo en cuenta la diversidad en cuanto a formato y contenidos de las preguntas. Esto, en sí mismo, resulta muy orientador para las prácticas pedagógicas.

El análisis de la prueba de Matemática nos deja una evidencia clara de que el currículo nacional en Chile incluye los contenidos que permitirían responder la

prueba PISA con mejores resultados. Los contenidos de la prueba de Lectura de PISA, también están presentes en el currículo nacional, sin embargo no incluyen contenidos de conocimientos conceptuales como sí lo hace la evaluación nacional SIMCE. En ambas pruebas (Lectura y Matemática), la manera de abordar la evaluación en PISA es tan distinta de la generalidad nacional en cuanto a formato y contextos que ello podría incidir en los menores resultados observados entre los estudiantes chilenos. En la prueba de Ciencias PISA, por su parte (que no puede ser comparada con las evaluaciones nacionales, pues no existe una prueba nacional de esta materia en Educación Media), se observa que los estudiantes chilenos siempre rinden menos que los de la OCDE, especialmente en las preguntas que evalúan las habilidades de explicar fenómenos científicos y de usar la evidencia científica¹⁰. En menor grado, también rinden menos en aquellas que evalúan la habilidad de identificar cuestiones científicas. La identificación de cuestiones científicas, reúne requerimientos de conocimiento relativo a los procesos científicos y a las principales áreas de contenido y también exige habilidades para identificar cuestiones susceptibles de ser investigadas científicamente y determinar palabras claves para buscar información. En cambio para explicar fenómenos científicos, se requiere aplicar conocimientos científicos y para usar evidencia científica, se requiere aplicar tanto conocimientos científicos como conocimientos sobre las ciencias.

En términos generales, se puede decir que algo pasa en el proceso de enseñanza-aprendizaje que a los estudiantes chilenos se les dificultan las preguntas que requieren habilidades para transferir, interpretar y aplicar determinados conocimientos. Esto debería ser objeto de estudio para determinar cuáles son las limitantes. Es posible que la forma de enseñar de los profesores en Chile no favorezca el interés de

9 Estas variables son medidas en PISA a través de la declaración de los directores, quienes declaran si en el año en curso fue posible llenar todas las vacantes de profesores que se requerían (disponibilidad) y si la enseñanza en su establecimiento se ve afectada por la escasez de profesores calificados.

10 Chile aparece junto a México, Argentina, Colombia y otros como un país con fortalezas relativas en identificar cuestiones científicas. OCDE: "PISA 2006. Science Competencies for Tomorrow's World", OCDE, 2007, Página 64.

los alumnos por desarrollar las competencias que se miden en esta prueba internacional, que predomine una interpretación academicista del currículo, que los contextos escolares no favorezcan la aplicación del conocimiento o el desarrollo de habilidades, entre otros posibles factores. Descubrir facilitadores para el desarrollo de habilidades sería un avance importante para el proceso de enseñanza aprendizaje.

En el artículo sobre Matemática, se señala que PISA opta por una profundización en el sentido de mirar una misma idea u objeto desde distintas perspectivas y esto lo contrasta con la evaluación SIMCE y con la evaluación de aula. Los profesores de Matemática podrían observar, por ejemplo, con qué énfasis desarrollan los contenidos de Geometría y Datos en Educación Básica y cuál es la relación entre profundidad y cantidad con la que se abordan los contenidos de Álgebra. También sería interesante para los docentes observar en su propia práctica, de qué manera desarrollan habilidades para razonar y para resolver problemas. Aparentemente, los aspectos que más se evalúan en el aula son los que con mayor probabilidad los estudiantes responden correctamente en PISA. Entonces, los profesores, aplicando el mismo principio de la evaluación externa, deberían evaluar a sus estudiantes en una diversidad de aspectos y, especialmente, en aquellos que los estudiantes demuestran, a través de las evaluaciones externas, haber tenido menos oportunidades de aprender. Los profesores deberían dar gran importancia al desarrollo de las habilidades de aplicar y transferir conocimientos, así como de razonar y resolver problemas, tanto en la enseñanza como en la evaluación, puesto que son requeridas en la sociedad actual. Del mismo modo, las facultades de educación deberían poner énfasis en las estrategias pedagógicas que se requieren para desarrollar estas habilidades.

Basándose en el artículo de Alfaro y Gormaz se pueden identificar aspectos de la Matemática que resultan difíciles para los estudiantes chilenos, pero no

porque no estén considerados en el marco curricular, sino por la forma de aplicar los conocimientos que se exige en la prueba PISA. En efecto, las preguntas PISA se plantean en situaciones novedosas, con cálculos numéricos de mayor dificultad, que exigen pensamiento reversible o razonamiento de múltiples pasos. También revisten mayor dificultad las preguntas referidas a información visual o a situaciones reales. Estos aspectos, de ser mejor desarrollados en el aula, ayudarían a los estudiantes a adquirir las habilidades necesarias para participar activamente en la sociedad, pero no implican necesariamente un cambio curricular. Este hallazgo puede ser un aporte a la discusión que se lleva en muchos países sobre la conveniencia de adecuar los currículos a PISA. Se sabe que los criterios para el diseño curricular no responden a exigencias o necesidades foráneas (por mucho que se deba tener en cuenta la globalización y la creciente tendencia hacia ella), pero lo que PISA parece estar mostrando, es una forma diferente de orientar los procesos de enseñanza y aprendizaje y de evaluar los contenidos.

La prueba de Lectura PISA tiene un enfoque más funcional que la prueba SIMCE en la misma área. Se observa que SIMCE tiene más textos literarios y menos textos discontinuos que PISA. El mayor número de textos discontinuos permite diversificarlos más. Así también, la prueba SIMCE no clasifica sus preguntas según el contexto, por lo que no incluye preguntas, por ejemplo, del contexto laboral. También en esta área, se alude a la diversidad de formatos de las preguntas de PISA, a diferencia de SIMCE, donde todas las preguntas de Lectura son cerradas (de selección múltiple). Este hecho reduce recursos y tiempos de corrección, pero dificulta evaluar procesos más complejos como la reflexión, la evaluación, la comparación y la elaboración de hipótesis. Esto nos lleva a pensar que SIMCE debería evaluar procesos de lectura más complejos y los profesores, deberían enfatizar complejidades de la comprensión lectora, más que agregar un gran número de contenidos. Por otra parte, PISA y SIMCE también se distinguen en

que esta última incluye, además de lectura, contenidos conceptuales que, si bien se corresponden con el carácter curricular de la prueba SIMCE, podrían determinar una mayor dificultad en ella. En este sentido, dado el avance de Chile en los resultados de PISA entre 2001 y 2006, cabe preguntarse si los profesores estarán más preocupados de desarrollar competencias para una lectura funcional que de enseñar contenidos conceptuales del currículo de Lenguaje y Comunicación. Complementariamente, habría que preguntarse si no hay además un aporte de la sociedad entera, con padres que poseen mayor escolaridad, que impulsa a una lectura más comprensiva. También sería importante analizar si las preguntas SIMCE con contenidos conceptuales efectivamente resultan más difíciles que las de exclusiva comprensión lectora.

En Ciencias, el trabajo realizado por Gubler y Williamson demuestra que en algunos aspectos lo que resulta comparativamente más difícil para los estudiantes OCDE, también resulta más difícil, aunque acentuado, para los estudiantes chilenos. Tal es el caso de las preguntas abiertas, de aquellas de contexto global y de las que se refieren a las áreas de aplicación salud o medio ambiente. En cambio, hay otros aspectos que para los estudiantes chilenos resultan particularmente difíciles, como las preguntas que requieren usar evidencia científica (especialmente las planteadas en contexto social). La práctica docente tiene valiosos elementos para nutrirse con este tipo de conclusiones. Las preguntas que resultan difíciles en todos los países pueden incluir contenidos complejos, pero aquellas que solo resultan difíciles para los alumnos chilenos habría que investigar cómo se enseñan y evalúan esos contenidos.

En los párrafos precedentes se ha intentado mostrar posibles reflexiones que serían útiles para inspirar la práctica docente. Sin embargo, en la misma sección 2, hay otro artículo (el de Cares) que, en un intento de llegar a los sectores más jóvenes y renovados del profesorado, analiza las percepciones de estudiantes

próximos a egresar de carreras de pedagogía. En este artículo se sugiere la necesidad de divulgar los resultados no solo de PISA, sino de las evaluaciones estandarizadas en general, y de las evaluaciones de aprendizajes en el aula que realizan los docentes. Esa difusión debería ocurrir en las facultades de educación y probablemente también entre los docentes en servicio. Además se señala la importancia de que las carreras de pedagogía entreguen métodos y técnicas para enfrentar la diversidad de logros.

Esta publicación como experiencia y aprendizaje

En la introducción se señaló que uno de los objetivos del proyecto que origina esta publicación es atraer el interés de investigadores y otros profesionales por analizar los resultados PISA con mayor profundidad y buscar formas de aprovechar el conocimiento generado para que sea útil tanto a la política pública como a la práctica pedagógica.

Pareciera que este objetivo empieza a cumplirse, aunque todavía queda mucho por hacer. La experiencia deja algunas certezas que vale la pena destacar. Una de ellas es que existen personas e instituciones interesadas en realizar análisis de este tipo y que con tiempo y esfuerzo de su parte y el adecuado apoyo metodológico y técnico, es posible realizar una tarea que no se puede abordar desde el Estado por sí solo. Seguramente esas personas y muchas otras, habiendo experimentado la utilidad de la información disponible, seguirán adelante, profundizando temas semejantes y utilizando los datos que proporcionan las evaluaciones internacionales y nacionales de aprendizaje. En ese sentido, este proyecto puede ser una primera iniciativa para seguir adelante en un esfuerzo colaborativo y articulado entre diversos actores. A este respecto, es necesario valorar y reconocer, en primer término, la buena disposición de los investigadores que siguieron adelante con sus análisis, a pesar de las múltiples dificultades, premuras y exigencias.

Otro aspecto que se considera facilitador, es que se disponga un apoyo de interlocución. En este caso, la comunicación no solo se dio en los talleres, sino en muchas otras oportunidades en que los investigadores quisieron consultar respecto de las bases de datos, tipos de variables o conveniencia de ciertos análisis. En esos casos, se recurrió al SIMCE, el cual posee la experiencia de haber analizado este tipo de datos y de haber participado indirectamente en los análisis que se realizan a nivel internacional. Sin duda lo más destacado de este proyecto fue el apoyo de destacados profesionales y académicos que estuvieron dispuestos a formar parte del Comité Técnico, comentando los trabajos en varias oportunidades, tanto en lo metodológico, como en el enfoque de los análisis, e incluso en la manera de comunicar los hallazgos de cada estudio.

El aspecto organizacional, que fijó un calendario y ciertos hitos a cumplir, es otro elemento que evidentemente ayudó a que los trabajos avanzaran al ritmo deseado, acercándose a la premura con que los diseñadores de política requieren información. Aunque a los investigadores y académicos pueda irritarles ciertas rigideces y también agobiarles en medio de sus tareas habituales, quedan con la satisfacción de avanzar al ritmo programado en este trabajo. Se agradece a los autores la capacidad, paciencia y buena disposición para responder a este tipo de pedidos.

Se espera que la continuidad de proyectos de este tipo, si la hubiere, permita a un mayor número de facultades y académicos incorporarse a esta experiencia y formular luego sus propios proyectos de investigación. Se debe reconocer que hubo dos situaciones que desestimularon a personas que inicialmente estuvieron interesadas. Por una parte, la restricción para mostrar ítemes no liberados tanto de PISA como de SIMCE, por razones de confidencialidad ya que estos se deben reservar para anclar los resultados con futuras evaluaciones. Por otra parte, el énfasis que se puso en los análisis cuantitativos, probablemente desincentivó

la presencia de estudios de carácter cualitativo que algunos investigadores y académicos hubieran estado dispuestos a realizar. En un próximo proyecto de este tipo, se debería considerar una mayor variedad en el carácter de los estudios.

Esta experiencia ha hecho evidente la potencialidad de un análisis acucioso de los ítemes de PISA, en contraste con evaluaciones nacionales. Los análisis presentados en la sección 2 de este libro, parecen mostrar que la intermediación entre los resultados expresados en puntajes y los especialistas en la evaluación de cada área, fuera necesaria para favorecer la aplicación práctica de los resultados en la sala de clases, siempre que se resguarden aspectos de confidencialidad. Dichos análisis permiten informar a los docentes sobre aspectos a observar y controlar en el aula, de una manera más detallada y adecuada a los resultados chilenos, que si solo se entregan puntajes, para incrementar el aprendizaje de los estudiantes. También es posible que este trabajo sea una experiencia útil para mejorar los productos SIMCE.

Finalmente, este proyecto deja un tremendo desafío tanto para el Ministerio de Educación, como para los docentes y para las facultades que forman profesores: la necesidad de desarrollar competencias de evaluación en el cuerpo docente, que las requieren tanto para aprovechar mejor las evaluaciones externas y estandarizadas, como para determinar el aprendizaje detallado de sus propios estudiantes.

Investigaciones pendientes

Como se ha dejado entrever en los párrafos anteriores, existe interés en prolongar este proyecto y repetirlo con otras evaluaciones internacionales, pues de los análisis aquí expuestos y sus conclusiones, surgen aspectos que requerirían ser investigados para un mejor e informado diseño de políticas educacionales a nivel nacional, local y de escuela. A continuación se presentan algunas propuestas.

El efecto par, medido por el nivel socioeconómico de la escuela o de los compañeros que, como se ha visto, resulta muy relevante para el aprendizaje, destaca lo nocivo del carácter segregado del sistema educacional chileno. Sin embargo, para mayor certeza respecto de esta asociación, sería conveniente hacer estudios más finos como, por ejemplo, analizar qué sucede con los estudiantes de distintos estratos en establecimientos con mayor o menor segregación. Asimismo, en relación a este aspecto tan gravitante en Chile, se deberían estudiar propuestas de estrategias exitosas en el manejo de la diversidad de logros entre los estudiantes de un mismo curso, y determinar qué competencias se requieren para aplicar estas estrategias, y desarrollarlas en la formación docente inicial.

Las disparidades entre hombres y mujeres en los aprendizajes, que se viene comprobando desde hace tiempo¹¹ y en estos estudios se reiteran, requiere de una acción decidida por parte de la política pública. Sin embargo, no se sabe cuándo, cómo ni por qué se producen. Las evaluaciones nacionales e internacionales tienen infinidad de información útil para caracterizar este fenómeno, que si bien no es exclusivo de Chile parece ser un rasgo distintivo. Estos análisis requieren complementarse con estudios cualitativos que observen las prácticas docentes, para determinar si algunas de estas favorecen o desfavorecen las disparidades por género. Este tema afecta la equidad y no debería seguir siendo ignorado. Sería importante encontrar respuestas para preguntas como ¿qué pasa en el aula?, ¿qué pasa con las representaciones de masculinidad y femineidad en relación con la enseñanza y el desempeño en Ciencias, Matemática y Lectura?, y confirmar si los colegios no coeducacionales disminuyen estas brechas y en qué situaciones o áreas de aprendizaje. Junto con lo anterior, el tema de las diferencias de género también trasciende a la escuela. Hay estudios que plantean que estas diferencias tienen que ver con las representaciones sociales de masculinidad

y femineidad, por lo que sería interesante investigar hasta qué punto es una responsabilidad exclusiva de la escuela o de qué manera la escuela podría contrarrestar estas tendencias sociales.

En relación a la competencia lectora y su evaluación, aún existen muchas preguntas abiertas. Sería de mucho interés hacer un estudio de tendencias equiparando los puntajes PISA y SIMCE. Asimismo, se podría estudiar la cohorte de 2° Medio 2006, determinando la correlación de puntajes en la evaluación nacional y en PISA, y si los estudiantes tienen peores resultados en las preguntas SIMCE con contenidos conceptuales que en aquellas de comprensión lectora exclusivamente. Este análisis podrá ser más robusto y profundo que el actual, con los resultados de PISA 2009 ya que habrá dos cohortes (2006 y 2009) que contarán con resultados SIMCE y PISA.

¿Dónde y cómo se aprende a leer, más allá de la escuela? Esta pregunta surge de la constatación de que la varianza explicada de los resultados de Lectura es inferior a la de Matemática y de otros análisis que indican que la competencia lectora depende de factores extra escolares. Esta pregunta también es atinente para saber por qué en PISA se mejoró el puntaje promedio y en SIMCE no. Es posible que los estudiantes adquieran competencias lectoras al margen de la escuela y del currículo, que les sean más útiles para PISA o también, que los profesores se estén acercando al desarrollo de la alfabetización lectora que evalúa PISA.

Estudiar y caracterizar las prácticas pedagógicas ayudaría muchísimo a determinar posibles medidas para fomentar el aprendizaje. Sucintamente, se puede señalar que ello permitiría estudiar: cómo se desarrollan habilidades para aplicar los contenidos, qué se evalúa y cómo se evalúa en el aula, cuánto se detecta y atiende a las disparidades de aprendizaje y a los factores que las determinan, el uso del tiempo y muchos otros aspectos que seguramente inciden

11 SIMCE: "Análisis de las diferencias de logro en el aprendizaje escolar entre hombres y mujeres", Julio 2005, extraído de www.mineduc.cl

en los aprendizajes, pero que permanecen en la “caja negra” de la sala de clases. Lo interesante sería hacer estos estudios relacionándolos con los resultados de las evaluaciones de aprendizaje, de tal manera de poder explicarlos sobre la base de los procesos de enseñanza y aprendizaje.

Junto con estas sugerencias de investigaciones que iluminan políticas y prácticas educativas, también se puede pensar en investigaciones sobre la misma evaluación en aspectos técnicos y psicométricos que serían muy útiles, como los artículos de la sección 3 de esta publicación. En este sentido, es indudable que el artículo de González y San Martín plantea inquietudes sobre la comparabilidad entre países dadas las diferencias de distribución de los resultados y el efecto de factores no considerados en los análisis jerárquicos. Esto motiva a seguir profundizando, con el fin de asegurar la mayor rigurosidad de las interpretaciones de los resultados. Por otra parte, el artículo de Fernández entrega un antecedente importante en cuanto a que la muestra PISA se puede considerar representativa de la población nacional, lo que justificaría otros análisis psicométricos como el de sesgo y equiparación de puntajes.

Se aprecia una ardua y variada tarea por delante, que solo es posible abordar con la colaboración más amplia y diversa posible.