



United Nations
Educational, Scientific and
Cultural Organization



UNESCO
INSTITUTE
FOR
STATISTICS



SUSTAINABLE
DEVELOPMENT
GOALS

Documento Informativo
No. 54
Septiembre de 2018
UIS/2018/ED/IP/54

Las brechas de aprendizaje: uso de datos para formular la política educativa



UNESCO

La constitución de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) fue aprobada por veinte países en la Conferencia de Londres, en noviembre de 1945, y entró en vigor el 4 de noviembre de 1946. La Organización cuenta actualmente con 195 Estados Miembros y 11 Miembros Asociados.

El principal objetivo de la UNESCO es contribuir a la paz y la seguridad en el mundo promoviendo, mediante la educación, la ciencia, la cultura y la comunicación, la colaboración entre naciones, a fin de asegurar el respeto universal de la justicia, el estado de derecho, los derechos humanos y las libertades fundamentales que la Carta de las Naciones Unidas reconoce a todos los pueblos del mundo sin distinción de raza, sexo, idioma o religión.

Para cumplir este mandato, la UNESCO desempeña cinco funciones principales: 1) estudios prospectivos sobre la educación, la ciencia, la cultura y la comunicación para el mundo del mañana; 2) el fomento, la transferencia y el intercambio de conocimientos mediante actividades de investigación, formación y educativas; 3) acciones normativas, para la preparación y aprobación de instrumentos internos y recomendaciones estatutarias; 4) conocimientos especializados que se transmiten a los Estados Miembros mediante cooperación técnica para que elaboren sus políticas y proyectos de desarrollo; y 5) el intercambio de información especializada.

Instituto de Estadística de la UNESCO

El Instituto de Estadística de la UNESCO (UIS) es la oficina de estadística de la UNESCO y es el depositario de la ONU en materia de estadísticas mundiales en los campos de la educación, la ciencia y la tecnología, la innovación, la cultura y la comunicación.

El UIS fue fundado en 1999. Se creó con el fin de mejorar el programa de estadística de la UNESCO, así como para desarrollar y suministrar estadísticas exactas, oportunas y políticamente relevantes, requeridas en un contexto social, político y económico cada vez más complejo y rápidamente cambiante.

Este documento fue redactado por J. Douglas Willms, Universidad de Nuevo Brunswick y Presidente de The Learning Bar Inc., doug.w@thelearningbar.com

Publicado en 2018 por:

Instituto de Estadística de la UNESCO
P.O. Box 6128, Succursale Centre-Ville
Montreal, Quebec H3C 3J7
Canadá

Tel.: +1 514-343-6880
Email: uis.publications@unesco.org
<http://www.uis.unesco.org>

Ref.: UIS/2018/ED/IP54

© UNESCO-UIS 2018

Esta publicación está disponible en acceso abierto bajo la licencia Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). Al utilizar el contenido de la presente publicación, los usuarios aceptan estar sujetos a los términos de uso del Repositorio UNESCO, de acceso abierto (<http://www.unesco.org/open-access/terms-use-ccbysa-en>).

Los términos empleados en esta publicación y la presentación de los datos que en ella aparecen no implican la adopción de postura alguna por parte de la UNESCO en lo atinente al estado jurídico de los países, territorios, ciudades, regiones o autoridades a los que se refieren, ni a sus fronteras o límites.



Las ideas y opiniones expresadas en esta publicación son las de los autores y no necesariamente reflejan las de la UNESCO ni comprometen a la Organización.



Reconocimientos

Agradezco a Danielle Durepos, Elizabeth Fairbairn, Alma López y Lucía Tramonte sus comentarios sobre los borradores de este escrito; a Jennifer Morgen, por haberse encargado de la programación de las cifras utilizadas en el informe; y a otros miembros del personal de The Learning Bar Inc., que brindaron apoyo a este trabajo en varias etapas. Muchas de las ideas del presente informe emanan de una colaboración a largo plazo con Lucía Tramonte, quien ayudó a esclarecer diversos conceptos a fin de que puedan aplicarse más directamente a la investigación en curso para la Agenda de Desarrollo Sostenible de la Organización de las Naciones Unidas.

El modelo de prosperidad educativa se elaboró con el apoyo de líderes escolares de Canadá y Australia, y posteriormente se adaptó a las necesidades del Programa para la Evaluación Internacional de Alumnos para el Desarrollo (PISA para el Desarrollo o PISA-D) de la Organización para la Cooperación y el Desarrollo Económicos (OCDE). En el proceso, recibí numerosos comentarios útiles de educadores y directores en diversas jurisdicciones. Agradezco las contribuciones de Michael Ward y del personal de la OCDE, así como de los miembros del Grupo de Expertos en Cuestionarios y el Grupo Asesor Internacional, y los gerentes y analistas de proyectos nacionales de los ocho países participantes en PISA para el Desarrollo.

Agradezco asimismo a Eduardo Backhoff, David Berliner, Harvey Sánchez, Valérie Tehio y Servaas Van der Berg, quienes realizaron revisiones exhaustivas del documento y externaron muchas opiniones importantes.

Vaya también mi agradecimiento para Silvia Montoya, Directora del Instituto de Estadística de la UNESCO, por apoyar mi investigación y alentarme a escribir el presente informe.



Índice

Reconocimientos.....	3
Introducción.....	6
I. Prosperidad educativa: un enfoque de ciclo vital	11
Prosperidad	11
Resultados de prosperidad	12
Fundamentos para el éxito.....	12
Cuatro procesos que determinan cómo se acumula el éxito.....	13
Igualdad, equidad y acceso.....	14
Importancia del nivel socioeconómico	17
II. Dos transiciones críticas para el desarrollo de las habilidades de lectura y escritura.....	19
Ingreso a la educación primaria baja	19
De “aprender a leer” a “leer para aprender”	21
III. Políticas sobre resultados e igualdad de los estudiantes	27
¿Cómo vamos?.....	27
¿Quiénes son vulnerables?	31
¿Dónde se encuentran los estudiantes vulnerables?.....	34
IV. Políticas sobre estrategias y su ejecución	42
Fortalecimiento de los fundamentos para el éxito	48
Modificación de las características estructurales de las escuelas.....	52
V. Seguimiento de la prosperidad educativa.....	55
Establecimiento de objetivos.....	55
Un ejemplo de programa de seguimiento	58
Nuevos rumbos.....	58
Referencias	62
Apéndice 1. Tasas de crecimiento de la competencia lectora en los resultados de PISA, 2000-2015	70
Apéndice 2. Estadísticas clave para formular la política educativa.....	71
Figura 1. Tasas de crecimiento anual de la competencia lectora según los resultados de PISA, 2000-2015....	9
Figura 2. Marco de prosperidad educativa.....	11
Figura 3. Cuatro procesos que determinan cómo se acumula el éxito.....	13
Figura 4. Efectos de la igualdad, la equidad y la escuela en la prosperidad educativa.....	15
Figura 5. Porcentaje de estudiantes de 15 años con habilidades lectoras del nivel 2 o inferiores (PISA 2015), en relación al ingreso nacional bruto (INB) (método Atlas, en dólares estadounidenses).....	18



Figura 6. Puntuaciones de la Evaluación Infantil Temprana en habilidades cognitivas y lenguaje, realizada en Uruguay en 2017	21
Figura 7. Puntuaciones de competencia lectora en Australia, basadas en el programa NAPLAN 2017.....	24
Figura 8. Distribución de la competencia lectora a los 15 años en la República Dominicana, México y Suecia.....	28
Figura 9. Relación entre la media de la competencia lectora y la asimetría.....	29
Figura 10. Gradiente socioeconómico de la competencia lectora en México	32
Figura 11. Gradientes socioeconómicos de la competencia lectora en México por primera lengua	34
Figura 12. Perfil escolar de competencia lectora en México.....	35
Figura 13. Perfil escolar de baja competencia lectora (nivel 1 o inferior) en México	36
Figura 14. Concentración de estudiantes con baja competencia lectora en escuelas mexicanas	39
Figura 15. Efectos de una estrategia universal con un tamaño de efecto de 0.50	43
Figura 16. Efectos de una estrategia centrada en el desempeño con un tamaño de efecto de 0.50, dirigida a estudiantes con una competencia lectora del nivel 1 o inferior	45
Figura 17. Efectos de una estrategia centrada en el riesgo con un tamaño de efecto de 0.50 dirigida a estudiantes con un NSE de -2 o inferior	45
Figura 18. Efectos de una estrategia compensatoria que eleva el NSE de los estudiantes de bajo NSE en desviaciones estándar de 0.25.....	46
Figura 19. Efectos de una estrategia de reubicación de los estudiantes de escuelas de bajo NSE en escuelas convencionales.....	47
Figure 20. Marco de prosperidad educativa para PISA para el Desarrollo	49
Figura 21. Gráficos de recursos escolares de México.....	51
Figura 22. Áreas de conocimiento y mediciones para un sistema nacional de seguimiento.....	59



Introducción

Tanto los países, incluidos los más ricos del mundo, como sus escuelas difieren considerablemente entre sí en cuanto a las habilidades lectoras de los estudiantes. Éstos son dos de los resultados del informe *Las brechas de aprendizaje: diez preguntas en materia de política sobre el desempeño y la equidad en las escuelas y los sistemas educativos* (Willms, 2006). El término “brechas de aprendizaje” se utilizó porque los análisis revelaron una alta variabilidad de la relación entre las habilidades lectoras de los estudiantes y el nivel socioeconómico (NSE) de sus respectivas familias en los distintos países. La fuerza de esta relación es un indicador de desigualdad, o “brecha de aprendizaje”, entre los estudiantes de contextos con un NSE bajo y alto. En promedio, los países con un bajo nivel de desigualdad mostraron niveles de habilidades lectoras más altos en general. Estas diferencias entre países y entre escuelas de distintos países fueron evidentes en el caso de los estudiantes que se encontraban al final del cuarto grado, y aumentaron conforme avanzaron en sus estudios.

El informe de 2006 se basó en análisis de datos del Estudio Internacional de Progreso en Comprensión Lectora (PIRLS) 2001 y el Programa para la Evaluación Internacional de Alumnos (PISA) 2000. El PIRLS es una evaluación de las habilidades lectoras de los alumnos de cuarto año de primaria y se lleva a cabo cada cinco años desde 2001, bajo los auspicios de la Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA). PISA es una evaluación de los conocimientos y aptitudes para la vida de jóvenes de 15 años que realiza la Organización para la Cooperación y el Desarrollo Económicos (OCDE). Esta evaluación pone énfasis en las habilidades en lectura, matemáticas y ciencias que los estudiantes necesitan cotidianamente cuando cursan estudios postsecundarios o ingresan a la fuerza laboral. Se ha llevado a cabo cada tres años desde 2000.

En *Brechas de aprendizaje* se usaron los gradientes socioeconómicos como una estructura general para abordar preguntas sobre la política educativa. Un gradiente socioeconómico o “barra de aprendizaje” simplemente describe la relación entre un resultado escolar o social y el NSE de los individuos en una jurisdicción específica, tal como una escuela, una región o un país (Willms, 2003a). El objetivo principal de la política educativa es mejorar los resultados de los estudiantes y reducir las desigualdades asociadas con los antecedentes familiares; en otras palabras, es “elevar y nivelar la barra de aprendizaje”. El informe mostró que un análisis detallado de gradientes socioeconómicos que tiene en cuenta la estructura jerárquica del sistema escolar puede proporcionar a los encargados de formular políticas información valiosa acerca las formas de intervención más adecuadas para mejorar las habilidades de los estudiantes y reducir las desigualdades. El informe también abordó la cuestión del “valor agregado”: “¿La variación entre escuelas es atribuible a los niveles de recursos escolares y a la política y las prácticas en escuelas y aulas?” (Willms, 2006, p. 54). Los análisis indicaron que después de considerar los antecedentes familiares de los estudiantes, se observó que las relaciones entre sus logros y las mediciones de las prácticas en escuelas y aulas eran pequeñas y, en la mayoría de los casos, estadísticamente insignificantes.



Este nuevo informe es una secuela del de 2006. Se basa en un marco de evaluación denominado “prosperidad educativa” que puede usarse para dar seguimiento al éxito de las familias, las comunidades y las instituciones públicas en el desarrollo de las habilidades cognitivas y al bienestar social, emocional, físico y espiritual de los niños. La prosperidad educativa adopta un enfoque de ciclo vital y considera los procesos que determinan cómo evolucionan los resultados de los niños desde la concepción hasta el final de la adolescencia. El marco incluye un conjunto básico de “resultados de prosperidad” para cada una de las seis fases de desarrollo y una serie de factores familiares, institucionales y comunitarios que determinan esos resultados y se denominan “fundamentos para el éxito”. Asimismo, distingue entre igualdad y equidad, y vincula explícitamente el seguimiento de los sistemas educativos con la política nacional y local. El marco de prosperidad educativa se describe con mayor detalle en otra publicación (Willms, 2018a).

El rol de los estudios internacionales a gran escala para orientar la formulación de la política educativa se ha basado principalmente en dos abordajes. Uno de ellos consiste en recopilar datos sobre un sinnúmero de factores de la escuela y del aula, y determinar sus relaciones con los resultados de la educación. Esta “búsqueda de efectos escolares” ha sido una característica central de todos los estudios internacionales a gran escala, incluidos los de PISA y los de la IEA. Los marcos contextuales se han fundado en un paradigma de “función de producción” de insumo-proceso-producto, que persigue captar los principales factores a nivel del estudiante, de la familia, de la escuela y del aula que explican los logros del estudiante. Los análisis de datos se han basado en modelos de regresión multinivel que examinan las relaciones entre los resultados de un estudiante, por ejemplo, en lectura, y una larga lista de factores de la escuela y del aula. Muy a menudo, estos modelos estadísticos se valoran por separado para cada país, con la idea de que los factores de la escuela y del aula pertinentes para las habilidades de los estudiantes varían entre países. Los resultados de los análisis sirven para respaldar diversas políticas nacionales.

El segundo enfoque consiste en comparar los resultados de distintos países. Los factores que se consideran pertinentes para el éxito de los estudiantes se agrupan en varios temas de política, principalmente: recursos escolares, rendición de cuentas, gobernanza escolar, prácticas de enseñanza y escolaridad selectiva. Así, un país puede determinar si sus políticas difieren de las de otros países que tienen un contexto socioeconómico similar. No obstante, la comunidad encargada de formular políticas internacionales tiende a considerar sobre todo los casos exitosos. Por ejemplo, los robustos resultados de Finlandia en el primer estudio PISA dieron lugar a numerosos informes sobre los factores de éxito de su sistema educativo, principalmente la pericia de los profesores y el enfoque de evaluación de los estudiantes (Grek, 2009; Simola, 2005).

Ambos enfoques son problemáticos. La cuestión más importante es que el desempeño de los estudiantes en las pruebas PIRLS de final del cuarto grado o en las pruebas PISA aplicadas a estudiantes de 15 años son el *resultado acumulativo* de incontables factores que influyen en el desarrollo de los niños, desde la concepción hasta el momento de la evaluación. Sus habilidades cognitivas y lingüísticas cuando ingresan al primer grado son variables predictivas sólidas de si tendrán o no éxito en la lectura dos o tres años después (Scarborough, 1989; Schatschneider et al., 2004). Las habilidades lectoras al final de la escuela primaria son variables predictivas sólidas de las habilidades lectoras a los 15 años (Adlof, Catts y Lee, 2010). Por lo tanto,



no debemos esperar que las mediciones de las prácticas de escuela o de aula, que se derivan de los cuestionarios administrados simultáneamente a las pruebas de aptitud, guarden una fuerte relación con la competencia lectora.

Además, las mediciones de los principales factores de la escuela que influyen en el desempeño del estudiante tienden a estar correlacionadas entre sí y fuertemente correlacionadas con el NSE promedio de la escuela. Es prácticamente imposible aislar con un estudio transversal los “efectos escolares” atribuibles a determinados recursos o procesos (Raudenbush y Willms, 1995). El segundo enfoque también es problemático por las mismas razones. Es posible que un país tenga mejores resultados en la lectura que otro porque cuenta con sólidos fundamentos para el desarrollo de los niños en sus primeros años; así que cualquier comparación de los efectos de las políticas escolares puede arrojar resultados espurios. En pocas palabras, no se pueden hacer afirmaciones causales basadas en los resultados de estudios nacionales o internacionales.

Las habilidades lectoras de los estudiantes no han mejorado en los últimos 15 años. Ésta es la razón de más peso para adoptar un nuevo enfoque en el uso de los datos para formular la política educativa. La **Figura 1** muestra la tasa de crecimiento anual correspondiente a los 28 países que participaron en PISA de 2000 a 2015. Para la evaluación de 2000, la escala de puntuaciones de la competencia lectora se ajustó a modo de obtener una media de 500 y una desviación estándar de 100 para todos los países miembros de la OCDE. Las escalas fueron calibradas en cada uno de los ciclos para poder estimar los cambios en el desempeño utilizando una escala común¹. Los puntos verdes indican que la tasa de crecimiento en los países en cuestión fue estadísticamente significativa. En el Cuadro 1 del Apéndice se desglosan los resultados.

En promedio, la tasa de crecimiento anual fue ligeramente negativa, de -0.004% de una desviación estándar. Las tasas de crecimiento anual de 15 de los 28 países no fueron estadísticamente significativas. De los países con un crecimiento anual significativo, nueve registraron una tasa de crecimiento anual negativa; cuatro, una tasa positiva; y todos los países, una tasa de crecimiento anual inferior al 2.5% de una desviación estándar, ya sea positiva o negativa. Esto equivale a 2.5 puntos en la escala PISA, que tiene una media de 500 y una desviación estándar de 100. Para situar estos resultados en un contexto más amplio, una intervención de lectura adecuadamente diseñada puede mejorar la competencia lectora de los estudiantes en 50 o 60 puntos, es decir, un tamaño de efecto de 0.50 a 0.60. Esta estimación se basa en el resumen de Hattie (2009) de 50 metanálisis de intervenciones de lectura. El tamaño de efecto promedio fue de 0.51; y, en el caso de los 14 estudios que implicaron una enseñanza directa de la conciencia fonológica, de 0.60.

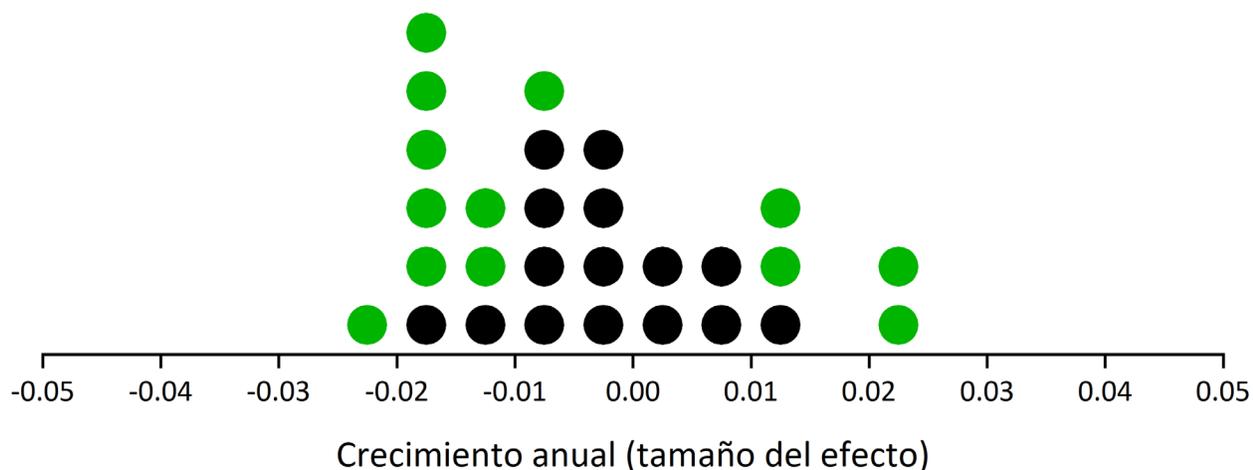
¹ Las estimaciones que se presentan en la Figura 1 se basaron en las muestras completas de datos recopilados para cada país en cada ciclo, y se ajustaron al NSE para tener en cuenta los cambios en este indicador en las muestras. Los análisis se basaron en un modelo de “estabilidad” jerárquico de dos niveles que se adecuó por separado para cada país, utilizando el NSE como covariable en el nivel 1 y el “año” en el nivel 2 (véase Willms y Raudenbush, 1989).



Algunos de los cambios observados en las puntuaciones PISA pueden atribuirse a un error de medición y muestreo que resultó en una “regresión a la media”: los países con puntuaciones altas en 2000 tuvieron una mayor probabilidad de obtener tasas de crecimiento anual negativas; y aquellos con puntuaciones bajas en 2000, una mayor probabilidad de obtener tasas de crecimiento positivas. La correlación entre las puntuaciones PISA de los países en 2000 y las tasas de crecimiento anual es de -0.39. Finlandia, el país con la puntuación más alta en 2000, registró el descenso más grande: sus puntuaciones ajustadas al NSE bajaron en casi un 2.5% de una desviación estándar por año.



Figura 1. Tasas de crecimiento anual de la competencia lectora según los resultados de PISA, 2000-2015



Fuente: PISA 2015.

El enfoque adoptado en este informe se basa en tres premisas. Primero, los sistemas de seguimiento educativo tienen que centrar la atención en el desarrollo de las habilidades lectoras de los niños. Esto es un requisito previo para el desarrollo de competencias académicas fuertes en la secundaria baja y alta, y un factor esencial para la terminación de estudios y la justicia social (Beswick, Sloat y Willms, 2008; Snow, Burns y Griffin, 1998; Willms, 2006). Segundo, la documentación sobre los efectos de la escuela y del aula nos ha proporcionado los conocimientos necesarios para construir sistemas de seguimiento educativo informativos. Para proseguir con la búsqueda de efectos escolares, no necesitamos estudios nacionales o internacionales a gran escala con numerosas medidas de los factores de la escuela y del aula. Lo que necesitamos es que esos estudios se centren en un pequeño número de factores, los midan con mayor detalle y les den un seguimiento longitudinal. Tercero, los resultados de grandes estudios internacionales, combinados con los de estudios nacionales y pequeños estudios experimentales controlados, pueden proporcionar a los administradores educativos información para establecer objetivos asequibles, asignar los recursos y evaluar los efectos de las políticas que modifican una o varias de las características estructurales de la educación.

Esta investigación no es un llamado a abandonar los estudios internacionales a gran escala; de hecho, muchos de los ejemplos utilizados en este informe se basan en datos de PISA. La mayoría de los países de ingresos bajos y medios aún no han participado en una evaluación internacional, y les beneficiaría saber qué tan bien les va a sus estudiantes en comparación con los de otros países. Además, los resultados de los estudios comparativos a menudo estimulan la voluntad política de un país para invertir recursos en la educación (Singer y Braun, 2018). En cambio, lo que pretendemos es desviar la atención del ordenamiento jerárquico de los países según sus resultados y de las afirmaciones causales basadas en datos transversales.



Los ejemplos que se presentan en las secciones III y IV de este informe usan datos de PISA 2015 relativos a México. Sin embargo, no se trata de evaluar el sistema educativo mexicano. Backhoff, Bouzas, Hernández y García (2007) hacen una evaluación detallada de los resultados de México de acuerdo con la estructura establecida en el informe de 2006 sobre las brechas de aprendizaje (Willms, 2006). Martínez y Díaz (2016) presentan resultados detallados de PISA 2015 sobre México.

El propósito general del presente informe es establecer una estructura para el uso de datos de seguimiento que contribuyan a la formulación de dos tipos de políticas educativas: las relacionadas con la mejora de los resultados educativos y la reducción de las desigualdades; y las que se refieren a estrategias para alcanzar los objetivos educativos. En este informe, el término “estrategias” se emplea en un sentido amplio a fin de incluir cursos de acción tales como la reasignación de recursos entre subpoblaciones de estudiantes o escuelas, la adopción de nuevos planes de estudios o enfoques de enseñanza, o cambios en una característica estructural clave de la enseñanza. Para ello se hace uso de quince estadísticas fundamentales y cinco enfoques gráficos –“herramientas de trabajo”–. Los detalles técnicos del cálculo de las quince estadísticas se presentan en el Apéndice 2 y a lo largo del informe se dan ejemplos de los enfoques gráficos.

La Sección I de este informe contiene un resumen del marco de prosperidad educativa. La Sección II aborda las dos transiciones más importantes para el desarrollo de las habilidades de lectura y escritura. La Sección III trata sobre las políticas relativas al desempeño de los estudiantes. ¿De qué manera se pueden usar los datos de seguimiento para responder a la pregunta “¿Cómo vamos”? Exponemos argumentos a favor de la identificación de los estudiantes vulnerables y al establecimiento de objetivos realistas para su desarrollo adecuado. La Sección IV examina estrategias para alcanzar los objetivos educativos. Se describen cinco tipos de estrategias y se analiza cómo se pueden utilizar los datos de seguimiento para discernir qué estrategias tienen más probabilidades de mejorar el desempeño de los estudiantes. Una estrategia o intervención puede estar destinada a fortalecer los fundamentos para el éxito, reducir las desigualdades en el suministro de recursos escolares, o modificar una o varias de las características estructurales de la educación. En esta sección se examina cada uno de dichos objetivos. La Sección V describe un método para usar los datos de seguimiento con la finalidad de establecer objetivos, ilustra un sistema de seguimiento basado en la prosperidad educativa, y concluye con un análisis de las implicaciones del seguimiento de escuelas y sistemas escolares basado en el marco de prosperidad educativa.



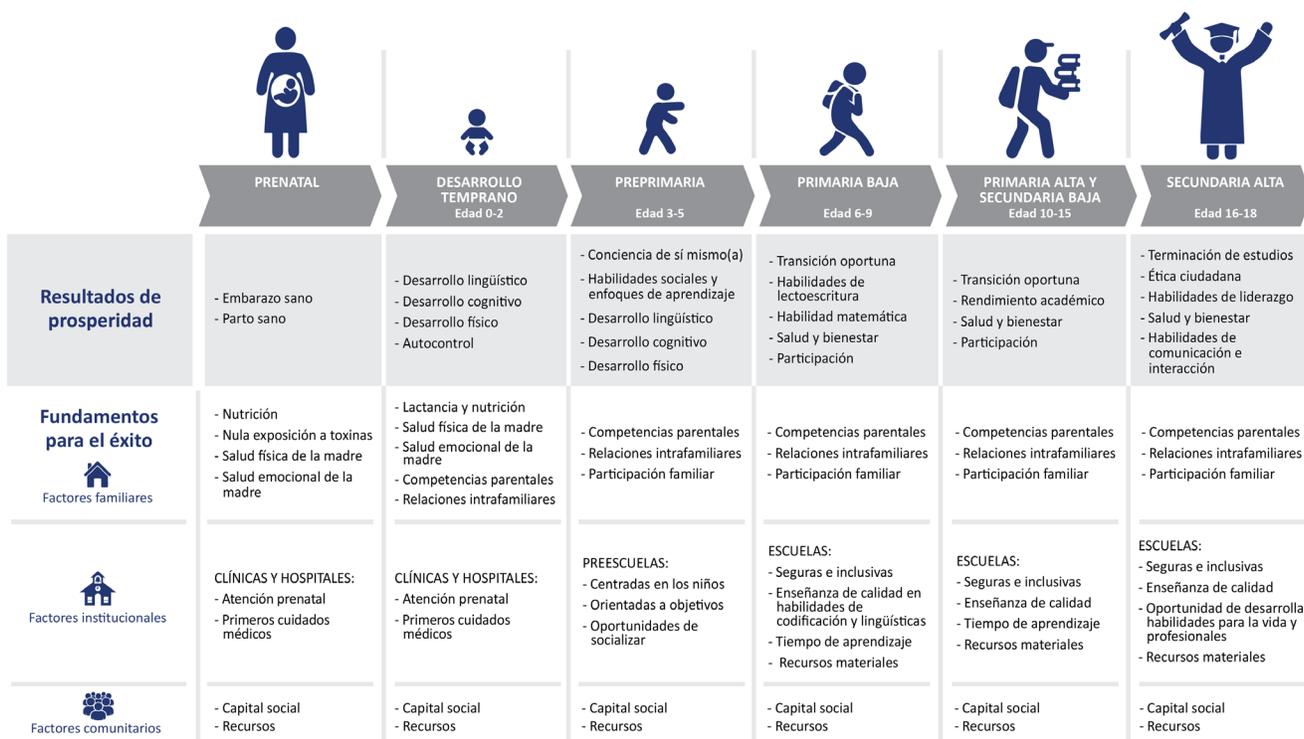
I. Prosperidad educativa: un enfoque de ciclo vital

Prosperidad

“Prosperidad” significa desarrollar al máximo el potencial propio. El objetivo es *brindar a todos los niños y todas las niñas la oportunidad de experimentar el éxito*. “Prosperidad educativa” significa que las familias, las comunidades y las instituciones públicas alcanzan ese objetivo al posibilitar que los niños desarrollen su bienestar académico, social, emocional, físico y espiritual. La prosperidad educativa es un marco de evaluación del que los países, las jurisdicciones escolares y las escuelas pueden servirse para vigilar el avance hacia dicho objetivo.

La **Figura 2** muestra el marco de prosperidad educativa (Willms, 2018a), el cual incluye seis fases de desarrollo, que van desde la concepción hasta la secundaria alta. Cada fase abarca un conjunto pequeño de “resultados de prosperidad” y una serie de factores familiares, institucionales y comunitarios denominados “fundamentos para el éxito”, que determinan esos resultados.

Figura 2. Marco de prosperidad educativa





Resultados de prosperidad

Los resultados de prosperidad se consideran universales en el sentido de que constituyen marcadores clave del desarrollo de los niños y son necesarios para que todos ellos prosperen. Es posible que las prioridades de los países difieran en cuanto a los resultados de las distintas fases, especialmente en la última, pero los resultados incluidos se consideran deseables en todos los contextos. Además, son congruentes con los Objetivos de Desarrollo Sostenible (ODS) de la UNESCO (Instituto de Estadística de la UNESCO, 2017).

Fundamentos para el éxito

Los fundamentos para el éxito también se consideran universales en la medida en que un amplio conjunto de investigaciones confirma que son condiciones necesarias para el éxito en cada fase de desarrollo. Se tuvieron en cuenta tres criterios para determinar qué factores incluir como fundamentos para el éxito: que fueran potentes, cercanos y generalizados.

Un factor “potente” es aquel que muestra una fuerte correlación con un resultado y que, según investigaciones previas, tiene una relación *causal* con ese resultado. Un factor se considera “causal” si se ha demostrado que es previo a un resultado y está correlacionado con él; y que cualquier cambio que sufra modificará el resultado (Kraemer et. al., 1997). La enseñanza de calidad es potente porque se correlaciona con el desempeño académico durante el periodo de escolaridad, y las estrategias que la mejoran elevan dicho desempeño (Anderson, 2004; Creemers y Kyriakides, 2006; Kyriakides, Christoforou y Charalambous, 2013; Rosenshine, 2010).

Un factor “generalizado” se correlaciona positivamente con una amplia variedad de resultados; e, idealmente, la investigación previa respalda las afirmaciones de que guarda una relación causal con cada resultado. Por ejemplo, una escuela “segura e inclusiva”, que es uno de los fundamentos para el éxito en las últimas tres fases del marco, afecta al desempeño académico, los logros educativos, la participación, la salud y el bienestar de los estudiantes, entre muchos otros resultados.

Un factor es “cercano” a los resultados en la medida en que guarda una relación directa con ellos; es decir, su efecto no depende de ningún otro factor. El tiempo destinado para el aprendizaje tiene un efecto directo sobre los resultados de los estudiantes; es decir, su efecto no depende de ningún otro factor. El desarrollo profesional de los maestros no se incluye entre los fundamentos para el éxito porque depende de otros factores. Por ejemplo, sólo es efectivo si da como resultado una mejora en la calidad de la enseñanza, un contexto más inclusivo, o un mayor apoyo familiar y comunitario.



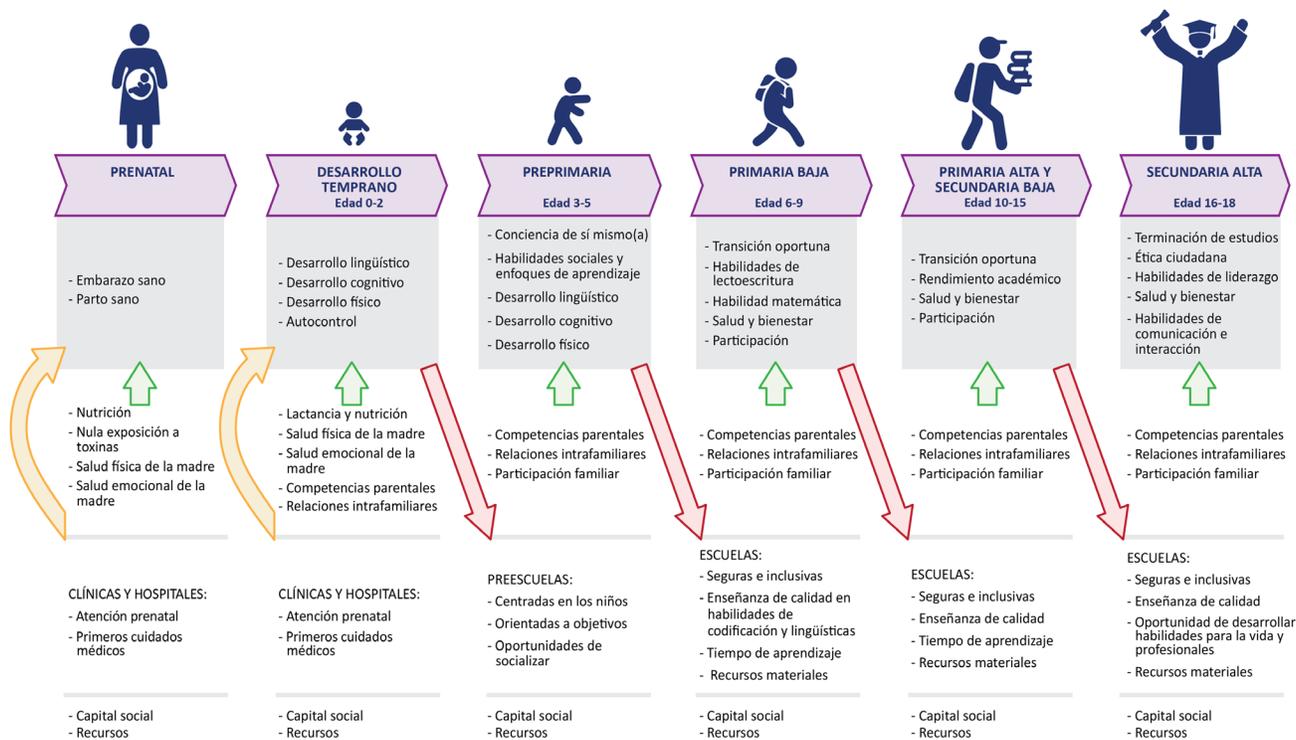
Figura 3. Cuatro procesos que determinan cómo se acumula el éxito

Flechas anaranjadas: incorporación biológica

Flechas verdes: fundamentos para el éxito

Flechas moradas: efectos acumulativos

Flechas rojas: efectos de la selección institucional



Cuatro procesos que determinan cómo se acumula el éxito

El marco de prosperidad educativa se basa en un modelo de desarrollo que incluye cuatro procesos que determinan cómo se acumula el éxito durante el ciclo vital de los estudiantes, desde la concepción hasta el final de la adolescencia. Esto se ilustra en la **Figura 3**.

Incorporación biológica. Un embarazo y parto sanos se ven afectados por los fundamentos para el éxito (flecha verde): nutrición, nula exposición a toxinas, y salud física y mental de la madre. Hasta cierto punto, algunos de estos resultados se incorporan biológicamente (flecha anaranjada) a través de procesos epigenéticos en los que las improntas químicas se vinculan a genes que predisponen al niño a ser vulnerable o resistente (Boyce y Kobor, 2015). Las primeras experiencias del niño interactúan con su disposición genética en formas que afectan al desarrollo del cerebro y otros sistemas neurológicos y biológicos asociados con un crecimiento sano (Boyce, Sokolowski y Robinson, 2012).



Al nacer, los niños tienen miles de millones de neuronas; y durante su desarrollo temprano las neuronas crean conexiones, llamadas sinapsis, en respuesta a los estímulos del entorno. En este proceso, muchas de las neuronas que no se están utilizando se descartan y las restantes se fortalecen. Las conexiones para la audición y la visión se establecen tempranamente, seguidas de cerca por las del control emocional, del lenguaje y del funcionamiento cognitivo (Center on the Developing Child, 2007; Knudsen, 2004; Shonkoff y Phillips, 2000). El proceso de formación de sinapsis y descarte de neuronas –esculpido del cerebro– es más rápido durante ciertos *períodos críticos* de los dos o tres primeros años de vida (McEwen y Schmeck; 1994; Cynader y Frost, 1999; Hertzman, 1999).

Fundamentos para el éxito. Después del nacimiento, los fundamentos para el éxito apoyan el desarrollo del niño. En el período del nacimiento a los 2 años de edad, por ejemplo, las interacciones con los padres son decisivas para el desarrollo de sus habilidades lingüísticas y de autocontrol (McClelland et. al., 2010). Por consiguiente, se considera que los estilos de crianza y las relaciones intrafamiliares son dos de los fundamentos para esta fase y las subsecuentes. Durante el período preescolar y los tres períodos siguientes, los fundamentos para el éxito son un entorno seguro e inclusivo, una enseñanza de calidad, el tiempo de aprendizaje, los recursos materiales y el apoyo familiar y comunitario.

Efectos acumulativos. Conforme avanzan de una fase a la siguiente, los niños acumulan resultados de prosperidad. Las habilidades lingüísticas se desarrollan a lo largo de la infancia, pero algunas habilidades son requisitos previos para el desarrollo de otras. Por ejemplo, el desarrollo de fuertes habilidades cognitivas durante los años preescolares determina si el niño aprenderá a leer bien en la escuela primaria (Leppänen et. al., 2004; Nation y Snowling, 2004). Aunque los factores familiares tienen una influencia significativa en los resultados de prosperidad, después de los cinco años de edad las escuelas desempeñan un papel de creciente importancia. Por ello, si el niño no adquiere habilidades fuertes en sus primeros años de vida, aumenta el riesgo de que fracase en la escuela.

Selección institucional. Cuando los estudiantes son exitosos en una fase de desarrollo, su ciclo vital puede verse alterado si se les selecciona para ser parte de ciertos grupos, programas académicos o escuelas. En muchos sistemas escolares, los niños que poseen grandes habilidades lectoras y lingüísticas son canalizados a clases o programas académicos con fundamentos fuertes. Estos niños tienen más probabilidades de beneficiarse de interacciones positivas con sus compañeros, una enseñanza de mayor calidad y otros factores que les permitirán desarrollar sus habilidades más rápido (Willms, 2006).

Igualdad, equidad y acceso

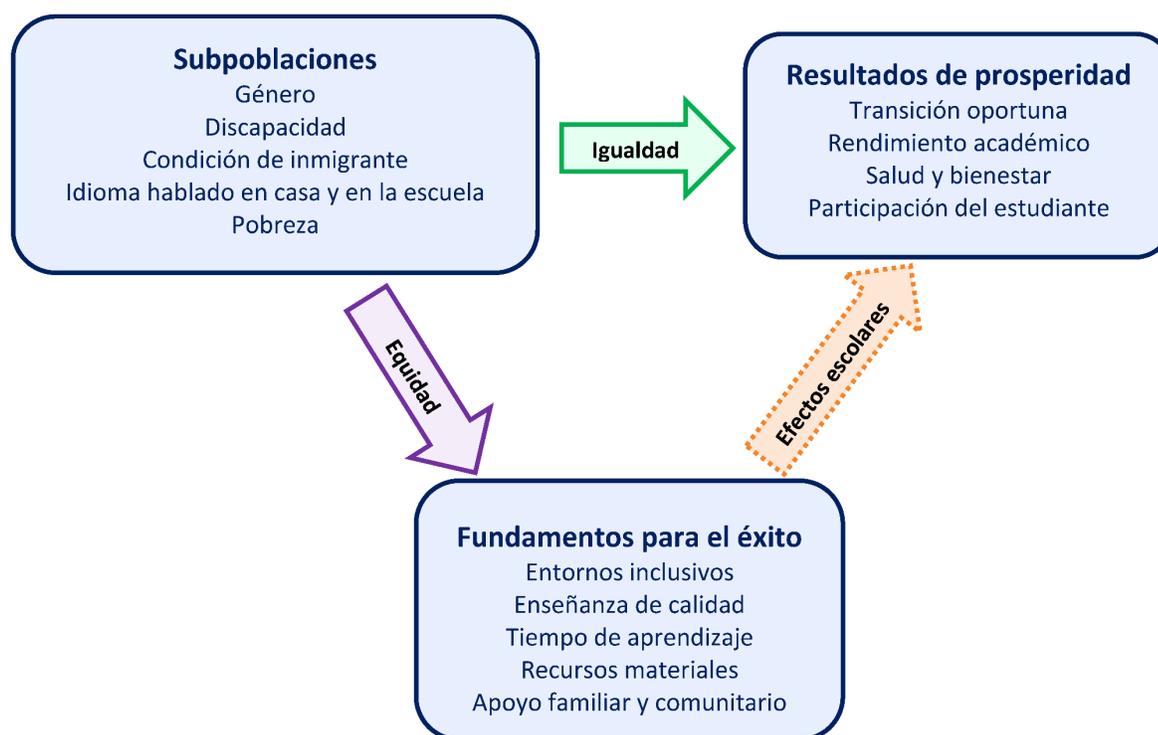
La Convención de las Naciones Unidas sobre los Derechos del Niño, firmada en 1989, reconoce los derechos del niño a un nivel de vida adecuado para su desarrollo físico, mental, espiritual moral y social; al disfrute del más alto nivel posible de salud y a servicios para la rehabilitación de la salud; y a una educación de calidad, a fin de que pueda ejercer estos derechos progresivamente y en condiciones de igualdad de oportunidades. La mayoría de los países reconoce que el desarrollo económico y social exige que se imparta educación desde la primera infancia hasta la adolescencia; por consiguiente, los gobiernos han establecido



garantías constitucionales y legales para la educación universal (Levin, 2009). No obstante, aun en los países más ricos, los niños que pertenecen a familias con un nivel socioeconómico alto tienen mayor acceso a recursos económicos, sociales y educativos que aquellos cuyas familias tienen un nivel socioeconómico bajo (Willms, 2006). Por ello, a los líderes educativos les ha interesado cada vez más cuantificar y monitorear las desigualdades en los resultados educativos entre subpoblaciones, y determinar los factores asociados.

Los términos “igualdad”, “equidad” y “acceso” han sido utilizados de distintas formas por investigadores, y muy a menudo los utilizan indistintamente. En este informe, en consonancia con Willms, Tramonte, Duarte y Bos (2012) y Willms (2011), se distingue entre igualdad y equidad y se considera que el acceso es un aspecto de la equidad. El modelo de prosperidad educativa fue adoptado por PISA para el Desarrollo (PISA-D), una iniciativa de la OCDE que tiene por objeto hacer las evaluaciones PISA más pertinentes para los países de ingreso bajo y medio (OCDE, 2017). Para ilustrar los conceptos, en la **Figura 4** se muestra el modelo de trayectorias de prosperidad educativa utilizado en PISA-D.

Figura 4. Efectos de la igualdad, la equidad y la escuela en la prosperidad educativa



Igualdad. La igualdad tiene que ver con las diferencias de distribución entre los resultados educativos por subpoblaciones. En la Figura 4 se representa como la trayectoria (flecha verde) que vincula los factores demográficos (género, discapacidad, etc.) con los resultados de prosperidad (flecha verde). La medición de la igualdad es relativamente directa. Por ejemplo, la diferencia entre las niñas y los niños en su puntuación



promedio en lectura en PIRLS es una medida de la igualdad. Cabe hacer notar que la definición se refiere a las diferencias en la *distribución* de los resultados. En el caso del resultado de puntuaciones en lectura, una importante diferencia entre los sexos es la distribución de las mismas. Un marcador pertinente de la igualdad es la prevalencia de niños y niñas que no alcanzan un nivel mínimo. Las diferencias de resultados entre los grupos socioeconómicos también es una medida de la igualdad. Más adelante en este informe se utilizan gradientes socioeconómicos para evaluar las desigualdades.

Equidad. La equidad tiene que ver con la imparcialidad –un trato justo a las personas que pertenecen a distintas subpoblaciones. Se refiere a las diferencias entre subpoblaciones en lo tocante al acceso a los recursos y los procesos escolares que afectan a los resultados educativos. En la Figura 4 se representa como la trayectoria (flecha morada) que vincula los factores demográficos con los fundamentos para el éxito. Medir la equidad ha sido difícil porque primero se tienen que identificar los recursos y los procesos escolares que afectan a los resultados educativos. En la Figura 4 se les denomina “efectos escolares” (flecha anaranjada). Se presume que son “efectos causales”. No se pueden usar estudios transversales a gran escala para determinar qué factores tienen efectos escolares. El marco de prosperidad se basa en la literatura más amplia para identificar los fundamentos para el éxito.

Esta distinción entre la igualdad y la equidad es útil porque las mediciones de la igualdad indican la magnitud de las diferencias entre subpoblaciones –¿qué tan grande es el problema?; mientras que las mediciones de la igualdad requieren políticas que aborden el problema –¿qué se necesita cambiar para reducir las desigualdades?

Acceso. El acceso a la educación se refiere a la libre disponibilidad de la enseñanza para los niños en una jurisdicción. El modelo de prosperidad educativa trata el acceso como un aspecto de recursos materiales, lo cual plantea interrogantes acerca de la equidad –por ejemplo, “¿Los niños discapacitados tienen el mismo acceso a la educación que los no discapacitados?”. Lo más relevante al respecto es la *provisión* de escolaridad, y que corresponde a los gobiernos y las instituciones educativas asegurar la oferta local de escuelas seguras y con materiales de enseñanza adecuados. “Acceso” también significa que las políticas educativas no crean barreras a la asistencia de los niños a la escuela. En su carácter de fundamento para el éxito, el acceso se califica como un factor de la oferta educativa.

El modelo de prosperidad educativa considera que la asistencia a la escuela y la posibilidad de que los estudiantes terminen niveles sucesivos de escolarización constituyen aspectos de la transición oportuna, uno de los resultados de prosperidad. Una medida de la asistencia a la escuela incorpora factores de demanda: si hay una provisión adecuada de enseñanza, no solo de recursos materiales sino de un entorno inclusivo, una enseñanza de calidad, tiempo de aprendizaje y apoyo familiar y comunitario, la pregunta pertinente es: “¿En qué medida asisten los estudiantes a la escuela?” Esto depende de varios factores culturales, sociales, religiosos, políticos y económicos.

Una de las metas de los Objetivos de Desarrollo Sostenible (párrafo 1.4) es la siguiente: “De aquí a 2030, asegurar que todas las niñas y todos los niños terminen la enseñanza primaria y secundaria, que ha de ser



gratuita, equitativa y de calidad y producir resultados de aprendizaje pertinentes y efectivos” (Naciones Unidas, 2015). Esta meta implica el acceso, que en el presente informe se refiere a la libre disposición de la enseñanza (recursos materiales), a la calidad escolar (escuelas seguras e inclusivas y enseñanza de calidad) y el rendimiento académico, un resultado de prosperidad. En el seguimiento del progreso hacia esta meta es importante distinguir entre equidad e igualdad, ya que así podemos preguntarnos primero acerca de la equidad de la provisión y después si ésta se relaciona con la igualdad en el desempeño y la transición oportuna. El modelo CREATE de Lewin (2015) es útil en este sentido porque subraya la importancia de asegurar una escolarización propicia para el desempeño académico y la transición oportuna, y de las “zonas de exclusión” asociadas a los estudiantes que pierden el rumbo durante su escolarización.

Subpoblaciones. Las subpoblaciones de interés pueden variar según el país y los resultados o la fase de desarrollo. Por ejemplo, las metas de la Agenda 2030 para el Desarrollo Sostenible (UNESCO, Instituto de Estadística, 2017) se centran en las desigualdades e inequidades relacionadas con el género. Los factores enumerados en la Figura 4 –género, discapacidad, condición de inmigrante, idioma que hablan en casa e idioma que hablan en la escuela y pobreza– fueron identificados por los países participantes en PISA para el Desarrollo como los más importantes en sus contextos.

Importancia del nivel socioeconómico

El NSE de la familia y la pobreza desempeñan un papel clave en el desarrollo del niño durante todo su ciclo de vida, ya que se relacionan con los fundamentos familiares y comunitarios en cada fase. El NSE y la pobreza también determinan si el niño tiene acceso a fundamentos institucionales sólidos, especialmente después de los 5 años, cuando la selección institucional es más prominente. La **Figura 5** presenta la relación entre el porcentaje de estudiantes con puntuaciones en lectura del nivel 2 o inferior en la prueba PISA 2015 y el ingreso nacional bruto (INB) per cápita, que demuestra claramente que los niveles de vulnerabilidad suben en gran medida cuando el INB desciende a menos de \$30.000.

El enfoque de ciclo vital asumido en el modelo de prosperidad educativa sugiere que muchos niños se muestran vulnerables a temprana edad, antes de empezar la escuela. Su potencial de convertirse en buenos lectores probablemente se hizo evidente cuando ingresaron a la primaria. En la mayoría de los países de ingreso medio y bajo, más del 60% del alumnado tiene habilidades lectoras deficientes a los 15 años; y en los países de alto ingreso, esa cifra se ubica entre el 30% y el 50%. Desconocemos la medida en que el potencial de desarrollo de habilidades de lectura y escritura se incorpora biológicamente durante la fase prenatal o las primeras fases de desarrollo y la preprimaria, pero los datos empíricos indican cada vez más que los primeros años son los más importantes. PISA para el Desarrollo incluye una medición exhaustiva del NSE, que se extiende hasta el extremo inferior de su rango a fin de identificar a los niños que viven en la pobreza extrema o severa.

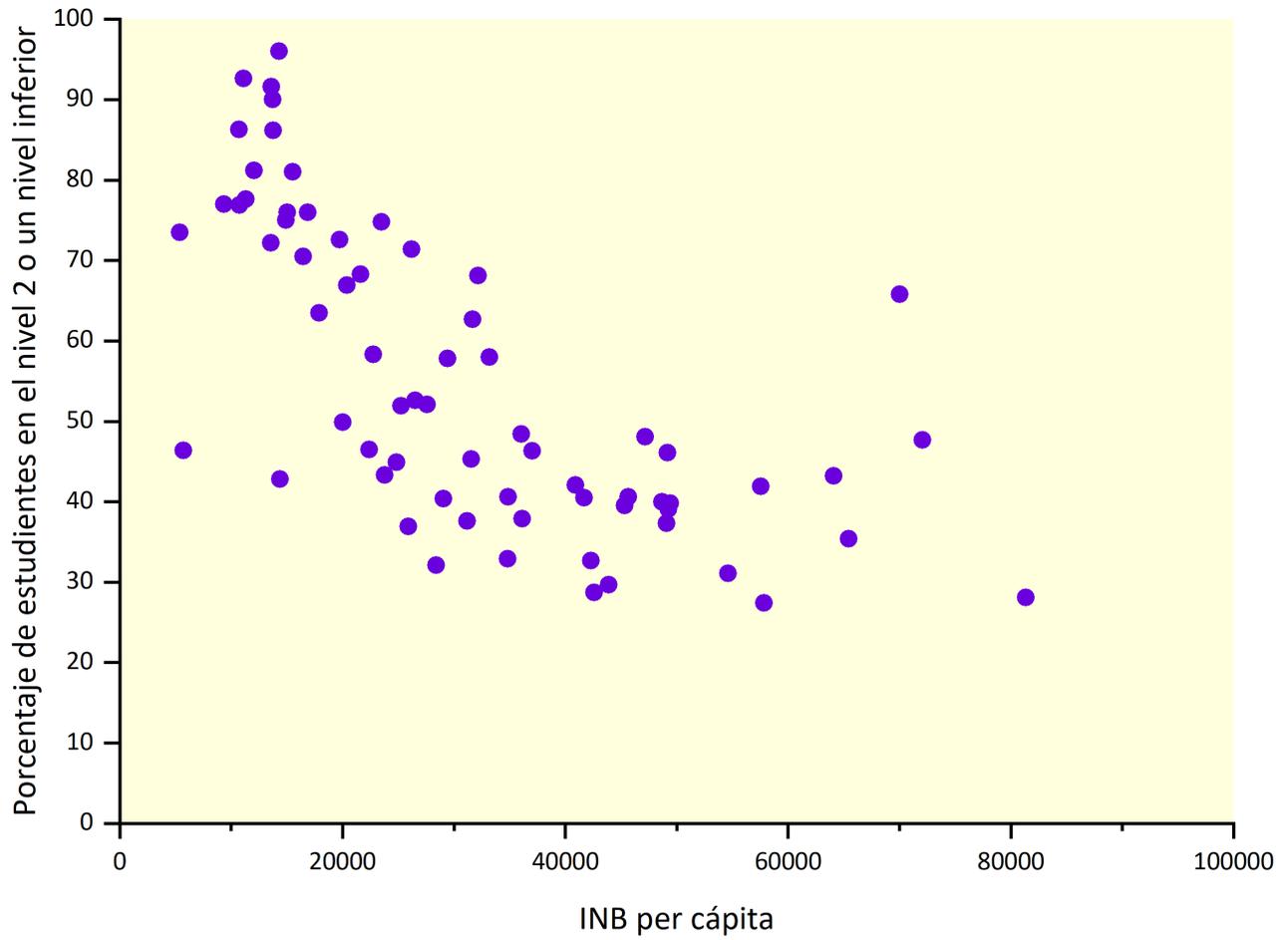
No se puede suponer que cuando los niños llegan a la preprimaria, su suerte está echada; más bien, en todos los países, cuando se preparan para iniciar la escuela primaria, ya cuentan con una amplia variedad de habilidades previas a la alfabetización. La fortaleza del sistema escolar y el apoyo familiar y comunitario



influyen en el desarrollo subsecuente de todos los niños y todas las niñas. La mayoría de ellos pueden aprender a leer si reciben enseñanza de calidad en un entorno inclusivo y con un tiempo de aprendizaje adecuado.

Se puede decir que todas las fases del marco de prosperidad educativa son “críticas” para un desarrollo adecuado de las habilidades de lectura y escritura. Sin embargo, para dar seguimiento del éxito de un sistema escolar con objeto de formular la política educativa, hay dos transiciones especialmente importantes a considerar. El ingreso a la primaria baja, que en la mayoría de los sistemas ocurre más o menos a la edad de seis años, marca un cambio del aprendizaje en casa, o de un programa de educación y guardería en la primera infancia, o de preprimaria al primer grado escolar, que generalmente es un entorno de aprendizaje estructurado con un plan de estudios más formal. La segunda transición crítica es de la primaria baja a la primaria alta, que implica el paso de aprender a leer a leer para aprender. En la siguiente sección se abordan en detalle estas dos transiciones.

Figura 5. Porcentaje de estudiantes de 15 años con habilidades lectoras del nivel 2 o inferiores (PISA 2015), en relación al ingreso nacional bruto (INB) (método Atlas, en dólares estadounidenses)



Fuente: PISA 2015 y Banco Mundial, 2018a.



II. Dos transiciones críticas para el desarrollo de las habilidades de lectura y escritura

Ingreso a la educación primaria baja

Los gobiernos reconocen cada vez más que las inversiones en programas para la primera infancia arrojan altos rendimientos en el desarrollo de los niños (Knudsen, Heckman, Cameron y Shonkoff, 2006). Puesto que el inicio de la escuela primaria es una transición clave, la medición de las habilidades de desarrollo de los niños a medida que hacen esa transición es valiosa por dos razones: sirve como indicador retrospectivo del éxito de la sociedad en la preparación de los niños para la escuela; y sirve como indicador prospectivo para determinar cuáles niños podrían necesitar apoyo adicional durante la escuela primaria.

El término “preparación para la escuela” suele emplearse para indicar el grado en que un niño está preparado para tener éxito en un entorno más estructurado (UNICEF, 2012). Una definición más amplia incorpora las habilidades de los niños y el papel de las familias y las comunidades (Rhode Island Kids Count, 2005):

Familias preparadas + Comunidades preparadas + Servicios preparados + Escuelas preparadas = Niños preparados para la escuela

Un antiguo modelo de preparación para la escuela, derivado del trabajo del U.S. National Education Goals Panel, abarca cinco campos: bienestar físico y desarrollo motor, desarrollo social y emocional, métodos de aprendizaje, desarrollo lingüístico, y cognición y conocimientos generales (Barnett, Ayers y Francis, 2015; Kagan, Moore y Bredekamp, 1995; National Early Literacy Panel, 2008).

El vínculo entre las habilidades cognitivas y las lingüísticas a los 5 años con las habilidades lectoras a los 8 o 9 años está bien establecido (Deary, Strand, Smith y Fernandes, 2007; Duncan *et al.*, 2007; McClelland, Morrison y Holmes, 2000; Raver *et al.*, 2011; Rose, 2006). Las habilidades de expresión oral y la capacidad cognitiva son especialmente importantes (Scarborough, 2001). La atención y el autocontrol también son clave y aparentemente más importantes que los comportamientos de externalización problemáticos (Trezesniewski, Moffit, Caspi, Taylor y Maughan, 2006).

La Evaluación Infantil Temprana (EIT) se usa para identificar las habilidades de los niños de 3 a 6 años mientras se preparan para la transición a la educación formal (The Learning Bar, 2011). Incluye mediciones de los cinco campos establecidos por el National Education Goals Panel, así como dos herramientas complementarias que ayudan a los docentes a monitorear el desarrollo general de los niños antes y durante la transición a la educación formal. La EIT-Evaluación Directa (EIT-ED) es una evaluación interactiva basada en juegos a cargo de un evaluador externo capacitado. La EIT-Valoración del Maestro (EIT-VM) provee un marco sistemático que puede servir a los maestros para estructurar sus observaciones y evaluaciones



informales. Generalmente la utilizan los maestros de pre primaria para obtener información de carácter formativo pertinente para la enseñanza.

La EIT-VM se empleó en una evaluación nacional que se efectuó en Uruguay en 2017, tras un proceso de implementación de tres años que incluyó un estudio piloto para examinar las propiedades psicométricas de la evaluación; la adaptación de ésta al contexto cultural después de consultar con expertos en planes de estudios y maestros; la participación de los maestros en el proceso de implementación y uso conforme a las mejores prácticas; y revisiones gubernamentales para asegurar que el contenido de la evaluación fuera pertinente para el desarrollo de los niños en los jardines de infancia y contribuyera a mejorar las prácticas de enseñanza (López, 2016; Willms, 2018b).

La **Figura 6** muestra el rango de puntuaciones de toda la población de niños uruguayos de 4 y 5 años en 2017. La EIT-VM se trata como una prueba referida a criterios, que indica si los niños han dominado las tareas necesarias para un desempeño exitoso en la escuela primaria. El rango va del 0 al 3. Los niños con una puntuación de 2 o superior se sitúan en el nivel 1 (verde). Se considera que han completado las tareas en esa área de desarrollo. El nivel 2 (amarillo, que denota una puntuación de 1 o superior, pero inferior a 2) indica que el niño está experimentando una dificultad regular para llevar a cabo las tareas; y el nivel 3 (rojo, que denota una puntuación inferior a 1), indica que está experimentando una dificultad significativa. El diagrama de cajas muestra el rango de puntuaciones en cada una de las edades, que se expresan con incrementos de medio año. El cuartil inferior indica el rango del 50. percentil al 25o.; el segundo, el rango del 25o. percentil a la mediana; el tercero, el rango de la mediana al 75o. percentil; y el cuartil superior, el rango del 75o. percentil al 95o.

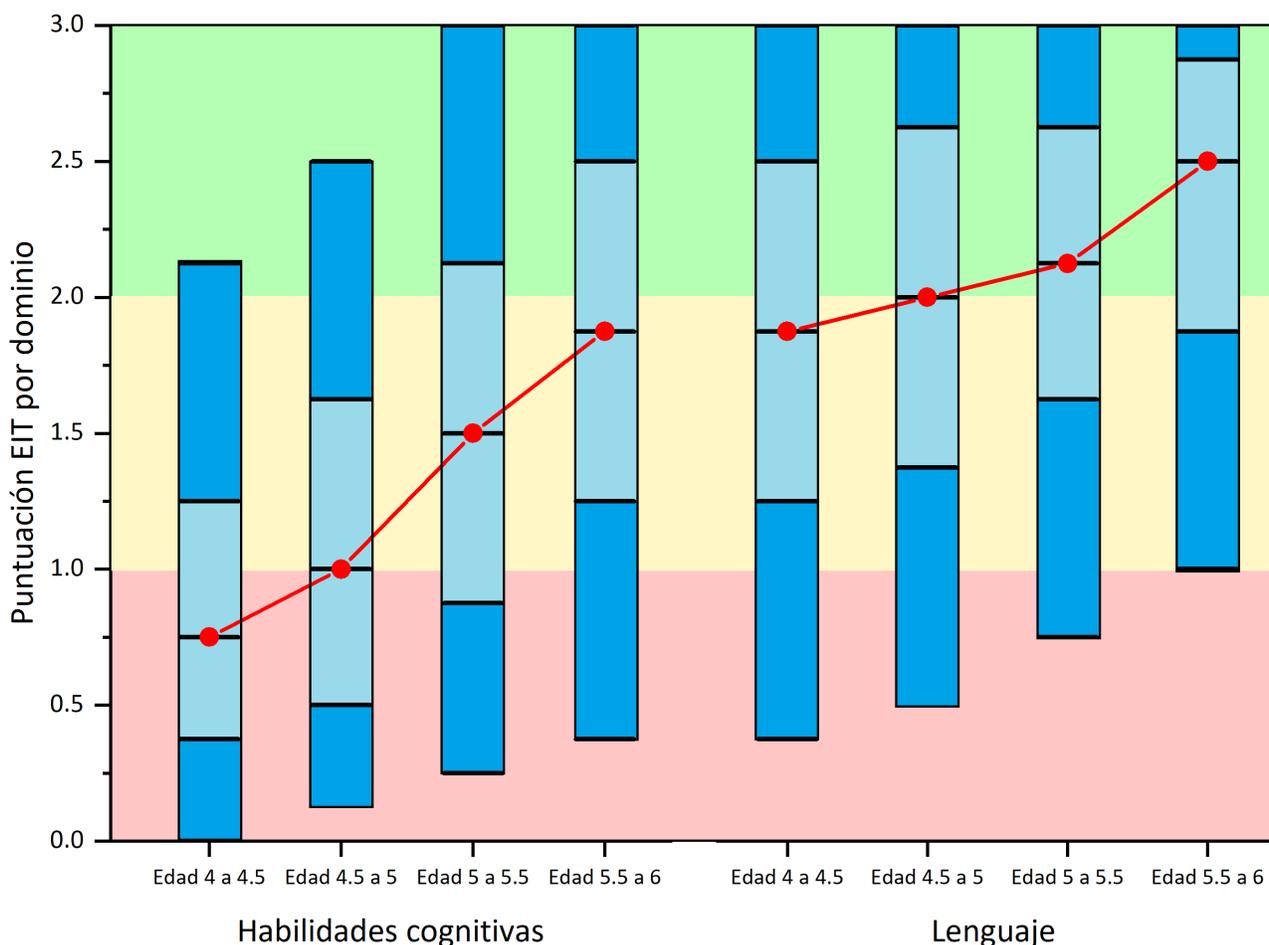
De lo anterior se desprende que la mayoría de los niños de 4 a 4.5 años tienen más habilidades lingüísticas que habilidades cognitivas. Sin embargo, las habilidades cognitivas se desarrollan con mayor rapidez. Estos resultados van de la mano con los de investigaciones recientes sobre el desarrollo del cerebro: la formación de sinapsis para las habilidades lingüísticas alcanza su punto máximo antes que la función cognitiva y a la edad de 5 años es más estable; la función cognitiva no se estabiliza hasta los 15 años, aproximadamente (Bhattacharjee, 2015).

Los análisis en cuestión indican que más de la mitad de los niños de 5 años son vulnerables en cuanto a habilidades cognitivas en el sentido de que están experimentando una dificultad regular o significativa: de los de 5 a 5.5 años, el 38% experimenta una dificultad regular y el 33% una dificultad significativa; mientras que de los de 5.5 a 6 años, el 34% experimenta una dificultad regular y el 49% una dificultad significativa. Si se considera a todos los niños de 5 años que están preparados para ingresar a la escuela primaria, los que se sitúan en el cuartil inferior tienen habilidades comparables con las de los niños de 4 a 4.5 años. En otras palabras, van más de un año atrás con respecto a sus pares en el desarrollo de habilidades cognitivas. En la siguiente subsección se aborda la importancia de las habilidades cognitivas para el desarrollo de las habilidades de decodificación, que constituye el factor más importante para determinar si los niños han adquirido una buena competencia lectora.





Figura 6. Puntuaciones de la Evaluación Infantil Temprana en habilidades cognitivas y lenguaje, realizada en Uruguay en 2017



Fuente: *The Learning Bar*, 2017.

De “aprender a leer” a “leer para aprender”

Una transición oportuna entre “aprender a leer” y “leer para aprender”, que en la mayoría de los casos tiene lugar entre los 8 y los 9 años, es esencial para el éxito académico, los logros educativos y el bienestar (Snow, Burns and Griffin, 1998). Durante los primeros años de enseñanza primaria los niños aprenden contenidos sobre las asignaturas y adquieren diversas habilidades; sin embargo, se pone énfasis en que aprendan a leer. Cuando pasan a la primaria alta, se da por sentado que pueden leer con fluidez y entender contenidos curriculares sobre las áreas de salud, estudios sociales y ciencias, entre otros. Se espera que aprendan el lenguaje de las asignaturas y que lo utilicen para pensar de manera crítica, resolver problemas y generar nuevos conocimientos. La necesidad de habilidades lectoras sólidas aumenta cuando llegan a la secundaria alta. Aquellos que carecen de las habilidades lectoras básicas se van quedando rezagados. Un hallazgo de



largo tiempo atrás es que la mayoría de los estudiantes a los que se les dificulta leer al final de la primaria siguen teniendo problemas en la secundaria baja y alta (Francis et. al., 1996; Warwick, 2005).

Aprender a leer es un proceso complejo que requiere conciencia fonológica, es decir, la capacidad de oír y manipular las estructuras de sonido del lenguaje; comprensión alfabética, que implica relacionar símbolos impresos con el lenguaje oral; y fluidez en el reconocimiento de las palabras utilizadas en los textos (Good, Simmons and Kame'enui, 2001; Storch and Whitehurst, 2002). El modelo de "concepción simple de la lectura", (Gough y Tunmer, 1986) ampliamente reconocido, postula que una adquisición adecuada de la lectura depende de dos componentes: la decodificación y la comprensión lingüística. La decodificación es la capacidad de reconocer eficientemente palabras conocidas y desconocidas. La capacidad de recuperar con precisión el código fonológico de una palabra escrita contribuye no solo al desarrollo de las habilidades lectoras del niño sino también al de su vocabulario y su comprensión de la lectura (Verhoeven, van Leeuwe y Vermeer, 2011). La comprensión lingüística es la capacidad de entender e interpretar el lenguaje oral y el escrito cuando forman parte de oraciones o de otro discurso. Para que un niño se convierta en un buen lector, no solo tiene que dominar la habilidad de decodificar con precisión las palabras escritas, sino también comprender el significado de las palabras y cómo se combinan en frases, oraciones y párrafos.

La habilidad de decodificar palabras es el "filtro crítico" durante la primaria. Para que los niños puedan entender qué están leyendo, tienen que ser capaces de identificar las palabras con precisión y eficiencia, y retener en su memoria de trabajo la información contenida en una frase u oración (Perfetti, Landi y Oakhill, 2005; Vellutino y Scanlon, 1987). El desarrollo de las habilidades de decodificación refuerza el desarrollo de las habilidades de comprensión; sin embargo, la mayoría de los niños que no aprenden a leer bien tienen dificultades con sus habilidades de decodificación (Storch y Whitehurst, 2002; Verhoeven, van Leeuwe y Vermeer, 2011). Los profesores eficaces elaboran un repertorio de estrategias para enseñar a los niños la conciencia fonémica y el reconocimiento de palabras; logran los mejores resultados cuando la enseñanza de habilidades de decodificación incluye oportunidades de leer textos conectados (Torgesen, Otaiba y Grek, 2005).

Confident Learners es un programa de alfabetización escolar integral que tiene por objeto mejorar las habilidades de lectoescritura de niños indígenas durante los primeros grados (The Learning Bar Inc., 2016). Está dirigido a educadores indígenas de Canadá y se elaboró en colaboración con ellos y con el apoyo del Indigenous Advisory Circle. Una de las características principales del programa es el "enfoque secuencial" de evaluación y enseñanza basado en la concepción simple de la lectura. Incluye un programa de aprendizaje profesional para maestros encaminado a incrementar sus conocimientos sobre el desarrollo de las habilidades de lectura y escritura de los niños, así como el uso de estrategias de enseñanza de alto rendimiento para desarrollar sus habilidades de decodificación y lingüísticas. El programa incluye 20 módulos o "pasos" para las habilidades de decodificación y 20 pasos para las habilidades lingüísticas; cada paso comprende 14 objetivos basados en habilidades. Estos objetivos se vinculan con una o varias actividades de aprendizaje divertidas, interesantes y pertinentes desde el punto de vista cultural. Los



maestros vigilan el progreso de los estudiantes por medio de evaluaciones periódicas, a fin de determinar su dominio de los objetivos de cada paso conforme avanzan en la secuencia.

Durante las fases iniciales de la creación de Confident Learners, tres expertos en alfabetización reunieron los planes de estudios de todas las provincias canadienses, algunos estados de los Estados Unidos, de Australia y el Reino Unido. Establecieron un modelo que describe las habilidades a nivel micro necesarias para decodificar palabras y desarrollar las habilidades lingüísticas requeridas para aprender a leer durante la primaria. El modelo arrojó más de 300 habilidades de codificación y 300 habilidades lingüísticas que abarcan el periodo entre la preprimaria y el tercer grado. Un hallazgo inesperado fue que, en la mayoría de los planes de estudios, más de la mitad de las habilidades de decodificación esenciales se enseñan desde el principio del primer grado hasta la primera mitad del segundo. El equipo de investigación se refiere a este fenómeno como la “montaña de la lectura”.

El problema que enfrentan los alumnos que empiezan la primaria sin las habilidades fundamentales previas a la alfabetización es que no pueden atravesar la montaña de la lectura. Los resultados de la Evaluación Infantil Temprana indican que, en muchas jurisdicciones y ciertas subpoblaciones, entre una cuarta parte y la mitad de los niños ingresan a la primaria con habilidades cognitivas y lingüísticas que sus pares aprendieron uno o varios años antes. Cuando en el primer año se enfrentan con un plan de estudios basado en grados, son incapaces de dominar las habilidades necesarias para pasar al siguiente grado. Las tres estrategias más prominentes para abordar este problema son la repetición de año, la clasificación de los niños como estudiantes con necesidades especiales y su separación en clases especiales, o “retrasar el fracaso”, que consiste en dejarlos pasar al segundo grado con la expectativa de que se pongan al corriente. Ninguna de estas estrategias sirve para incrementar la prevalencia de niños que hacen una transición exitosa de “aprender a leer” a “leer para aprender”.

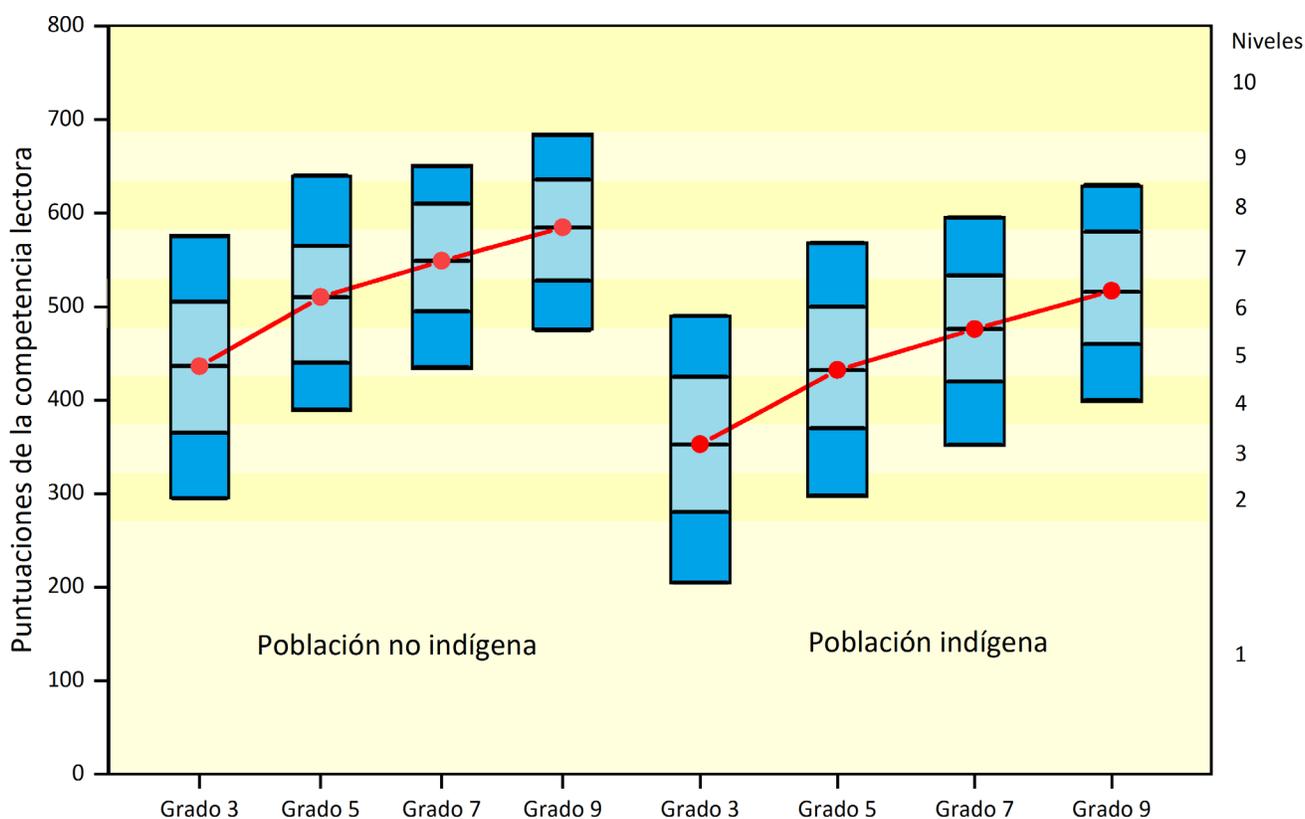
El fenómeno de la “montaña de la lectura” y la baja tasa de mejoramiento de las habilidades lectoras después de la escuela primaria son evidentes en los resultados del National Assessment Program of Literacy and Numeracy de Australia (NAPLAN; Australian Curriculum, Assessment and Reporting Authority, 2017). En la **Figura 7** se observa el rango de puntuaciones en lectura, que va del grado 3 (edad promedio de 8 años 3 meses) al grado 9 (edad promedio de 14 años 7 meses) de los estudiantes no indígenas e indígenas. El diagrama de cajas muestra el rango por grado. El cuartil inferior indica el rango del 5º percentil al 20º; el segundo, el del 20ºpercentil a la mediana; el tercero, el de la mediana al 80º percentil; y el cuartil superior, el del 80º percentil al 95º.

Tres aspectos de estos resultados son pertinentes para el desarrollo de las habilidades lectoras de los estudiantes. Primero, la tasa de crecimiento es lenta del grado 3 al grado 5 y más aún en lo sucesivo. La puntuación promedio de los estudiantes no indígenas en el grado 3 fue de 436.3 (DE=84.1) y en el grado 5, de 510.2 (DE=74.4). El crecimiento anual promedio en términos de tamaño de efecto es de 0.44. El crecimiento anual promedio del grado 5 al grado 7 (media=548.9, DE=66.7) fue de 0.26; y el del grado 7 al grado 9 (media=584.6, DE=636.7) fue de 0.24.



Segundo, las tasas de crecimiento correspondientes a los estudiantes indígenas fueron comparables: del grado 3 (media=352.8, DE=92.6) al grado 5 (media=432.1, SD=81.3, de 0.43; del grado 5 al grado 7 (media=476.0, SD=75.0) fue de 0.27 y del grado 7 al grado 9 (media=516.9, SD=73.0) fue de 0.27. La principal diferencia en las puntuaciones de los estudiantes indígenas comparados con los no indígenas estriba en las obtenidas en el grado 3. Esta brecha de logro es atribuible a las diferencias entre sus experiencias de vida antes de ingresar a la primaria o sus experiencias en la primaria baja. No se dispone de datos para determinar la medida en que esa brecha se hizo evidente durante su estancia en la preprimaria o en la que aumentó entre ese período y el grado 3. Comprender esta relación y la oferta educativa durante los años de primaria tiene implicaciones importantes para el monitoreo del sistema educativo.

Figura 7. Puntuaciones de la competencia lectora en Australia, basadas en el programa NAPLAN 2017



Fuente: ACARA, 2017.

Los estudiantes del grado 9 con puntuaciones de lectura en el quintil inferior mostraron habilidades comparables con los situados cerca de la mediana en el grado 5. Esto se aplica tanto a los estudiantes indígenas como a los no indígenas. Si los estudiantes clasificados en el quintil inferior en el grado 5 recibieran



una intervención que aumentara sus puntuaciones en una desviación estándar de un cuarto por grado, se ubicarían en la parte media de la distribución en el grado 9.

Los efectos de grado en PISA 2015. El incremento relativamente pequeño de las puntuaciones en lectura después del grado 9 en Australia y en los países de la OCDE en general también se hace evidente en PISA. La edad objetivo de los estudiantes que participan en PISA es de 15 años. En la mayoría de los países, de los estudiantes que no han repetido ningún grado, hay estudiantes de 15 años por lo menos en dos grados. Si excluimos a los estudiantes que han repetido un grado, los inscritos en el grado inmediatamente inferior al grado modal están ahí por su fecha de nacimiento; es decir, son representativos de los estudiantes de 15 años que están avanzando en la escuela según lo programado. En Australia, por ejemplo, el 9% de los estudiantes que no habían repetido un grado estaban en el grado 9, el 76% en el grado 10 (el grado modal) y el 15% en el grado 11. Por lo tanto, es posible estimar un “efecto de grado” en las puntuaciones de lectura asociadas con un aumento de grado, del grado anterior al modal y de éste al inmediatamente superior.

En todos los países miembros de la OCDE, el efecto asociado a la matriculación de un estudiante en el grado modal en comparación con un grado inferior fue de 55 puntos; y el efecto asociado a su matrícula en el grado inmediatamente superior fue de 33 puntos. El efecto de grado global fue de 43 puntos. En los países no miembros de la OCDE, los efectos de grado fueron como sigue: 41 puntos por estar en el grado modal en vez de un grado más abajo, y 24 puntos por estar en el grado superior. El efecto de grado global fue de 29 puntos. En promedio, el efecto de grado global para todos los países participantes en PISA 2015 fue de 36 puntos².

En los resultados de PISA, la puntuación se traduce directamente en tamaños de efecto: 100 puntos representan un tamaño de efecto de 1.0. También podemos considerar que una diferencia de 43 puntos representa un año de escolaridad, o que 4.3 puntos representan un mes de escolaridad en un ciclo académico de 10 meses.

Aprender a leer en idiomas distintos al inglés. La tasa a la que los niños adquieren competencia para decodificar palabras depende de la profundidad ortográfica del idioma. El inglés y el francés tienen una ortografía relativamente profunda; esto significa que la relación entre las letras y los sonidos (fonemas) es

² Para este análisis se utilizó la técnica de modelos lineales jerárquicos (HLM), con estudiantes anidados dentro de los países. Así, el efecto estimado es un efecto promedio entre todos los países. El análisis reveló una variación considerable del efecto de grado entre los países, ya que resultó inferior en 5 puntos en cuatro de ellos y mayor de 65 puntos en ocho. En el caso de Australia, los efectos de grado PISA se situaron en 16 puntos en el grado 9 y en 12 puntos en el periodo del grado 10 al grado 11. Estos resultados son congruentes con los de NAPLAN antes mencionados, que revelan tasas de crecimiento descendentes después del grado 5. (En el modelo HLM también se utilizó la edad de los estudiantes como variable y los resultados indicaron que los efectos de edad eran insignificantes; los estudiantes de menor edad para su cohorte quedaron un punto abajo de aquellos de mayor edad para su cohorte.)



menos congruente que en los idiomas con una ortografía superficial, como el finés, el italiano y el español. Por consiguiente, aprender a leer en inglés es más lento que hacerlo en un idioma con una ortografía superficial, como el español (Caravolas et. al., 2013). La decodificación eficaz de las palabras es un requisito previo para leer en todos los idiomas alfabéticos; pero en el caso de los que tienen una ortografía superficial, la comprensión lingüística es más importante para el desarrollo temprano de la competencia lectora. Los estudios sobre el desarrollo temprano de la competencia lectora entre los estudiantes de habla hispana, que se basan en la concepción simple de la lectura, confirman que la comprensión lingüística es más importante que la decodificación (Polo, Araujo y Salceda, 2017; Ripoll, Aguado y Castilla-Earls, 2014).

Estos resultados tienen implicaciones significativas para las intervenciones o los programas de lectura destinadas a países de ingreso bajo y medio. En muchos países africanos, por ejemplo, se enseña a los niños a leer en el idioma nacional oficial que puede o no ser su primera lengua; o se les enseña en su primera lengua durante los años de primaria y posteriormente, después de la primaria baja, en el idioma nacional. Si el idioma inicial de enseñanza es la primera lengua de los niños, una intervención de lectura como Confident Learners podría privilegiar las habilidades de decodificación sobre las de comprensión lingüística durante la primaria. La altura de la “montaña de la lectura” difiere según la profundidad ortográfica del idioma. Si el idioma de la enseñanza no es la primera lengua de los niños, la intervención tendría que poner mayor énfasis en las habilidades de desarrollo lingüístico porque es menos probable que se refuercen en el hogar o en la comunidad local. Un mayor énfasis en las habilidades de decodificación que en las habilidades de comprensión lingüística es clave para las políticas educativas de Sudáfrica. Cuando se abolió el apartheid en 1994, la Constitución confirió un estatus oficial a nueve lenguas indígenas, junto con el afrikaans y el inglés (Heugh, 2013). Ahora, en la mayoría de los casos, a los niños se les enseña a leer en su primera lengua hasta el tercer grado y después en inglés (Manyike, 2012). Esto puede presentar dificultades para la mayoría de los estudiantes que aún no han llegado a dominar las habilidades de decodificación necesarias para aprender a leer en inglés.

La pobreza también es un factor importante para determinar la idoneidad de una intervención oportuna de preparación para la alfabetización y la enseñanza de la lectura durante la primaria baja. Ya se ha establecido claramente el vínculo entre la pobreza y la preparación de los niños para la alfabetización (Banco Mundial, 2018b). Los niños que viven en la pobreza tienen más probabilidades de sufrir deficiencia de nutrientes, enfermedades infecciosas y ambientes tóxicos durante las dos primeras fases de desarrollo, desde la concepción hasta los 2 años, y las consecuencias de estos factores de riesgo pueden incorporarse biológicamente (Black *et al.*, 2008). En el proceso de maduración de los primeros años, los niños también corren el riesgo de recibir menores estímulos, orientación y apoyo (Black *et al.*, 2017). Los resultados de Uruguay en la Evaluación Infantil Temprana antes descrita indican que cuando los niños entran al primer grado, más de la cuarta parte de ellos están atrasados más de un año completo con respecto a sus pares en cuanto al desarrollo de habilidades cognitivas. El porcentaje de tales niños vulnerables es mayor en los Departamentos que tienen un porcentaje más alto de niños que viven en la pobreza. Por lo tanto, en el diseño de una intervención de lectura para niños uruguayos, probablemente habría que dar más



importancia al desarrollo de las habilidades de decodificación en los Departamentos con altos niveles de pobreza; y al desarrollo de habilidades de comprensión lingüística, en aquéllos con mayor riqueza relativa.



III. Políticas sobre resultados e igualdad de los estudiantes

El objetivo general de la política educativa es asegurar que todos los niños alcancen su potencial, es decir, que experimenten la *prosperidad*. Las políticas establecen un curso de acción que implica establecer objetivos e identificar estrategias para lograrlos. Los datos obtenidos a partir de un monitoreo o seguimiento pueden servir para fundamentar el proceso de política, ya que caracterizan el desempeño del sistema escolar. La pregunta básica más importante es: “¿Cómo vamos?”. Por lo general, los indicadores sociales adquieren significado en una de tres formas: mediante comparaciones con ciertas normas, por ejemplo, un promedio nacional o internacional; mediante comparaciones entre jurisdicciones, por ejemplo, de escuelas de una provincia o estado; y mediante un análisis temporal de tendencias. Entre otras preguntas, podemos plantearnos las siguientes: “¿Cómo vamos en comparación con las normas internacionales?”, “¿Cómo vamos en comparación con otros países que tienen niveles de recursos similares?”, “¿Cómo estamos mejorando nuestros resultados con el tiempo?”. Las respuestas a estas preguntas inevitablemente llevan a más preguntas: “¿Quiénes son vulnerables –existen desigualdades en los resultados?” y “¿Dónde se encuentran nuestros niños más vulnerables?”. Con las respuestas a estas preguntas, los administradores de la educación pueden establecer objetivos realistas y medibles de la mejora de los resultados de los estudiantes y la reducción de las desigualdades. Esta sección aborda el tema de los resultados de prosperidad. Su propósito es examinar los tipos de indicadores más útiles, así como los análisis y las herramientas de información que pueden aportar datos para formular las políticas en materia de resultados de los estudiantes e igualdad. En la Sección V se explica cómo puede usarse esta información para establecer objetivos.

¿Cómo vamos?

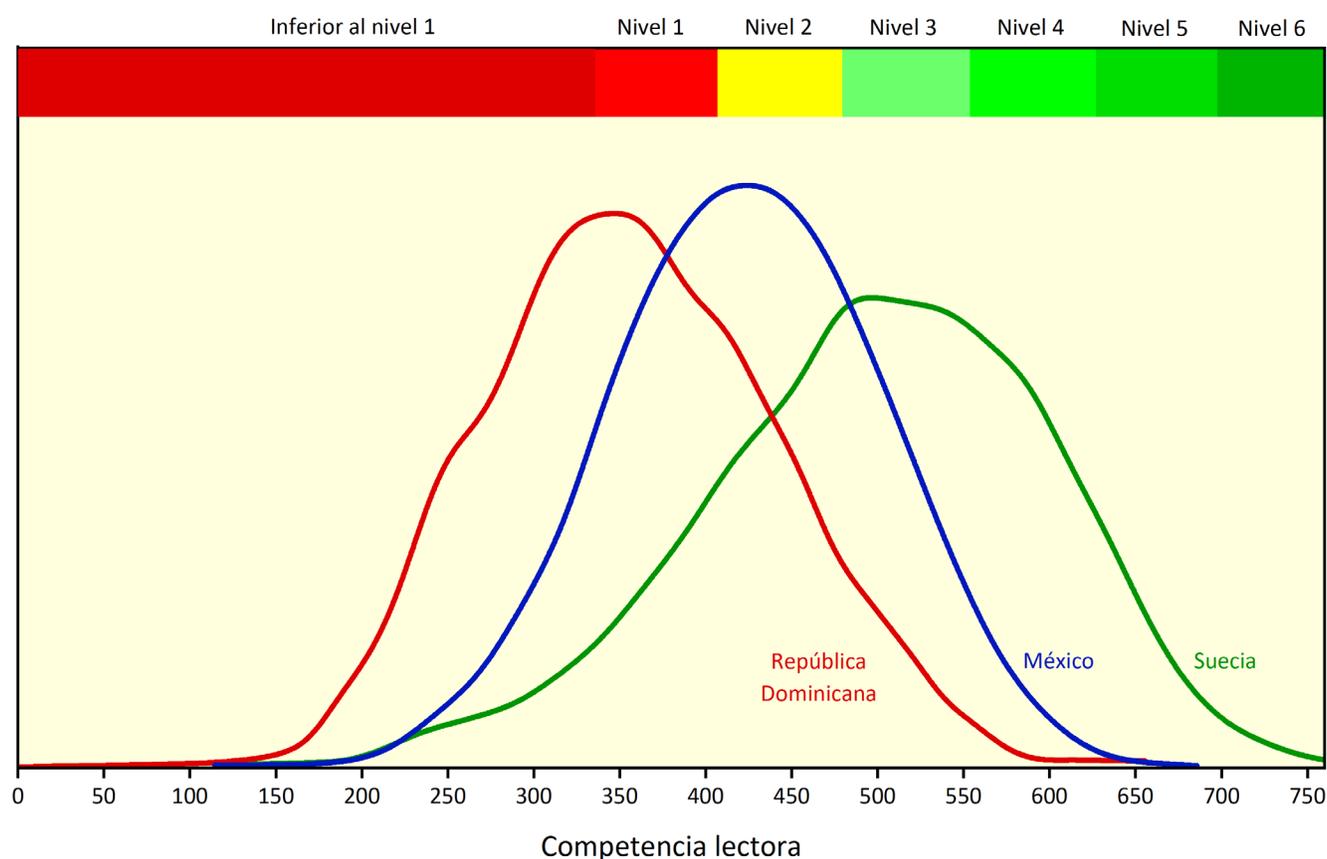
Énfasis en los resultados. El mejor punto de partida es el alcance que logran los estudiantes en su trayectoria, es decir, qué tan lejos avanzan en su escolaridad. La primera pregunta para los países de ingreso bajo y medio es si los estudiantes inician o cursan la primaria. La mayoría de los niños se inscriben en la primaria y las cohortes sucesivas por lo general los superan en términos de avance en los estudios (Banco Mundial, 2018b). Sin embargo, algunos nunca se inscriben en la escuela, en algunos casos porque viven en zonas remotas o padecen una discapacidad. En muchos países de bajo ingreso las tasas de inscripción de las niñas son mucho más bajas que las de los niños (Banco Mundial, 2018b). Así, dos marcadores clave son simplemente el porcentaje de estudiantes que se inscriben en la preprimaria y el de los que se inscriben en la primaria. En esta fase también son importantes las mediciones de las habilidades de preparación para la alfabetización, porque sirven como indicadores retrospectivos del desarrollo de los niños desde la concepción y como indicadores prospectivos de los recursos requeridos durante la primaria baja. Una vez que se dispone de estos indicadores, es posible determinar qué subpoblaciones son vulnerables y dónde se encuentran los niños vulnerables.

El siguiente indicador clave de una transición oportuna a lo largo de la trayectoria escolar es el porcentaje de niños que pasan de la primaria baja a la primaria alta. Cuando hacen esa transición, también se tienen que medir sus habilidades en lectura, escritura y matemáticas, así como su salud, bienestar y participación.



Además, se requiere un indicador del porcentaje de estudiantes que hacen la transición; sin él, los indicadores de dichas habilidades tienen poco valor porque probablemente corresponden a un determinado grupo de estudiantes. El mismo principio se aplica a la transición de la primaria alta a la secundaria baja, y de la secundaria baja a la alta. Las puntuaciones derivadas de PIRLS, TIMSS, PISA u otras evaluaciones internacionales tienen poca utilidad para fines comparativos a menos que los resultados se consideren junto con los indicadores del porcentaje de niños que no asisten a la escuela.

Figura 8. Distribución de la competencia lectora a los 15 años en República Dominicana, México y Suecia



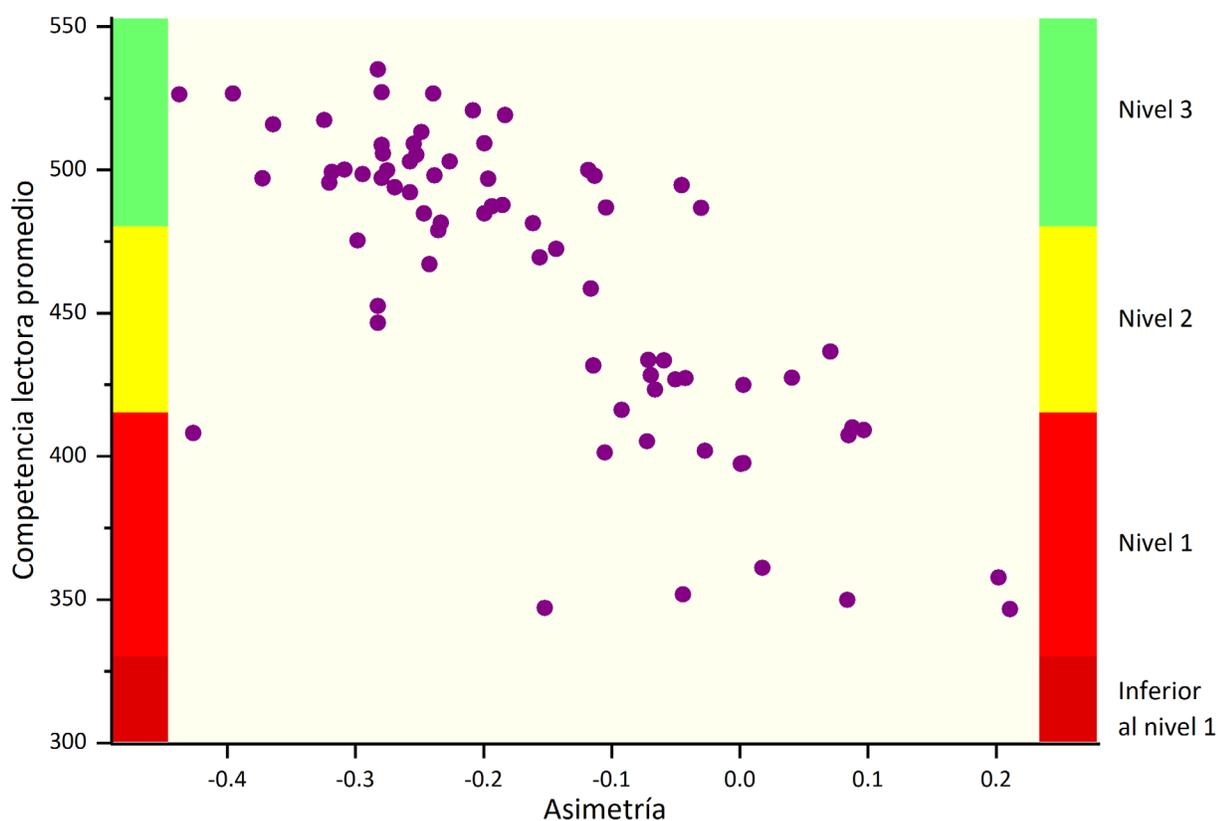
Fuente: PISA 2015.

El propósito de las mediciones continuas mediante pruebas o exámenes, por ejemplo, es describir la distribución de las puntuaciones con un conjunto pequeño de estadísticas resumidas que pueden ser útiles para dar seguimiento al progreso de un año a otro, o para calibrar el desempeño en comparación con una norma o con otras jurisdicciones. En la **Figura 8** se presenta la distribución de las puntuaciones para República Dominicana, México y Suecia. Los dos marcadores que se usan más comúnmente para vigilar el progreso son la media y la mediana; pero también son útiles otras tres estadísticas: la desviación estándar,



la asimetría y el porcentaje de estudiantes “vulnerables”. La desviación estándar es un indicador del grado de variación o dispersión de un conjunto de puntuaciones. El rango entre cuartiles –del 25º percentil al 75º– antes descrito con referencia a las puntuaciones EIT y NAPLAN también puede servir como indicador en este sentido. La asimetría indica el grado de desigualdad de la distribución. Las distribuciones sesgadas negativamente indican puntuaciones bajas que se extienden en sentido descendente más allá de la media, mucho más de lo que las puntuaciones altas se extienden por encima de ella; lo inverso se aplica a las distribuciones sesgadas positivamente.

Figura 9. Relación entre la media de la competencia lectora y la asimetría



Fuente: PISA 2015.

La asimetría es importante porque conforme los sistemas escolares se desarrollan, no cambia por completo la distribución de las puntuaciones, que pueden estar sesgadas positivamente o incluso ser simétricas (es decir, no sesgadas); más bien, tienden a cambiar las puntuaciones de los estudiantes más capaces, dejando atrás a una subpoblación vulnerable. En el caso de los tres países a los que alude la Figura 8, las puntuaciones de República Dominicana están sesgadas positivamente (sesgo = 0.20), las de México son casi simétricas (sesgo = -0.07), y las de Suecia están sesgadas negativamente (sesgo = -0.31). La **Figura 9** muestra la relación entre las puntuaciones medias en lectura y la asimetría de la distribución en los países que participaron en PISA 2015. La correlación entre las puntuaciones medias de lectura y la desviación estándar



es muy pequeña, de 0.12, pero la correlación entre ellas y la asimetría es de -0.74. Jordania es un valor atípico en este análisis porque presenta una puntuación media baja (408) y un sesgo negativo (-0.43). Cuando se excluye, la correlación entre las puntuaciones medias y el sesgo es de -0.80. La mayoría de los países con puntuaciones bajas por lo general tienen un ingreso bajo o medio (véase la Figura 5) y registran tasas de transición inferiores en la secundaria baja. Si se obtuvieran las puntuaciones de lectura de los estudiantes que no asisten a la escuela, la relación negativa entre las puntuaciones medias y la asimetría sin duda sería más fuerte.

Énfasis en la vulnerabilidad. En el seguimiento del desarrollo de los niños, la alternativa de poner énfasis en sus fortalezas o en sus déficits ha sido un tema de debate (Ben-Arieh y Frønes, 2007).

La literatura sobre desarrollo infantil hace hincapié sobre todo en los déficits, por ejemplo, los problemas de comportamiento y de salud. Los esfuerzos de seguimiento de la efectividad de las escuelas y los sistemas escolares han enfatizado principalmente las fortalezas de los niños con mediciones del desempeño académico, aunque a veces ponen el acento en las fallas de lectura o en la “deserción escolar”.

La alternativa de utilizar medidas dicotómicas o continuas está estrechamente relacionada con el tema de debate antes mencionado. Para los efectos de las políticas, a menudo se prefieren las medidas dicotómicas porque llaman la atención sobre las cuestiones sociales y económicas que afectan a los niños vulnerables. Además, son más fáciles de entender para la gente. Por ejemplo, la mayoría podría interpretar mejor el enunciado “La prevalencia de jóvenes que sufren una depresión aumentó del 8% al 10%”, que el enunciado “La puntuación media de la depresión aumentó de 8.5 a 8.9 en una escala de 10 puntos”.

El enfoque de la prosperidad educativa favorece las medidas basadas en las fortalezas, puesto que el objetivo general es brindar a todos los niños la oportunidad de prosperar. A los responsables de la formulación de políticas puede resultarles más fácil recabar apoyo público para incrementar la prevalencia de niños en edad escolar con una buena condición física que reducir la obesidad infantil. Los objetivos de la Agenda 2030 para el Desarrollo Sostenible (Instituto de Estadística de la UNESCO, 2017) están redactados en términos positivos. Sin embargo, algunos resultados, tales como problemas de salud física o mental, se tratan mejor como déficits. Sufrir ansiedad o depresión no es lo mismo que obtener una puntuación baja en un índice de felicidad. A los responsables de la formulación de políticas también puede resultarles más fácil centrarse en la prevalencia de niños con puntuaciones de lectura bajas que en la de niños que han logrado traspasar cierto umbral.

Generalmente, las medidas continuas son preferibles a las dicotómicas porque abarcan más información. No obstante, es posible que un país logre reducir la prevalencia de niños con puntuaciones de lectura bajas y que esto no sea evidente en los cambios de las puntuaciones medias. Además, en muchos de los países que participan en evaluaciones internacionales, las puntuaciones de un número considerable de estudiantes se sitúan en el nivel mínimo de la prueba o cerca del mismo. Incluso, la “verdadera puntuación” de esos estudiantes puede ser inferior a su puntuación estimada de la prueba (Nonoyama-Tarumi y Willms, 2010). En este caso, es preferible centrarse en la vulnerabilidad. Algunas medidas tienen puntos de corte bien



establecidos para definir la vulnerabilidad, como un bajo peso al nacer o la obesidad infantil. Los estudios internacionales sobre educación incluyen puntos de corte para parámetros de referencia o niveles de desempeño. La determinación de la idoneidad de un punto de corte depende de los resultados y de la distribución de la población que se están considerando.

Los datos de PISA 2015 indican que la puntuación media de México en lectura es de 423. La desviación estándar es de 78 y el sesgo, de -0.07. El porcentaje de estudiantes considerados vulnerables, aquellos que no alcanzaron al menos el nivel 3, es del 76.0%. El diseño de las muestras de los principales estudios internacionales generalmente implica primero el muestreo de las escuelas y después el de los estudiantes. Además, las muestras de la mayoría de los países se estratifican con el objeto de asegurar que estén representados adecuadamente determinados sectores o regiones. Para estos estudios se utiliza un diseño de prueba matricial, en la que los estudiantes llevan a cabo solo una parte de las tareas en evaluación. Por consiguiente, para considerar el diseño de la prueba, los cálculos estadísticos requieren el uso de un conjunto de *ponderaciones replicadas*; y cuando se refieren a las puntuaciones de la prueba, también se necesita usar *valores plausibles*. El Apéndice 2 presenta las técnicas de cálculo estadístico y describe brevemente las ponderaciones replicadas y los valores plausibles.

¿Quiénes son vulnerables?

Gradientes socioeconómicos. Un *gradiente socioeconómico* o “barra de aprendizaje” describe la relación entre un resultado social y el nivel socioeconómico de los individuos bajo estudio en una jurisdicción específica, tal como una escuela, una comunidad, una provincia o un estado, o un país (Willms, 2003a; 2006). El resultado puede ser de carácter social, por ejemplo, el estado de salud, las puntuaciones de exámenes o el máximo nivel de estudios de los niños. El NSE generalmente se refiere a la posición relativa de una familia o una persona en una estructura social jerárquica debido a su acceso a la riqueza, el prestigio y el poder o a su control sobre los mismos (Mueller y Parcel, 1981). La medición del NSE es importante para el monitoreo de la educación porque facilita un contexto para establecer objetivos asequibles y evaluar el progreso hacia su logro. Además, constituye un método para evaluar la igualdad de los resultados y la equidad en la provisión; considerar el potencial de estrategias encaminadas al mejoramiento de los resultados de las estudiantes y la reducción de las desigualdades; y fortalecer la validez de los estudios de investigación (Willms y Tramonte, 2018).

La **Figura 10** muestra los gradientes socioeconómicos para las puntuaciones de lectura de los estudiantes de 15 años en México, según los datos de PISA 2015. El NSE se evalúa con la medida compuesta denominada “nivel socioeconómico y cultural” que se utiliza en PISA. Un gradiente socioeconómico consta de tres componentes: el nivel, la pendiente y la robustez de la relación entre los resultados y el NSE.

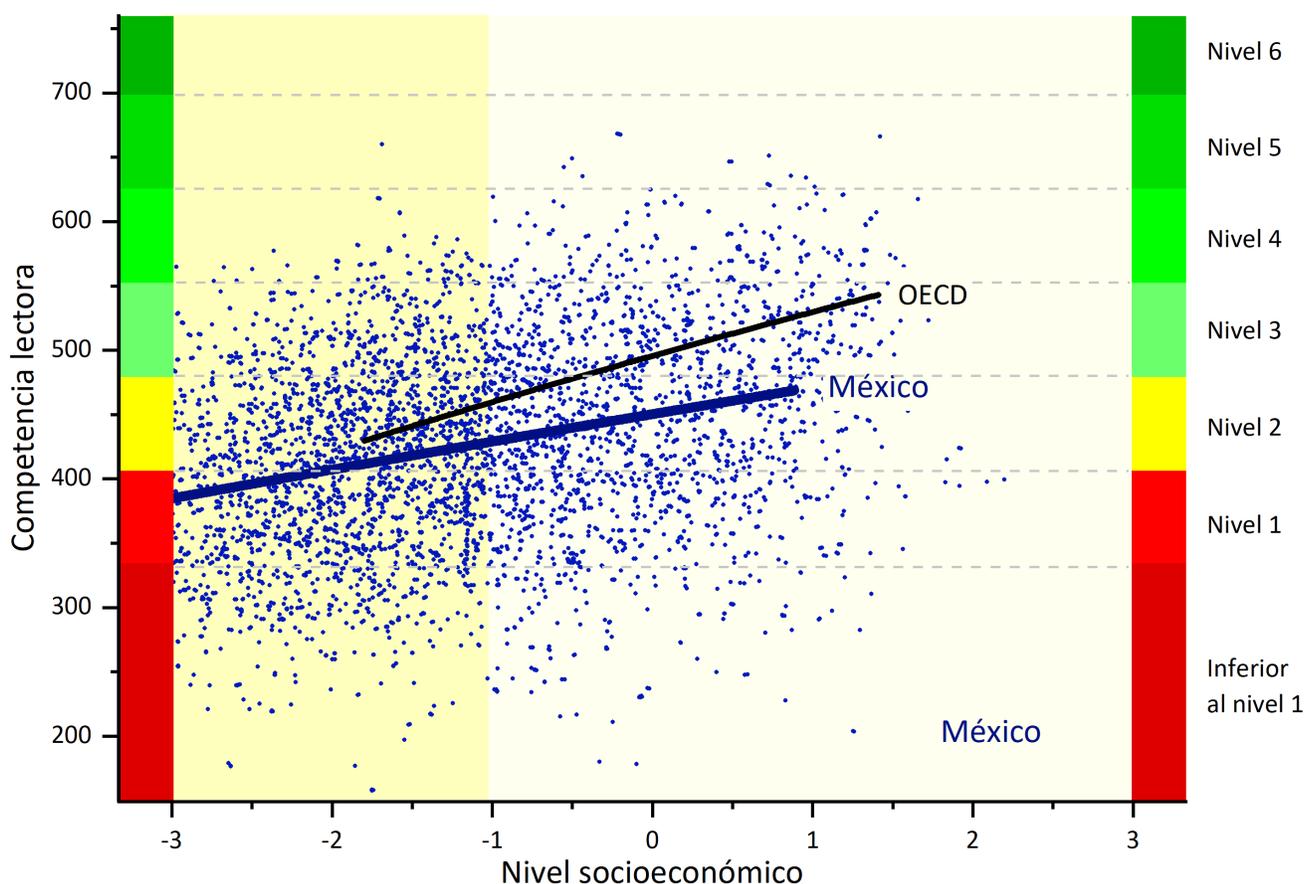
- a. El *nivel* del gradiente se define como la puntuación prevista en la medición de los resultados de un individuo con un NSE medio. El nivel del gradiente de un país (o una provincia, un estado o una escuela) es un indicador de su desempeño promedio, después de considerar el nivel socioeconómico de los estudiantes. El nivel del gradiente mexicano es de 450.3. Para este análisis, la medición del NSE



se ajustó de manera que el cero representara la media de la OCDE. Por lo tanto, el nivel se refiere a la puntuación prevista de un estudiante con un NSE comparable al de un estudiante hipotético con un NSE promedio igual a la media de la OCDE.

- b. La *pendiente* del gradiente indica el grado de desigualdad atribuible al NSE. Una mayor inclinación de los gradientes indica un mayor impacto del NSE en el desempeño de los estudiantes –es decir, más desigualdad–, en tanto que una mayor graduación de los gradientes indica un menor impacto del NSE –es decir, menos desigualdad. La pendiente de la gradiente mexicana es de 21.2 (en el centro de los datos), lo que indica que la competencia lectora esperada aumenta 21.2 puntos con cada incremento de una desviación estándar del NSE. La gradiente mexicana es ligeramente curvilínea, y su pendiente se reduce un poco cuando sube el NSE. Sin embargo, el coeficiente del componente curvilíneo no es significativo desde el punto de vista estadístico.
- c. La *fuerza* del gradiente se refiere a la proporción de la variación del resultado social atribuible al NSE. Una relación muy robusta sugiere que una gran parte de la variación en la medida de los resultados se debe al NSE; y una relación débil, que solo una parte de la variación se debe a este indicador. La medida que se utiliza más comúnmente para evaluar la robustez de la relación es la R cuadrada, que para México es de 0.12.

Figura 10. Gradiente socioeconómico de la competencia lectora en México



Fuente: PISA 2015.

La línea de gradiente que corresponde a México, en color azul, va del 5° al 95° percentil de puntuaciones del NSE. Para México, los percentiles 5° y 95° son -3.08 y 0.88, respectivamente. Por ende, el 90% de los estudiantes mexicanos caen dentro de este rango; y, en promedio, tienen un NSE más bajo que el de los estudiantes de otros países. La gráfica también muestra la competencia lectora y el NSE de una muestra representativa de México conformada por 5.000 estudiantes, que están comprendidos dentro de los pequeños puntos azules que aparecen abajo y arriba de la línea de gradiente. Los puntos indican que hay una variación considerable de la competencia lectora en todos los niveles socioeconómicos. La línea de gradiente para todos los países miembros de la OCDE se muestra en negro. Los percentiles 5° y 95° para los estudiantes de todos estos países son -1.78 y 1.40, respectivamente.

El término “barra de aprendizaje” se emplea metafóricamente para denotar el gradiente socioeconómico. La pregunta central para la mayoría de las escuelas y países es: “¿Cómo podemos subir y nivelar la barra de aprendizaje?”. El mejoramiento del desempeño académico y la reducción de las desigualdades entre los estudiantes que provienen de distintos medios socioeconómicos se pueden lograr de diversas maneras. El mejor enfoque posible depende de cuestiones sociales y políticas, pero también de la distribución de los



resultados de los estudiantes y del NSE en las escuelas y entre ellas, así como de la manera en que estos factores se relacionan e interactúan con los fundamentos para el éxito.

Rendimientos decrecientes. De acuerdo con la hipótesis de los rendimientos decrecientes, a mayor nivel socioeconómico, menores efectos sobre los resultados sociales vinculados a este indicador. Se podría predecir, por ejemplo, que arriba de cierto NSE habría un incremento ligero o nulo en la competencia lectora de los estudiantes asociada a este indicador. Parece ser que éste no es el caso de México, puesto que la línea de gradiente revela que la competencia lectora promedio aumenta de manera lineal cuando se eleva el NSE. La estimación de la línea de gradiente incluye un término cuadrático para el NSE, que es la estadística que se utiliza para determinar si los rendimientos de este indicador aumentan o disminuyen cuando se eleva. En el caso de México, el coeficiente del término lineal (es decir, la pendiente) fue de 21.2 y el término cuadrático, de -0.3. Lo anterior sugiere que a medida que el NSE sube, los rendimientos decrecen, pero en este caso no es algo estadísticamente significativo.

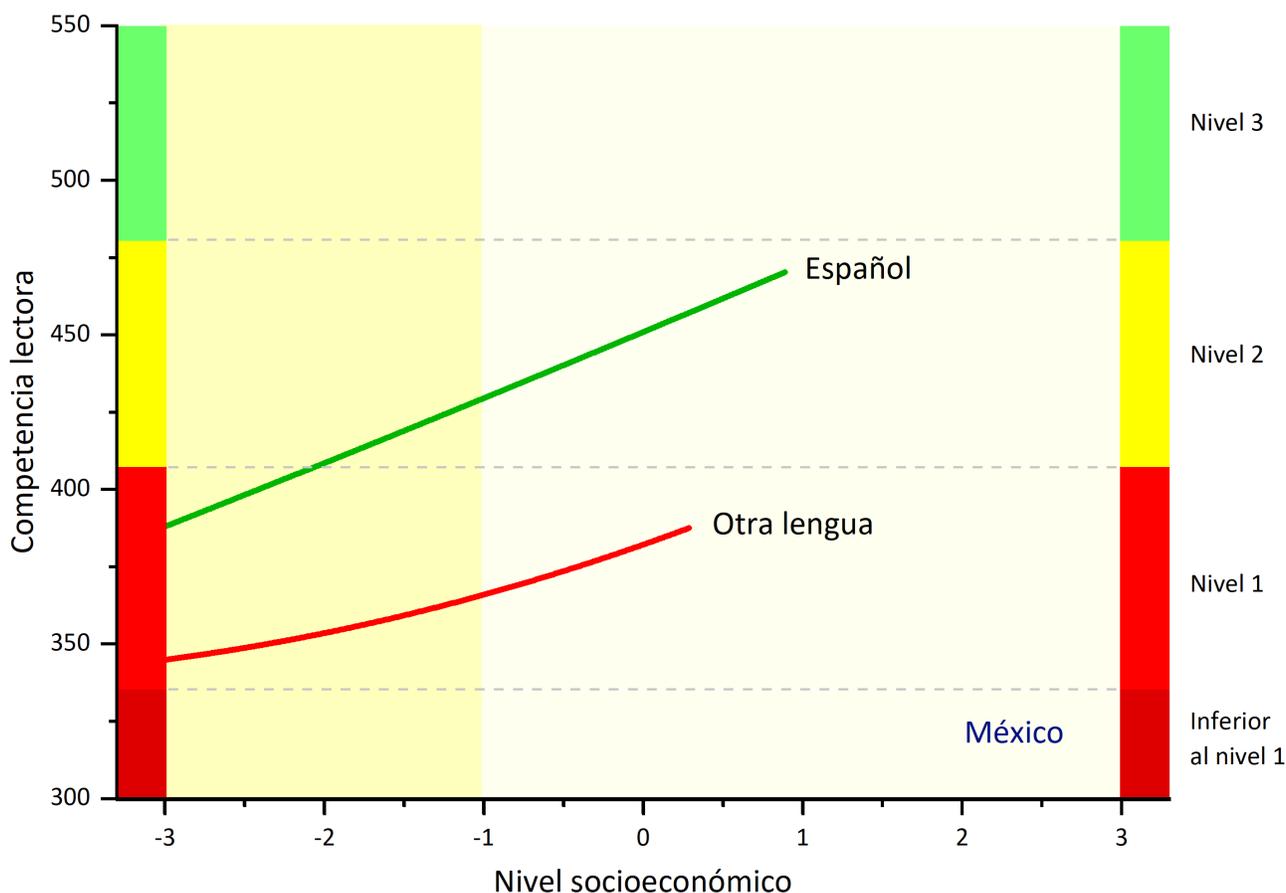
El Apéndice 2 describe las técnicas estadísticas para el cálculo de los gradientes socioeconómicos.

Willms y Somers (2001) observaron en varios países latinoamericanos rendimientos *crecientes* en los logros de lectura y matemáticas de los estudiantes de tercer y cuarto grado según su nivel socioeconómico. Sugirieron que el hecho de que los padres hubieran terminado la secundaria representaba una ventaja. En el caso de los estudiantes de 15 años, es posible que un indicio de rendimientos decrecientes se deba a que no hicieron una transición adecuada entre la fase de aprender a leer y la de leer para aprender al final de la primaria baja, y por ende no pudieron beneficiarse tanto de la enseñanza en grados superiores. Esto puede también atribuirse a un efecto piso en la prueba PISA: en la mayoría de los países de ingreso bajo y medio, un gran porcentaje de estudiantes con un nivel socioeconómico bajo no pueden responder bien ni siquiera las preguntas más fáciles de la prueba.

Igualdad. Una medida directa de la igualdad es la diferencia en las puntuaciones de la prueba entre dos subpoblaciones. La diferencia entre las habilidades lectoras de los estudiantes cuya primera lengua es el español con las de aquellos que tienen otra primera lengua es de aproximadamente 74 puntos. El método del gradiente del NSE permite determinar qué tan grande es la diferencia después de que se controla este indicador. También permite discernir si hay una interacción significativa asociada con él; es decir, ¿la diferencia varía en distintos niveles socioeconómicos? La **Figura 11** muestra las líneas de gradiente de las dos subpoblaciones. En este caso, el NSE se relaciona sólo parcialmente con las diferencias entre las dos subpoblaciones; la diferencia de los estudiantes con un NSE promedio es de aproximadamente 69 puntos. La figura también indica que la diferencia es bastante uniforme a lo largo del rango del NSE.



Figura 11. Gradientes socioeconómicos de la competencia lectora en México por primera lengua



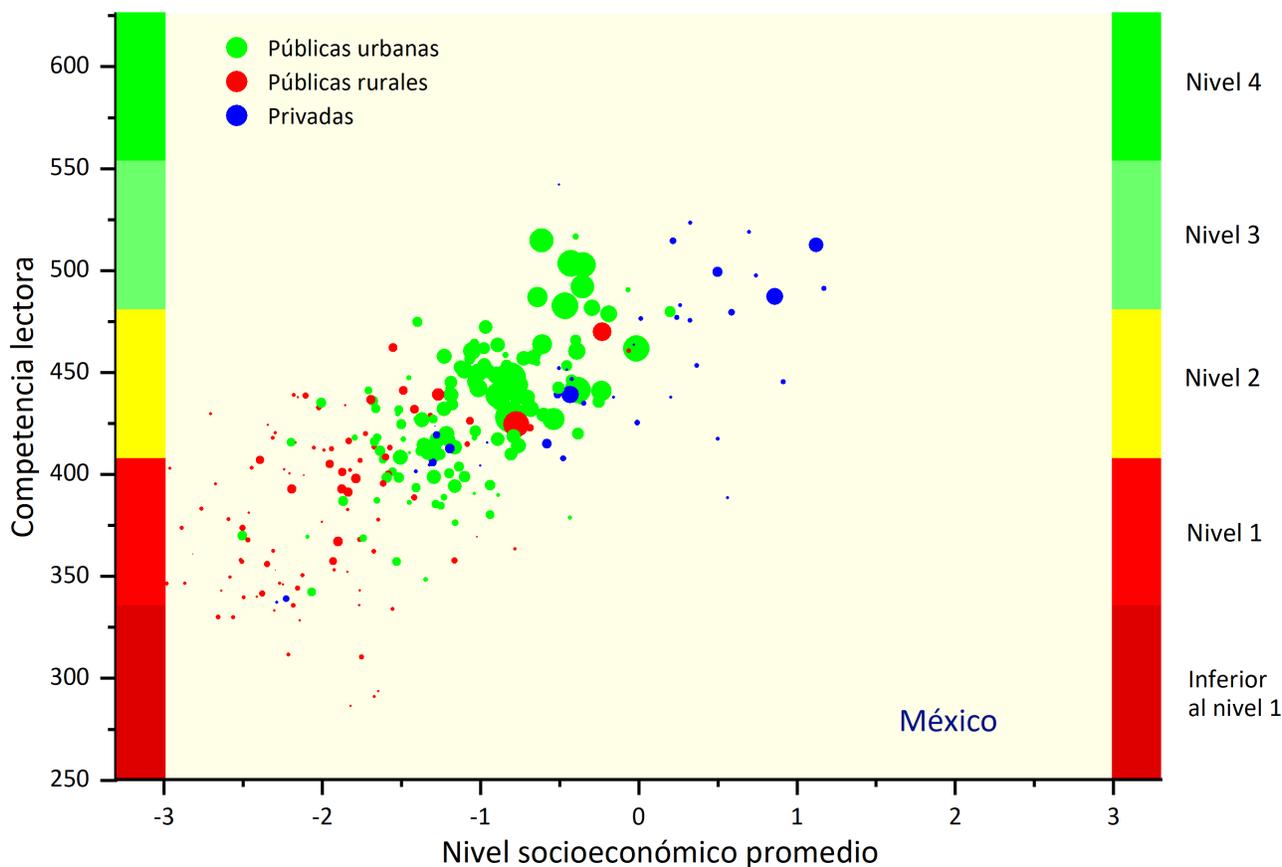
Fuente: PISA 2015.

¿Dónde se encuentran los estudiantes vulnerables?

Perfiles escolares. Un perfil escolar es simplemente un diagrama de dispersión de las puntuaciones promedio de las escuelas en un resultado, relacionado con su nivel socioeconómico medio. La **Figura 12** presenta el perfil escolar de los resultados de México en lectura. Cada punto representa a una de las escuelas que participaron en el estudio PISA 2015. El tamaño de los puntos es proporcional a la raíz cuadrada de la matrícula de las escuelas. Los colores de los puntos indican si se trata de una escuela pública rural (rojo), pública urbana (verde) o privada (azul).



Figura 12. Perfil escolar de la competencia lectora en México



Fuente: PISA 2015.

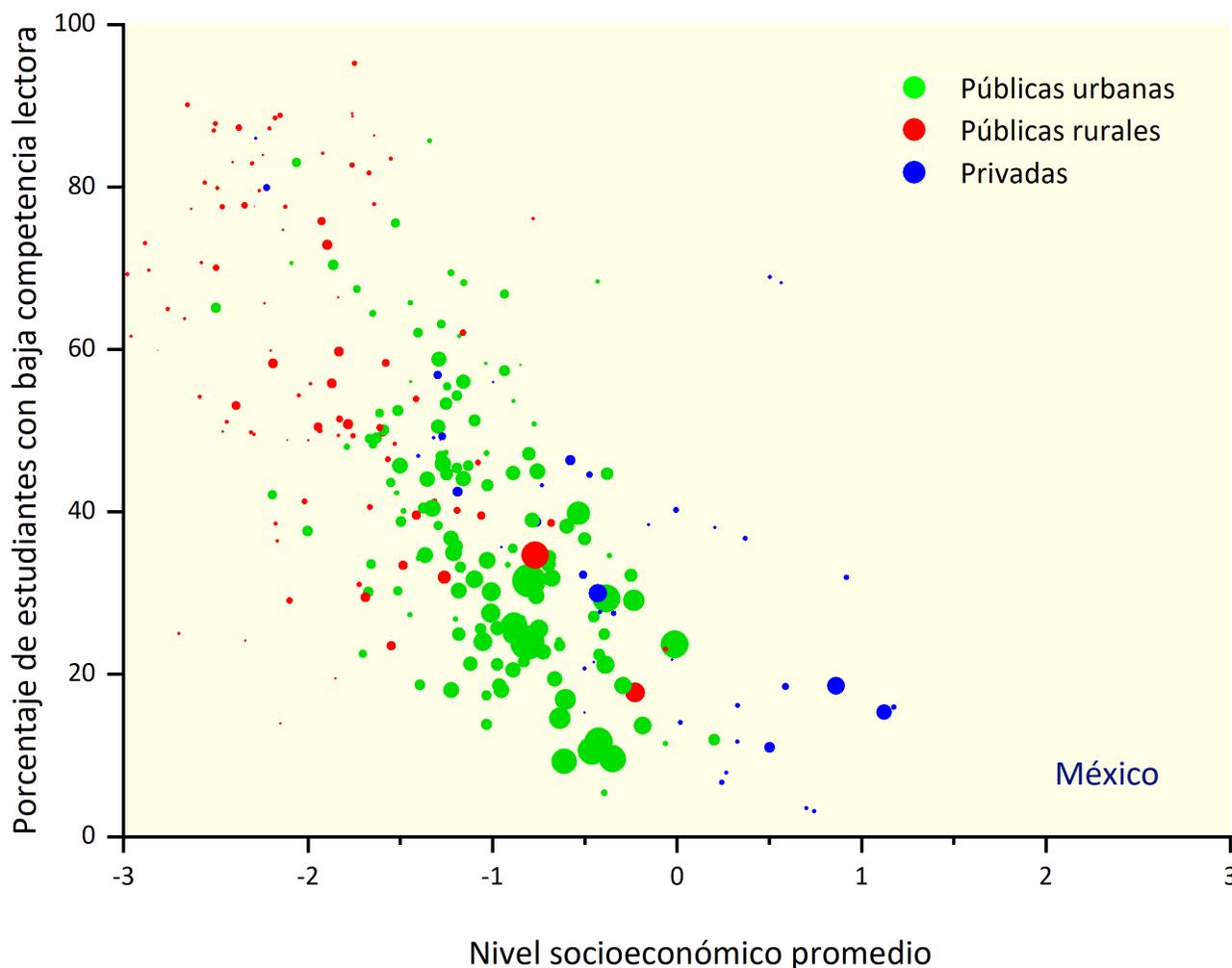
De este perfil escolar inmediatamente saltan a la vista tres hallazgos.

- a. Las escuelas difieren sustancialmente en sus puntuaciones de lectura promedio. En todos los NSE, el rango entre las escuelas con las puntuaciones de lectura más bajas y aquellas con las puntuaciones de lectura más altas es de 100 puntos aproximadamente. Esto es una desviación estándar, o un tamaño de efecto de 1.00.
- b. Si establecemos de manera informal que el NSE con un valor de -1.0 define la “línea de pobreza”, el perfil escolar revela que casi las dos terceras partes (el 65%) de las escuelas tienen un NSE inferior a la línea de la pobreza. Casi todas ellas son establecimientos rurales pequeños. Para PISA-D se establecieron una medida más específica del NSE y una medida de la pobreza (Tramonte y Willms, 2018).



- c. El nivel socioeconómico medio de la mayoría de las escuelas privadas es considerablemente más alto que el de las escuelas públicas urbanas; no obstante, su puntuación de lectura media es comparable con la de las escuelas públicas urbanas con un NSE medio superior a -0.5.

Figura 13. Perfil escolar de baja competencia lectora (nivel 1 o inferior) en México



Fuente: PISA 2015.

La **Figura 13** presenta el perfil escolar de baja competencia lectora en México. Para los fines del presente análisis, la baja competencia lectora se define como aquella que tiene una puntuación PISA del nivel 1 o inferior. Este tipo de perfil, que muestra el “porcentaje vulnerable”, es útil para los resultados dicotómicos. Los resultados pueden ser negativos, por ejemplo, pocos logros o una mala salud mental; o positivos, por ejemplo, la terminación de los estudios. Este perfil va de la mano con los gráficos de concentración de la vulnerabilidad, que se abordan más adelante.



Sistemas escolares inclusivos Los sistemas escolares inclusivos tienen escuelas en las que todos los niños y todas las niñas pueden prosperar. “Todos” significa absolutamente todos los estudiantes, sin distinción de género, discapacidades, clase social, etnicidad, origen nacional, orientación sexual y religión (Willms, 2009a). “Prosperar” significa explotar todo el potencial académico, social, emocional y físico propio. La inclusión implica un conjunto de valores y creencias que sustentan la idea de que todos los niños y todas las niñas tienen el derecho legal de recibir educación en un entorno seguro junto con sus pares vecinos. La inclusión entraña “una visión común que abarca a todos los niños en edad escolar y la convicción de que es responsabilidad del sistema educativo regular educar a todos los niños y todas las niñas”. (UNESCO, 2005, p. 13). Para lograr la inclusión es necesario adaptar la educación a los estudiantes con necesidades especiales o problemas de conducta, pero el principio general es que tengan el derecho de ser educados en el entorno menos restrictivo. La creación de un sistema escolar inclusivo requiere que los padres, los educadores, los políticos y otros líderes comunitarios adopten una postura sobre la inclusión contrarrestando las actitudes discriminatorias y apoyando las iniciativas que fomentan la participación de estudiantes con necesidades diversas (Riehl, 2000).

Los sistemas escolares inclusivos tienen mejores resultados estudiantiles y menos desigualdad (Willms, 2010). Cuando los sistemas escolares son más inclusivos, los fundamentos para el éxito – enseñanza de calidad, tiempo de aprendizaje, recursos materiales y apoyo familiar y comunitario– se distribuyen de manera más equitativa entre las escuelas. Además, los estudiantes mismos son un recurso clave. Los de baja capacidad tienen mayores probabilidades de éxito si participan en clases de habilidades mixtas con compañeros que tienen altas expectativas y están intelectualmente comprometidos.

Inclusión horizontal. A nivel de sistema, podemos considerar dos tipos de inclusión: horizontal y vertical. En los sistemas escolares horizontalmente inclusivos, la variación del nivel socioeconómico del estudiantado es relativamente pequeña. Por ejemplo, el NSE medio de la población estudiantil mexicana, basado en el perfil escolar, fluctúa entre -3.0 y 1.2. La segregación residencial es el factor más importante de la segregación horizontal. En México existen notables diferencias socioeconómicas entre las colonias o vecindarios de las grandes ciudades tanto en las zonas urbanas como en las rurales.

Inclusión vertical. En los sistemas escolares verticalmente inclusivos, la variación en los resultados de los alumnos de escuelas con un NSE similar es pequeña. El perfil escolar de México indica que el rango de puntuaciones de competencia lectora en la mayoría de las escuelas va de 350 a 550. En los sistemas escolares “verticalmente segregados”, es decir, que no son verticalmente inclusivos, los estudiantes se distribuyen entre ciertos tipos de escuelas o programas de enseñanza según su desempeño académico previo o una determinada medida de sus competencias cognitivas o lingüísticas. También es posible que se les seleccione para programas o cursos particulares, como clases para estudiantes con necesidades especiales. La repetición de grados, los programas de inmersión a un nuevo idioma y los planes de estudio diferentes contribuyen a la segregación vertical.



Las escuelas privadas también pueden fomentar ambos tipos de segregación. Si seleccionan a los estudiantes con base en su desempeño académico previo, contribuyen a la segregación vertical. Si cobran una inscripción elevada, lo que las hace prohibitivamente caras para los estudiantes con un NSE bajo, contribuyen a la segregación horizontal.

Para resumir el nivel de inclusión vertical y horizontal, se utilizan dos parámetros estadísticos. La variación en el desempeño de los estudiantes se divide en un componente intraescolar y un componente interescolar. La inclusión vertical es la proporción de la variación en los resultados de los estudiantes que se presenta *dentro* de las escuelas. La variación en el nivel socioeconómico de los estudiantes también se divide en un componente intraescolar y un componente interescolar. La inclusión horizontal es la proporción de la variación en el NSE *dentro* de las escuelas. Teóricamente, estos índices pueden variar entre 0% y 100%, pero los de la mayoría de los países se sitúan dentro del rango del 10% al 50% (Willms, 2010). El Apéndice 2 describe las técnicas estadísticas para el cálculo de estos índices.

Cabe señalar que la medida de inclusión vertical depende de los resultados que se estén considerando. Los niveles medios del bienestar de los estudiantes tienden a ser muy similares entre las escuelas; así que las estimaciones de inclusión vertical basadas en una medición de la autoestima, por ejemplo, arrojarían un índice de inclusión vertical mucho más alto que las basadas en el desempeño en lectura. Dada la importancia de las habilidades lectoras, un índice de inclusión vertical fundado en una prueba de lectura nacional o internacional sería útil para dar seguimiento a la inclusión vertical. El nivel de inclusión vertical también puede variar según el nivel de escolaridad, especialmente en los sistemas que seleccionan a los estudiantes de acuerdo con su desempeño académico o su capacidad a determinadas edades.

Concentración de la vulnerabilidad. Un gráfico de concentración de la vulnerabilidad es una curva de Lorenz, que se usa en economía para representar la distribución del ingreso o la riqueza (Lorenz, 1905). Este tipo de gráfico muestra la concentración de los estudiantes vulnerables entre las escuelas de un país. La **Figura 14** contiene un ejemplo. Los resultados de PISA 2015 indican que en México hay 44,335 escuelas con estudiantes de 15 años en el séptimo grado o uno superior. El tamaño de la población es de 1,392,995 estudiantes. El 41.8% de ellos tiene habilidades lectoras del nivel 1 o inferiores. La Figura 14, que es un histograma, consta de 443 barras, cada una de las cuales representa 100 escuelas. La altura de las barras indica el número estimado de estudiantes vulnerables en cada conjunto de escuelas, que están ordenadas por número de niños vulnerables en un conjunto de 100 escuelas. Por ejemplo, la altura de la barra 400ª indica que en ese conjunto hay 3,826 estudiantes, o un promedio aproximado de 38 estudiantes por escuela que tienen una competencia lectora del nivel 1 o inferior. El grupo con el número más alto tiene 10,838 estudiantes, o alrededor de 108 estudiantes por escuela.

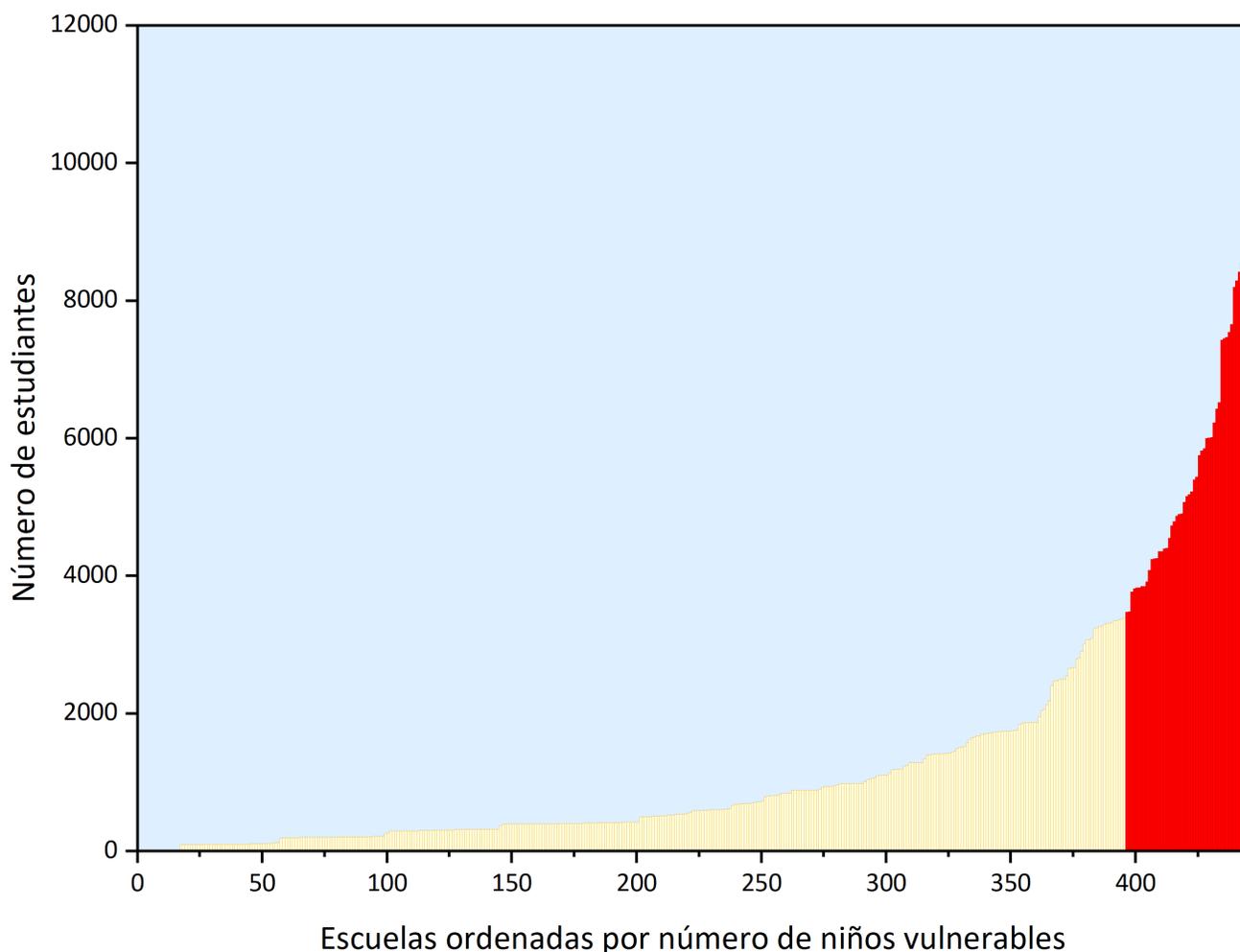
Este diagrama de concentración de la vulnerabilidad revela que el 50% de los estudiantes con una competencia lectora baja se encuentra en alrededor del 11% de las escuelas, que representa menos de 5.000 escuelas. El 75% de los estudiantes vulnerables se encuentra en el 18% de las escuelas, es decir, en poco menos de 8.000 escuelas. En la Sección IV del presente informe, uno de los cinco tipos de estrategias se



refiere a las focalizadas en los resultados. En este caso, una estrategia viable consistiría en implementar una intervención en 5,000 a 8,000 escuelas. Esto puede resultar desalentador, pero quizá no tanto si se considera que en México existen más de 44,000 escuelas a las que asisten estudiantes de 15 años.



Figura 14. Concentración de estudiantes con baja competencia lectora en escuelas mexicanas



Fuente: PISA 2015.

Los diagramas de concentración de la vulnerabilidad también pueden servir para evaluar la concentración de cualquier subpoblación, por ejemplo, la de estudiantes indígenas o la de estudiantes con una discapacidad. La subpoblación no tiene que ser “vulnerable”. Por ejemplo, en México, alrededor del 4.5% de los estudiantes mostraron habilidades lectoras del nivel 4 o un nivel superior; y el 50% de ellos asistían a menos del 2% de las escuelas del país.

Gradientes socioeconómicos intra e interescolares. La Figura 10 indica el gradiente socioeconómico general de México. El gradiente socioeconómico se puede estimar para cualquier jurisdicción, tal como un país, región o distrito escolar. Sin embargo, dentro de cada jurisdicción, las escuelas tienen sus propios gradientes de NSE, cada uno con una pendiente y una fuerza distintas. Un modelo de gradientes multinivel



describe simultáneamente los que corresponden a todas las escuelas y las relaciones entre ellos. El gradiente intraescolar medio de México es de 7.7 y el interescolar, de 41.0. Esto sugiere que los estudiantes de cada escuela por lo general tienen niveles de competencia lectora similares. Las desigualdades en la competencia lectora asociadas con el nivel socioeconómico tienen que ver principalmente con las diferencias entre las escuelas.

Composición de la escuela. “Efecto de composición de la escuela” se refiere al efecto que tienen sobre los resultados de los estudiantes las características agregadas de una escuela, tales como su NSE medio, el porcentaje de estudiantes con una discapacidad, o el porcentaje de estudiantes que hablan en casa un idioma distinto al idioma de la enseñanza (Alexander y Eckland, 1975; Bryk y Driscoll, 1988; Willms, 1986; Willms, 2010). La presencia de un efecto de composición indica que además de los efectos asociados al NSE, existe un efecto asociado a la composición de la escuela. Si el efecto de composición asociado al NSE de la escuela es positivo, indica que los estudiantes que asisten a escuelas con un NSE alto por lo general muestran mejores resultados que los que asisten a escuelas con un NSE bajo, incluso después de considerar el NSE de cada estudiante. En el caso de México, el efecto de composición asociado al NSE medio de la escuela es de 35.2.

El efecto de composición debería considerarse como una variable sustitutiva de diversos factores correlacionados con la composición de la escuela, porque los datos transversales no permiten separar los efectos asociados a la composición de la escuela de los efectos atribuibles a factores fundamentales u otros procesos escolares (Raudenbush y Willms, 1995). Hasta cierto punto, los efectos de composición pueden atribuirse a los efectos entre pares: cuando colaboran entre sí estudiantes brillantes y motivados, aprenden unos de otros y establecen estándares de desempeño más altos (Robertson y Symons, 1996; Zimmer y Toma, 1997). Además, las escuelas cuyo estudiantado tiene un NSE alto tienden a atraer y retener a maestros talentosos y motivados, disponen de mayores recursos para la enseñanza y reciben más apoyo de los padres (Willms, 1986; Willms y Somers, 2001; OCDE, 2001). En otras palabras, tienen fundamentos para el éxito más sólidos. El término “efectos” connota una relación causal, pero de los estudios transversales como PISA no se puede inferir una relación causal (Alexander, Fennessey, McDill y D’Amico, 1979).

Variabilidad de la población estudiantil. Varias políticas y prácticas educativas se basan en el argumento de que a mayor homogeneidad de los estudiantes que asisten a una escuela o un aula, mayor ritmo de aprendizaje. Este argumento sustenta las políticas que apoyan la transición de los estudiantes a escuelas o programas escolares de orientación académica o vocacional, así como la práctica de que los estudiantes repitan un grado cuando su desempeño académico es muy inferior al exigido.

Al igual que el efecto de composición de la escuela, el efecto de composición del estudiantado se calcula incluyendo una medida de la variabilidad de ésta, por ejemplo, la desviación estándar del NSE de la escuela como una variable a nivel de escuela en un modelo multinivel. El efecto de la variabilidad de la población estudiantil de México es de 5.9, un efecto pequeño, estadísticamente insignificante.



Gradientes convergentes. Según los resultados de investigaciones anteriores basadas en datos de PIRLS 2001 y PISA 2000, las habilidades lectoras de los estudiantes con un NSE alto no variaron sustancialmente entre los países en comparación con las de los estudiantes con un NSE bajo (Willms, 2006). En otras palabras, los gradientes tendieron a converger de la mano con el aumento del NSE. Esta tendencia también fue evidente entre los jóvenes de 16 a 25 años que participaron en la Encuesta Internacional sobre Alfabetización de la Población Adulta de 1997 (Willms, 2003b). Para cada país, es importante determinar si las gradientes de las escuelas convergen a medida que se eleva el nivel socioeconómico. Si lo hacen, significa que los estudiantes que provienen de un medio con un NSE alto por lo general se desempeñan bien en la mayoría de las escuelas; mientras que las habilidades de aquellos que provienen de un medio con un NSE bajo varían considerablemente, dependiendo de la escuela a la que asisten.

La estadística para estimar el grado de convergencia es la correlación entre los niveles de los gradientes y sus pendientes. Puede obtenerse aplicando un modelo de regresión de dos niveles que incluye una medida del NSE de la familia de cada estudiante y permite variar tanto las intersecciones como las pendientes de este indicador entre escuelas. En el caso de México, dicha correlación es de 0.37, lo que indica que las pendientes divergen cuando se eleva el nivel socioeconómico.

En el Apéndice 2 se explica cómo se calculan las estadísticas basadas en un modelo de gradientes multinivel.



IV. Políticas sobre estrategias y su ejecución

Los datos obtenidos a partir de un proceso de seguimiento pueden ayudar en la formulación de las preguntas de política relativas al desempeño del sistema escolar: ¿Cómo vamos? ¿Quiénes son vulnerables? y ¿Dónde se encuentran los niños vulnerables? En la sección anterior se identificaron varias estadísticas y gráficas clave que pueden utilizarse para analizar de manera sintética el desempeño de un sistema escolar. Esos instrumentos constituyen un método coherente para evaluar los niveles de desempeño estudiantil y su igualdad entre subpoblaciones de estudiantes. También se pueden usar datos de seguimiento a fin de establecer objetivos realistas y medibles de mejoramiento de resultados de los estudiantes y reducción de las desigualdades.

La política educativa implica establecer metas y un curso de acción para alcanzarlos. El curso de acción requiere identificar un pequeño conjunto de estrategias para lograr los resultados y un plan para su ejecución. También requiere fijar prioridades, establecer objetivos a corto y largo plazo alineados con las metas y monitorear el progreso hacia esos objetivos, así como formular políticas para asignar de una manera idónea los recursos disponibles. Los datos de seguimiento son esenciales para formular un conjunto de estrategias y planes para su ejecución.

En esta sección se consideran tres preguntas sobre las estrategias y su ejecución. Primero, ¿qué tipos de estrategias tienen más probabilidades de inducir mejoras en el sistema? Segundo, ¿cómo se pueden asignar los recursos con miras a fortalecer los fundamentos para el éxito y tener un sistema escolar más equitativo? Tercero, ¿cómo se pueden utilizar los datos de seguimiento para evaluar los efectos de las políticas que modifican una o varias de las características estructurales clave del sistema escolar?

Cinco tipos de estrategias. Willms (2006) describió cinco tipos de estrategias susceptibles de implementarse en un país, una provincia o estado, una jurisdicción o una escuela³. Las relaciones entre los resultados estudiantiles y el NSE, que se describen con los gradientes socioeconómicos y los perfiles escolares, pueden ayudarnos a discernir qué tipo de estrategia o combinación de estrategias tiene más probabilidades de elevar y nivelar la barra de aprendizaje. Dichos tipos de estrategias son: universal, centrada en el desempeño, centrada en el NSE, compensatoria y de reubicación. A continuación se abordan utilizando como ejemplo los datos de PISA 2015 para México. Por cada estrategia, se considera el efecto potencial de una estrategia hipotética. En los diagramas que muestran los efectos hipotéticos, la línea de gradiente roja denota el estatus de “anterior a la intervención”, que se ajusta al gradiente observado con los datos de PISA 2015. La línea verde denota el gradiente hipotético “posterior a la intervención”.

Estrategia universal. Una estrategia universal está encaminada a mejorar los resultados de todos los estudiantes en una jurisdicción. Las reformas curriculares, la reducción del tamaño de los grupos, la

³ En la versión anterior de Las brechas de aprendizaje (Willms, 2006) se utilizó el término “intervenciones”, pero este término a menudo se emplea para connotar una estrategia específica, tal como la introducción de un nuevo programa o técnica pedagógica de lectura.



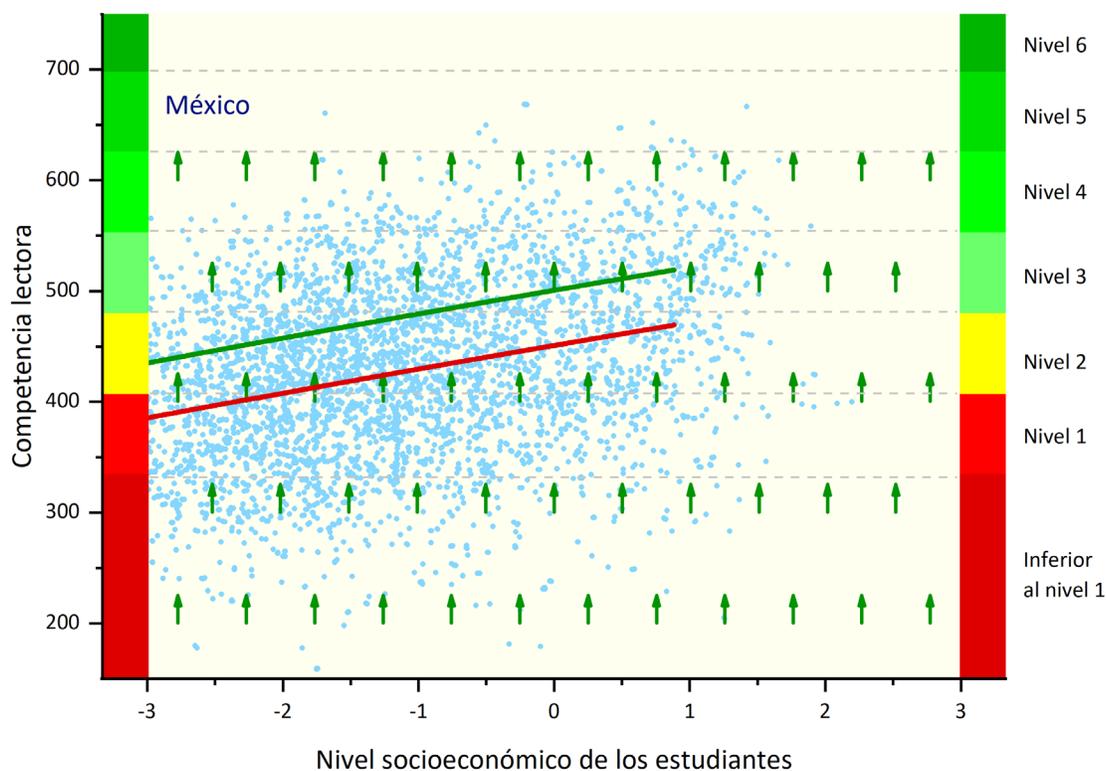
modificación de la edad de ingreso a la preprimaria y el aumento del tiempo dedicado a la enseñanza en lectura son estrategias universales porque se aplican a todos los estudiantes, sin importar su nivel socioeconómico.

La **Figura 15** muestra el efecto de una estrategia universal destinada a los estudiantes de todos los niveles socioeconómicos. El tamaño de efecto de la estrategia hipotética es de 0.5, equivalente a 50 puntos en la escala PISA. Algunas estrategias universales tienen efectos más significativos en los estudiantes de un NSE bajo que en los de un NSE alto, o viceversa.

Estrategia centrada en el desempeño. Una estrategia destinada a los estudiantes con un bajo nivel de desempeño en un resultado específico es una intervención centrada en el desempeño. Por ejemplo, la Evaluación Infantil Temprana se usa en varios países para identificar las habilidades de desarrollo de los niños de 3 a 6 años mientras se preparan para la transición a la educación formal (The Learning Bar, 2011). Los datos recopilados se usan para clasificar a los estudiantes en tres grupos según sus puntuaciones en cinco dominios. Esta clasificación facilita a los maestros información sobre *el tipo y la cantidad* de apoyo necesario para cada niño.

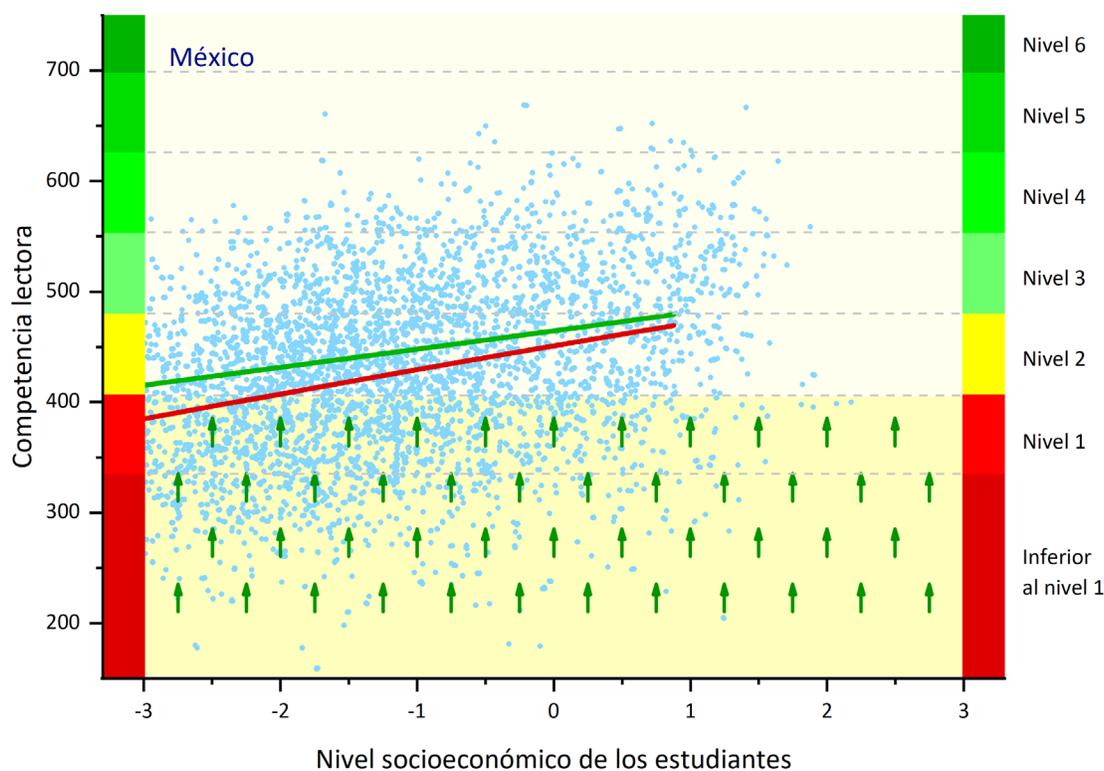
Una estrategia centrada en el desempeño también es susceptible de implementarse a nivel de escuela. Por ejemplo, se puede administrar un programa de lectura en una muestra de escuelas con un desempeño promedio bajo. En los sistemas escolares que tienen un índice de inclusión vertical bajo, es eficiente implementar una estrategia escolar integral en un pequeño número de escuelas. Se puede usar un diagrama de concentración de la vulnerabilidad para calcular la cantidad de niños que abarcaría una intervención en cierto número de escuelas.

Figura 15. Efectos de una estrategia universal con un tamaño de efecto de 0.50



La **Figura 16** muestra el efecto de una estrategia centrada en el desempeño aplicada a todos los estudiantes que obtuvieron una puntuación del nivel 2 o inferior. Tal estrategia hipotética tiene un tamaño de efecto de 0.50. En este caso, eleva y nivela la barra de aprendizaje porque conforme baja el nivel socioeconómico, aumenta desproporcionadamente el número de estudiantes con un nivel de desempeño bajo.

Figura 16. Efectos de una estrategia centrada en el desempeño con un tamaño de efecto de 0.50, dirigida a estudiantes con una competencia lectora de nivel 1 o inferior



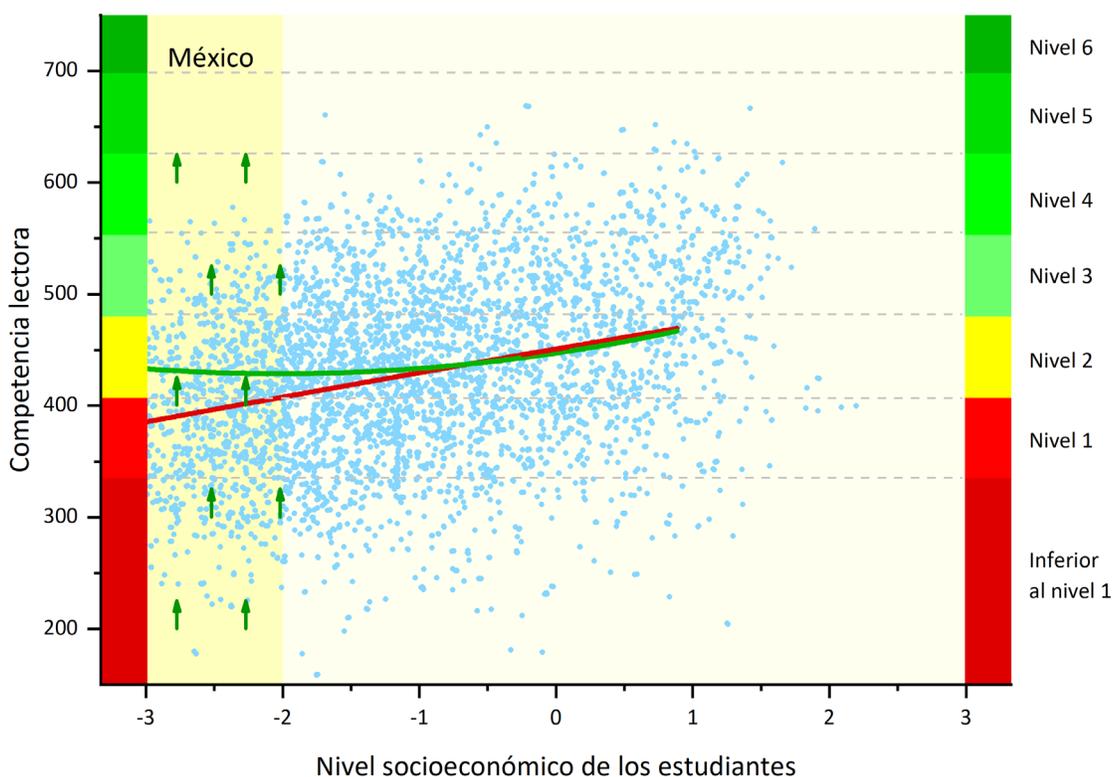
Estrategia centrada en el riesgo. Una estrategia centrada en el riesgo está encaminada a brindar apoyo o recursos adicionales a los niños que se consideran en riesgo de abandonar la escuela, como los que tienen un nivel socioeconómico bajo. La diferencia entre una estrategia centrada en el riesgo y una centrada en los resultados es que la primera se focaliza en los niños considerados “en riesgo” por alguna característica particular, en vez de en los que a todas luces han tenido un resultado de desarrollo bajo. Un programa preescolar ventajoso para niños de familias de bajos ingresos es un buen ejemplo de una intervención centrada en el riesgo. Este tipo de estrategia también se puede dirigir a una subpoblación considerada vulnerable, como la de nuevos inmigrantes o la de estudiantes que pertenecen a una minoría étnica.

La **Figura 17** muestra un ejemplo de una estrategia centrada en el riesgo con un tamaño de efecto de 0.50 para estudiantes con una puntuación de NSE inferior a -2.0. Esta estrategia eleva la barra para los estudiantes con un NSE bajo y por lo tanto la nivela, al reducir la brecha entre ellos y los estudiantes con un NSE alto. Un problema de las estrategias centradas en el riesgo es que los estudiantes con un NSE alto con puntuaciones bajas no reciben la intervención. Este problema se puede resolver implementando una estrategia centrada en el desempeño junto con una estrategia centrada en el riesgo. Sin embargo, no siempre es posible implementar una estrategia centrada en el desempeño. Por ejemplo, no se pueden medir fácilmente los resultados de desarrollo de los niños durante las primeras fases de su ciclo vital y por ende es más práctico aplicar una estrategia dirigida a aquellos a los que se considera en riesgo. El factor de riesgo podrían ser las familias con un nivel socioeconómico bajo o los niños prematuros o con bajo peso al nacer. Tampoco es posible implementar una estrategia centrada en el desempeño para ciertos tipos de resultados,



como un comportamiento sexual riesgoso o el abandono de la escuela. Por lo tanto, para identificar a la población objetivo, tenemos que echar mano de la información disponible sobre los factores de riesgo y de protección.

Figura 17. Efectos de una estrategia centrada en el riesgo con un tamaño de efecto de 0.50 dirigida a estudiantes con un NSE de -2 o inferior



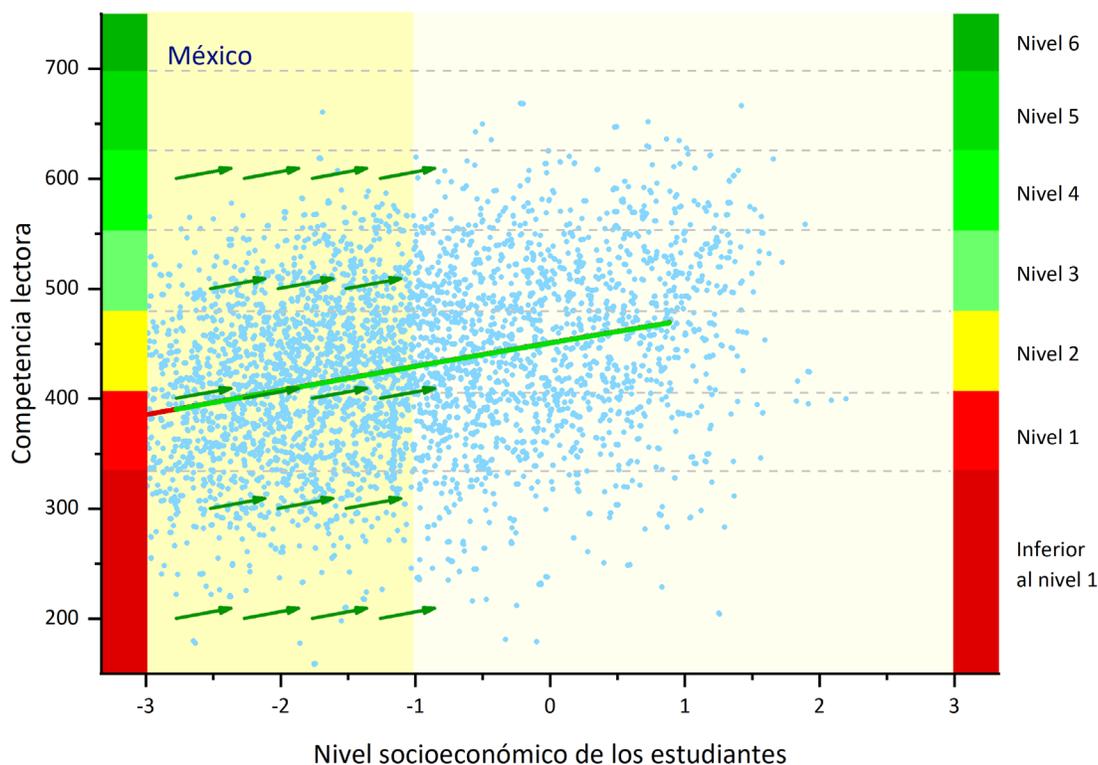
Otra preocupación relacionada con las estrategias centradas en el riesgo es que los estudiantes que son considerados en riesgo pero tienen una alta puntuación en los resultados también reciben la intervención. Podría decirse que esta estrategia no es eficiente en costos. Pero este argumento podría refutarse con la afirmación de que como esos niños mejoran sus habilidades, sirven de modelos para otros estudiantes con antecedentes familiares parecidos.



Estrategia compensatoria. Una estrategia compensatoria brinda recursos educativos adicionales a los estudiantes que provienen de un medio con un nivel socioeconómico bajo o que son considerados “en riesgo” por otras razones, por ejemplo, por no haber tenido un cierto resultado de desarrollo o haber tenido un resultado de desarrollo insuficiente. La subpoblación objetivo puede ser la misma que la de una intervención centrada en el NSE; pero la diferencia estriba en que una estrategia compensatoria tiene la intención de mejorar las circunstancias socioeconómicas de los niños con la expectativa de que mejoren sus resultados educativos. Los programas de desayunos o almuerzos escolares o libros de texto gratuitos son estrategias compensatorias.

En la **Figura 18** se pueden observar los efectos de una estrategia compensatoria que eleva las puntuaciones del nivel socioeconómico en desviaciones estándar de 0.25. En este ejemplo se supone que los niños beneficiarios de la estrategia mejorarán sus puntuaciones en lectura de una manera proporcional a la pendiente del gradiente de NSE. La pendiente de la gradiente de México se situó en 21.2 puntos, lo que sugiere que las puntuaciones aumentan en 21.2 puntos con cada incremento en el nivel socioeconómico. Por lo tanto, si el NSE del estudiante aumentó en desviaciones estándar de 0.25, podríamos prever un incremento aproximado de 5 o 6 puntos de sus puntuaciones en lectura. Una estrategia compensatoria también se puede implementar a nivel de escuelas. Sin embargo, ninguna estrategia mejora notablemente los resultados de los niños.

Figura 18. Efectos de una estrategia compensatoria que eleva el NSE de los estudiantes de bajo NSE en desviaciones estándar de 0.25

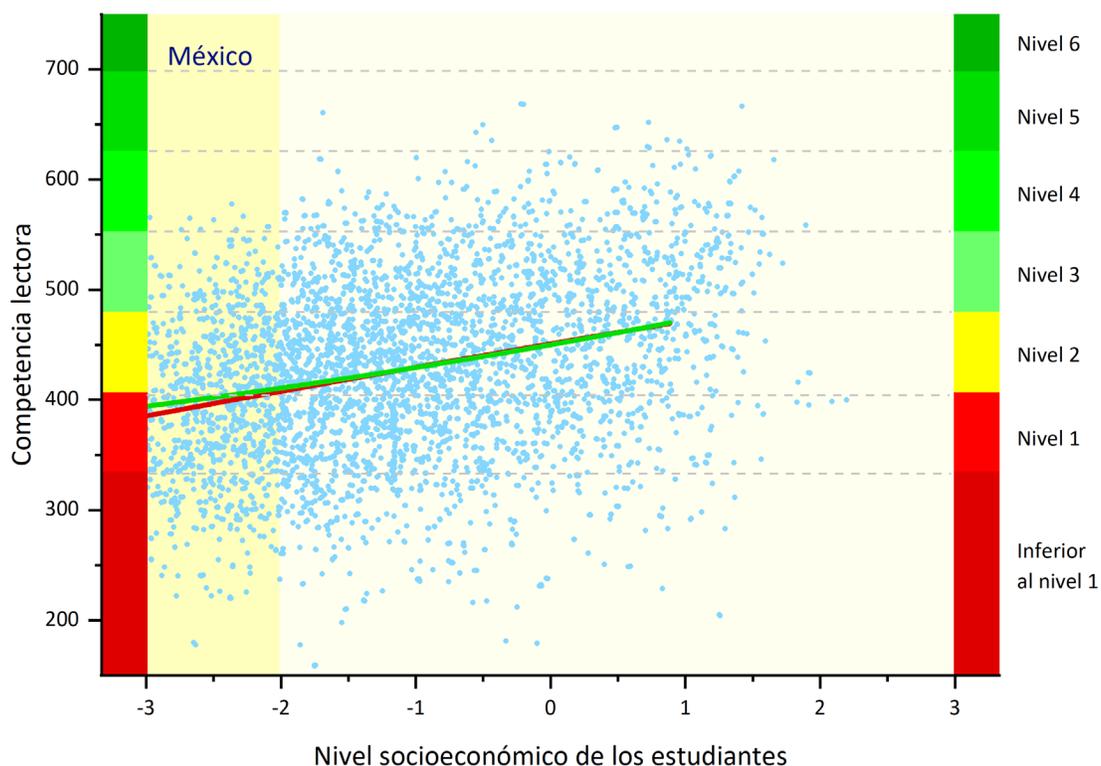


Estrategia de reubicación. Una estrategia de reubicación busca incluir en escuelas convencionales a los niños marginados o desfavorecidos, con base en el argumento de que así se beneficiarán con los efectos de composición de la escuela. Una estrategia consiste en cerrar escuelas pequeñas con un nivel socioeconómico bajo y reubicar a los estudiantes en otras escuelas. Esta estrategia no produce grandes efectos en el desempeño, y en algunas jurisdicciones puede influir negativamente en la transición oportuna de los estudiantes a lo largo de su trayectoria escolar. Por ejemplo, hay mayores probabilidades de que los niños permanezcan más tiempo en una escuela que les ofrezca una enseñanza completa desde la preprimaria hasta la secundaria alta. Otra estrategia factible en las grandes jurisdicciones escolares urbanas, consiste en modificar las zonas geográficas de actuación con vistas a reducir la segregación por nivel socioeconómico.

El perfil escolar de México, que se muestra en la Figura 12, indica que varias escuelas tienen un NSE inferior a -2.0. En la **Figura 19** pueden observarse los efectos de una estrategia de reubicación de alumnos de escuelas con un NSE inferior a -2.0 a escuelas con un NSE superior. Sin embargo, en México no es viable ninguna de las estrategias de reubicación antes descritas porque la mayoría de las escuelas con un nivel socioeconómico muy bajo son pequeños planteles rurales. Además, su efecto de elevación y nivelación de la barra de aprendizaje es mínimo.



Figura 19. Efectos de una estrategia de reubicación de los estudiantes de escuelas de bajo NSE a escuelas convencionales



Fortalecimiento de los fundamentos para el éxito

En el modelo de prosperidad educativa, los fundamentos para el éxito abarcan un conjunto de factores familiares, institucionales y comunitarios para cada una de las seis fases de desarrollo. En las tres últimas fases, cuando los niños cursan estudios formales, los fundamentos escolares abarcan una escuela segura e inclusiva, una enseñanza de calidad, tiempo de aprendizaje y recursos materiales. Cuando los educadores interactúan con las familias y los líderes comunitarios e influyen en ellos, contribuyen al capital social destinado al desarrollo infantil. Ésta es la razón por la que en el presente informe incluimos un quinto fundamento, el apoyo familiar y comunitario, como un factor escolar.

El modelo de prosperidad adoptado por PISA para el Desarrollo abarca estos cinco fundamentos (OCDE, 2017). La **Figura 20** muestra los resultados de prosperidad, los fundamentos para el éxito y las subpoblaciones de igualdad-equidad. Asimismo, resume las medidas incluidas en los cuestionarios para estudiantes, maestros y escuelas. Los óvalos verdes del centro de la figura presentan los cuatro resultados de prosperidad, cada uno de los cuales incluye una o varias medidas. Por ejemplo, el rendimiento académico

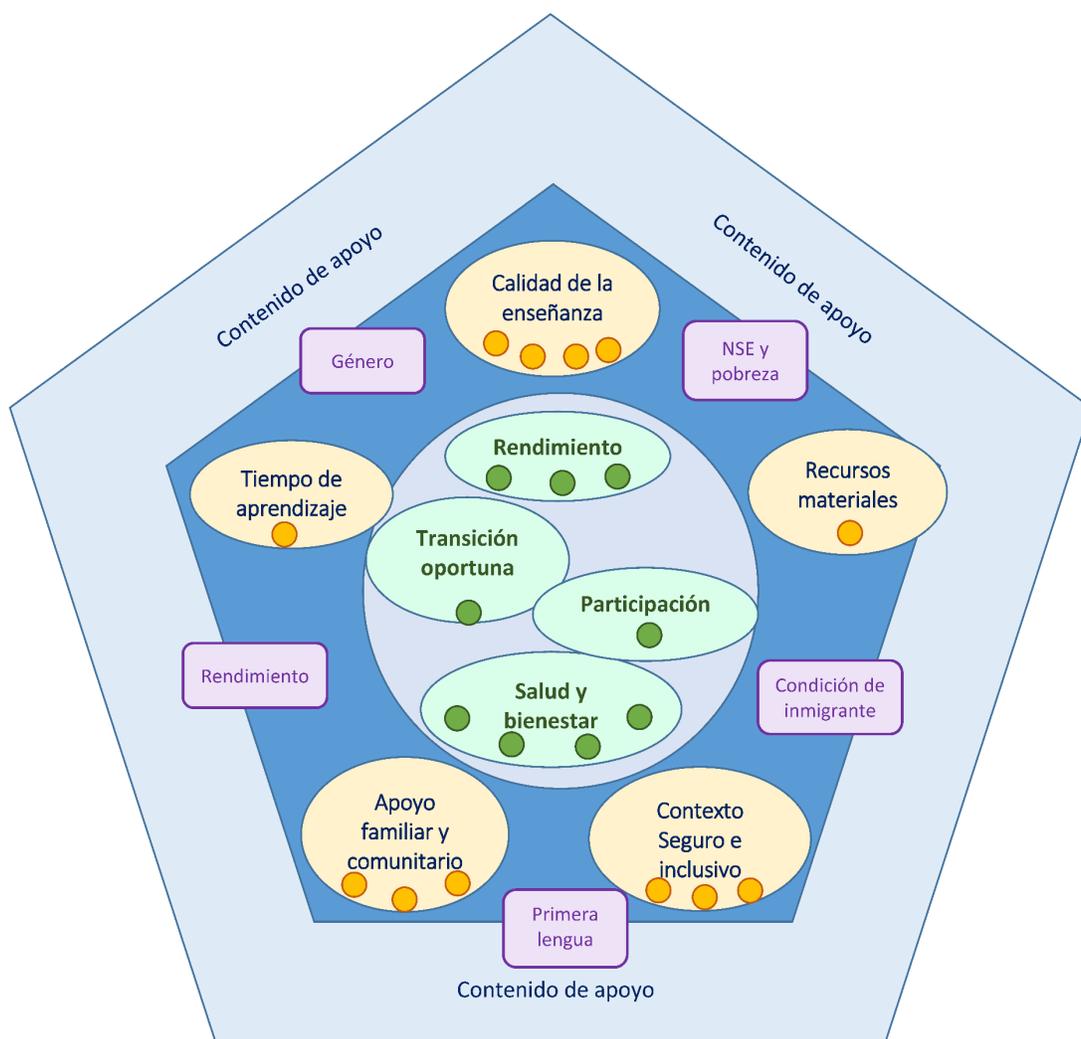


se evalúa con tres áreas: lectura, matemáticas y ciencias. Cada uno de los fundamentos para el éxito, que aparecen dentro de los óvalos anaranjados, incluye una o varias medidas. Los rectángulos morados muestran las subpoblaciones de interés para los países participantes. Los cuestionarios contextuales también incluyen otros factores pertinentes para esos países, que aparecen bajo el rubro “Contenido de apoyo”. El estudio PISA-D principal incluye un cuestionario para el estudiante, de 49 preguntas; un cuestionario para el profesor, de 33 preguntas; y un cuestionario para la escuela, de 28 preguntas (OCDE, 2017).

La idea clave detrás del diseño de este marco es enfocar la atención en un pequeño número de parámetros críticos para cada fase de desarrollo y lograr un consenso entre todas las partes interesadas del sistema escolar. Este podría ser un “cuadro sinóptico para el Ministro de educación”. Las medidas subyacentes a cada fundamento para el éxito proporcionan información sobre las fuerzas y debilidades del sistema escolar. Estos parámetros pueden utilizarse para establecer metas a corto y largo plazo y formular un plan de acción para alcanzarlas, que sienta las bases de un plan de comunicación del rumbo de la política educativa general, para todos los actores en todos los niveles del sistema escolar. El propósito es alinear las metas con las estrategias a nivel ministerial, jurisdiccional y escolar.

Una diferencia importante entre este enfoque y el enfoque “lógico” tradicional de insumo-proceso-producto en que se basan las investigaciones tradicionales sobre los efectos escolares es que no tiene la intención de capturar absolutamente todos los factores que contribuyen al éxito de los estudiantes. Es obvio que el “liderazgo de los directores” constituye un factor importante del éxito de los estudiantes, pero sus efectos dependen de los fundamentos centrales. Si en una jurisdicción se lanzara un programa de capacitación en liderazgo para los directores, sus efectos dependerían de la capacidad de éstos para efectuar cambios en la calidad de la enseñanza en las aulas, vigilar que la escuela sea segura e inclusiva, etc. Con esta perspectiva, el diseño de un programa de liderazgo para directores se centraría en cómo se definen y miden los fundamentos para el éxito, así como en los indicadores clave asociados a ellos y las estrategias para construir fundamentos sólidos. El enfoque también ayuda a identificar estrategias que tienen pocas probabilidades de operar cambios. Muy a menudo, los administradores tienen “proyectos favoritos” que en muchos casos convendría llevar a cabo pero que en realidad no reforzarían los fundamentos que producen resultados de prosperidad. El objetivo es que todos siempre mantengan la atención centrada en la construcción de los fundamentos para el éxito.

Figura 20. Marco de prosperidad educativa para PISA para el Desarrollo



Otra diferencia importante es que los “efectos” de los fundamentos se basan en una amplia literatura que respalda los efectos causales en los resultados, en vez de la estimación de sus relaciones a partir de los resultados de algún estudio transversal reciente o de cualquier estudio en particular. Los fundamentos son constructos independientes con medidas fiables y válidas que los representan. Estos parámetros pueden utilizarse para establecer metas a corto y largo plazo y formular un plan de acción para alcanzarlos. El siguiente ejemplo extremo sirve para ilustrar este punto. Fumar es malo para la salud. Podríamos tener un indicador de salud y bienestar que muestre la prevalencia de jóvenes no fumadores. No nos interesa estimar los efectos de fumar en las medidas de salud y bienestar evaluadas con el sistema de seguimiento, ni determinar si el “riesgo relativo” de fumar varía entre los países. Simplemente, lo correcto es reducir la prevalencia de jóvenes fumadores.

El enfoque de prosperidad educativa es congruente con los principios de la “correlación de resultados” (Earl, Carden y Smutylo, 2001). En particular, “no se basa en un marco de causa-efecto, sino que reconoce que el



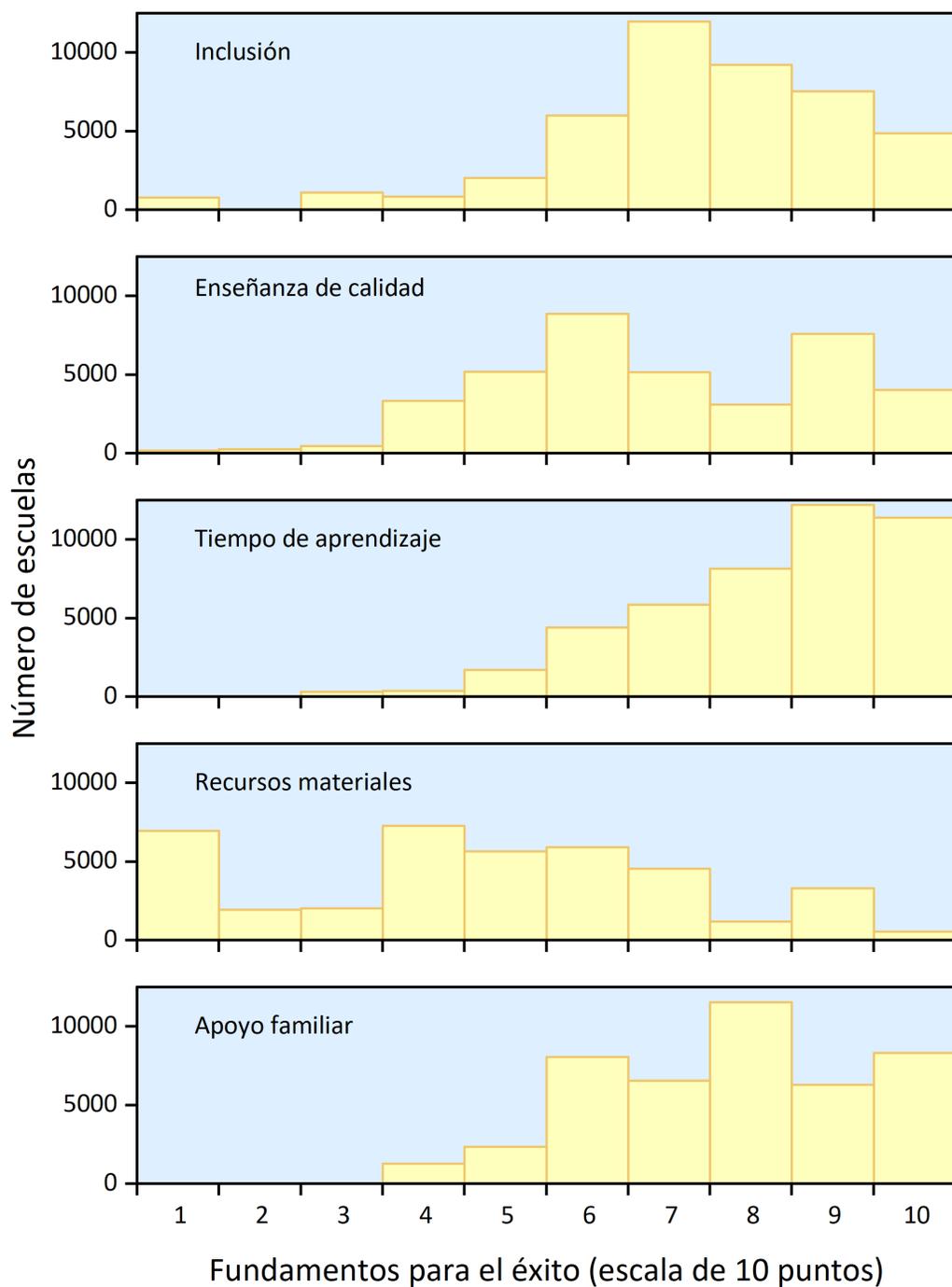
cambio es inducido por múltiples efectos no lineales. No intenta atribuir los resultados a una intervención o serie de intervenciones en particular. Más bien, considera los vínculos lógicos entre las intervenciones y los cambios conductuales” (Earl et al., p. 12). A manera de herramienta de planificación estratégica, se puede formular una cadena de estrategias emanadas de cada fundamento. Podríamos preguntarnos, por ejemplo: “¿Cuáles son las condiciones *necesarias* y *suficientes* para mejorar la calidad de la enseñanza?”. Podría decirse que una condición necesaria es que los maestros necesitan comprender a fondo la “concepción simple de la lectura” y crear un repertorio de estrategias de enseñanza que impliquen enseñar a los niños a leer. Después, esta condición podría vincularse con uno o varios proyectos específicos. No se puede atribuir a un proyecto o una intervención determinados una conexión causal, ya que siempre hay otros factores admisibles. Más bien, lo que se pretende es monitorear la “enseñanza de calidad” con un pequeño conjunto de indicadores clave.

Es ventajoso informar a las escuelas los resultados de la evaluación de los fundamentos para el éxito a medida que se desarrollan los indicadores. Aunque muchos de los factores involucrados en los resultados de la enseñanza varían adentro de las escuelas y entre ellas (Rowan, Raudenbush y Kang, 1991), la fijación de objetivos de mejora de los fundamentos para el éxito es más claro y más directo cuando se cuenta con un número pequeño de indicadores. También es beneficioso informar cada uno de los indicadores en una escala de 10 puntos y explicar las reglas de asignación de los puntajes. La prueba de fuego consiste en determinar si el director/la directora de cada escuela sabe cómo se evalúan los fundamentos y qué se requiere, digamos, para que su escuela suba de una puntuación de 4 a 5 o de 7 a 8 en cada fundamento.

La **Figura 21** contiene un gráfico de recursos escolares, nuevamente basado en los datos de PISA 2015 sobre México, como ejemplo. Un gráfico de recursos escolares es simplemente una serie de histogramas que resumen en una página los indicadores clave de los fundamentos para el éxito. Esta figura se presenta con fines ilustrativos, ya que PISA 2015 incluye solo un pequeño número de medidas que se relacionan con los fundamentos. Para este ejemplo, a cada escuela se le asignó una puntuación de inclusión derivada de la medida del “sentido de pertenencia a la escuela” por alumno. Cada punto de la escala indica que el 10% de los estudiantes de la escuela tienen un sentido de pertenencia positivo. Una puntuación de 6 indica que entre el 50% y el 60% de los estudiantes de la escuela tienen un alto sentido de pertenencia. Se utilizó un método similar para determinar el porcentaje de estudiantes que se consideró que recibían un fuerte apoyo familiar; y, con respecto al tiempo de aprendizaje, una tasa positiva de asistencia a la escuela y a clases. Las medidas de calidad de la enseñanza y recursos materiales se derivaron de las variables de recursos humanos y recursos materiales de los datos de PISA 2015 por escuela.



Figura 21. Gráficos de recursos escolares de México





Las técnicas antes descritas en este informe pueden servir para llevar a cabo análisis sobre la equidad de la provisión. Específicamente, en lo que atañe a las medidas a nivel de estudiante, los gráficos de gradientes socioeconómicos, perfiles escolares y gráficos de concentración de la vulnerabilidad proporcionan información útil para las políticas de asignación de recursos. Los perfiles escolares brindan información sobre las medidas a nivel de escuela. Para ciertas subpoblaciones, como aquellas que definen la etnicidad y la condición de inmigrante, es posible construir variables a nivel de escuela que denoten los porcentajes bajos, medios y altos de la subpoblación de interés; y después elaborar perfiles escolares similares a los del ejemplo presentado en la Figuras 12 y 13.

En este ejemplo, el gráfico de recursos escolares de México indica una variación considerable entre las escuelas en lo que se refiere a recursos materiales y enseñanza de calidad; mientras que las medidas de tiempo de aprendizaje y apoyo familiar varían menos y tienen puntuaciones más altas⁴. Obviamente, los datos del estudio PISA principal no permiten hacer una evaluación exhaustiva de los fundamentos para el éxito. Además, cubren únicamente una muestra de escuelas y un ciclo de tres años.

Modificación de las características estructurales de las escuelas

En la mayoría de las sociedades, las escuelas y los sistemas escolares siguen largas tradiciones en lo referente a la organización de la educación, las asignaturas y las formas de proveer enseñanza, estas representan las características estructurales de las escuelas. Dichas características incluyen, entre otros, los siguientes aspectos de la organización escolar a nivel de sistema: cuándo deben empezar los niños a asistir a la escuela, los grados que abarca cada nivel educativo, el tamaño de las escuelas, el plan de estudios, el idioma de la enseñanza, las estrategias de evaluación, el desarrollo profesional y las formas de selección de estudiantes para determinados programas escolares y escuelas. También incluyen los procedimientos internos de la escuela relativos a la organización y el horario de clases, los ciclos académicos, los mecanismos de enseñanza y la formación de grupos para la enseñanza. Las reformas escolares generalmente entrañan la modificación de una o varias de las características estructurales de las escuelas.

Cuando los gobiernos llevan a cabo una reforma para reestructurar las escuelas, raras veces se dan el lujo de realizar un *verdadero* experimento, que implique una intervención bien definida y el sujetar a las escuelas o a los estudiantes a ciertas condiciones de tratamiento y a un grupo de control. Más bien, las reformas tienden a exudar lentamente un cambio gradual de enfoques más burocráticos, con funciones laborales

⁴ La medida que utiliza PISA para evaluar los recursos materiales se basa en la siguiente pregunta del cuestionario para directores: “¿La capacidad de su escuela para proveer enseñanza se ve obstaculizada por alguno de los siguientes factores?”, la tiene cuatro opciones de respuesta: “No, en absoluto”, “Muy poco”, “Hasta cierto punto” y “Mucho”. PISA-D usa una medida más amplia que se basa en un esquema elaborado por Murillo y Román (2011), el cual plantea a los profesores preguntas sobre la disponibilidad y la condición de los recursos didácticos, y cómo los utilizan los profesores, y otras relacionadas con la disponibilidad y condición de los servicios básicos (OCDE, 2017; Tramonte y Willms, 2018).



especializadas y diferenciadas y líneas de autoridad formales descendentes, a formas comunitarias con funciones laborales variadas y responsabilidad compartida para un conjunto común de objetivos (Lee y Smith, 1993 y 1995). Fácilmente se puede clasificar a las escuelas en tradicionales y progresistas, y comparar los resultados académicos y sociales; pero no se puede atribuir fácilmente la causalidad a ningún determinado elemento de la reforma, como los grupos de niños con habilidades mixtas, los horarios de clases flexibles o dar más tiempo a los profesores para que preparen sus lecciones. Es difícil controlar los efectos de la selección y la variación en la fidelidad de la implementación. Además, cuando la reestructuración se da rápidamente, los eventos críticos o los cambios de dirección pueden impedir u obstaculizar los esfuerzos de la reforma, haciendo que sea imposible evaluar sus efectos. Comúnmente ocurren cambios en el gobierno e interferencias de los sindicatos de maestros, y a menudo los padres y los líderes comunitarios presionan para que se reinstaure el statu quo (Willms, 2008).

Para evaluar el impacto de las innovaciones estructurales, es esencial un sistema de seguimiento con medidas fiables y bien definidas de los resultados de prosperidad y los fundamentos para el éxito. Las mejoras positivas mejoran los resultados y reducen las desigualdades; en consecuencia, refuerzan los fundamentos. Los datos de seguimiento se pueden aprovechar para llevar a cabo cuasi experimentos sólidos con el fin de evaluar los efectos asociados a la modificación de las características estructurales de las escuelas.

Imaginemos que una jurisdicción escolar está convencida de que la estructura de la educación primaria se puede mejorar eliminando el sistema de organización de los alumnos en grados (1ro, 2ndo, 3ro, etc.), especialmente en las escuelas que tienen un alto porcentaje de alumnos vulnerables. La jurisdicción cuenta con un sistema de seguimiento que arroja datos útiles para el diseño y la evaluación de una intervención. Los administradores deciden implementar la intervención con el objeto de evaluar los efectos de la eliminación del sistema de grados en un pequeño número de escuelas. Los datos del seguimiento de este cambio estructural se pueden aprovechar por los menos de cinco formas.

Primero, una jurisdicción que cuenta con un buen sistema de seguimiento puede usar sus datos como base para llevar a cabo un estudio de intervención. Por ejemplo, puede implementar una intervención de lectura, como Confident Learners, en una muestra pequeña de escuelas sobre las que tiene datos básicos. El estudio puede consistir en una especie de experimento consistente en implementar la intervención solo en un determinado país o región; o en un experimento aleatorizado con grupos de tratamiento y de control. El principio medular es que los datos de seguimiento constituyen un “armazón” de insumo-proceso-producto que provee medidas básicas de los resultados de prosperidad y los fundamentos para el éxito. Estas medidas sirven para rastrear los cambios durante el proceso de reforma. Una ventaja que no hay que pasar por alto es que los costos de puesta en marcha de la intervención se reducen al mínimo porque la jurisdicción no tiene que invertir uno o dos años en el diseño de indicadores o en la identificación o definición de las subpoblaciones. La energía de la jurisdicción puede aprovecharse en el análisis de las dificultades inherentes a la implementación de la reforma.



Segundo, los datos de seguimiento pueden servir para identificar muestras estratégicas de escuelas; y los perfiles escolares y los gráficos de concentración de la vulnerabilidad, para identificar a las escuelas objetivo. El nivel socioeconómico medio de la escuela o el porcentaje de estudiantes vulnerables se pueden utilizar para definir muestras estratificadas (Sändal, Swensson y Wretman, 1992). Por ejemplo, una muestra de escuelas participantes potenciales se divide en estratos basados en sus perfiles, y después se toman muestras dentro de cada estrato.

Tercero, si es factible realizar una evaluación relativamente formal, los perfiles escolares se pueden usar en un diseño de bloques aleatorios. Por ejemplo, las condiciones de tratamiento asociadas a la intervención se asignan aleatoriamente a conjuntos de escuelas con un perfil similar (Matts y Lachin, 1988).

Cuarto, después de la recopilación, los datos de seguimiento pueden ser útiles para corroborar la validez interna de la evaluación. Las medidas demográficas pueden servir como controles estadísticos en el cálculo de los efectos de la intervención, y también para evaluar las interacciones de tratamiento por sujeto, por ejemplo: comparados con los programas que utilizan un sistema de organización por grados, ¿los programas que no aplican un sistema de grados tienen mayores efectos en los estudiantes con un NSE bajo que en aquéllos con un NSE alto?

Quinto, el sistema de seguimiento constituye una estrategia fiable para evaluar los cambios en la igualdad de los resultados de subpoblaciones clave.

Por último, podemos preguntarnos si la intervención tiene efectos secundarios deseables o indeseables. Aunque su objetivo principal sea mejorar las habilidades lectoras de los estudiantes, ¿también fomenta su participación? ¿Cómo afecta a la asistencia? ¿Es más o menos probable que los profesores apliquen estrategias de enseñanza de alto rendimiento?



V. Seguimiento de la prosperidad educativa

El marco de prosperidad educativa es una potente herramienta que los ministros de educación pueden usar para monitorear un conjunto básico de resultados de los estudiantes y los factores que los motivan. La utilidad de este marco ya se ha demostrado en países de ingreso bajo, medio y alto (Willms, 2018b), y tiene repercusiones positivas en el seguimiento del progreso hacia los Objetivos de Desarrollo Sostenible (ODS) para 2030. Esta sección del informe aborda las estrategias de establecimiento de metas, presenta un ejemplo de sistema de monitoreo exhaustivo y examina los rumbos de la evaluación en educación que pueden derivarse de la prosperidad educativa.

Establecimiento de objetivos

Un marco para establecer objetivos en los campos de la administración y la educación se denomina con el acrónimo SMART, y generalmente su concepción se le atribuye a Doran (1981). Según su definición original, los objetivos deben ser específicos, medibles, asignables, realistas y acotados en el tiempo. En una versión revisada de uso común en la educación se sustituyó “asignables” por “alcanzables” y “realistas” por “pertinentes”. Ésta es la versión que utilizamos en este informe.

Objetivos específicos. Un objetivo específico es aquel que enuncia claramente lo que se pretende lograr. Los objetivos son más alcanzables cuando son pocos y están bien definidos. Para los fines de seguimiento, se pueden establecer objetivos específicos para cada uno de los resultados de prosperidad, por ejemplo, mejorar el desempeño escolar y reducir las desigualdades.

Consideremos la meta 4.1 de la Agenda 2030 para el Desarrollo Sostenible. “De aquí a 2030, asegurar que todas las niñas y todos los niños terminen la enseñanza primaria y secundaria, que ha de ser gratuita, equitativa y de calidad y producir resultados de aprendizaje pertinentes y efectivos”. Esta meta claramente requiere una mayor especificidad, que se puede lograr desglosándola en varias metas separadas que definan los resultados de aprendizaje efectivos. Por otra parte, el marco de prosperidad educativa distingue entre los resultados y los factores que los producen. En consecuencia, los fundamentos para el éxito incluyen una educación gratuita, equitativa y de calidad.

Objetivos medibles. Es necesario formular los objetivos en términos absolutos y vincularlos con escalas invariables en el tiempo. Por ejemplo, un país podría establecer objetivos para reducir la prevalencia de los niños con bajo peso al nacer. “Bajo peso al nacer” se define como un peso al nacer inferior a 2500 g, independientemente de la edad gestacional (Organización Mundial de la Salud, 2010). La mayoría de las variables educativas utilizadas para denotar un resultado se definen de acuerdo con escalas estandarizadas o como variables dicotómicas. Si la escala es invariable en el tiempo, un parámetro de medición útil es el “tamaño de efecto”; por ejemplo, la media y la desviación estándar se fijan para un año base y no se reajusta su escala para obtener una nueva desviación estándar cada año o por ciclo de evaluación. La escala NAPLAN es ejemplar en el sentido de que abarca una amplia gama de edades y no se reajusta cada año. A menudo



es preferible usar medidas basadas en un criterio fijo, tal como el porcentaje de estudiantes que alcanzan el punto de referencia internacional intermedio de PIRLS o el nivel 2 o uno superior en PISA. Las medidas de ansiedad y depresión de los estudios OurSCHOOL de The Learning Bar (The Learning Bar Inc., 2009), que se usan en PISA para el Desarrollo establecen criterios fijos, invariables en el tiempo para los niveles moderados y severos de ansiedad y depresión.

Los objetivos establecidos conforme a un orden jerárquico de los países u otras jurisdicciones son inútiles porque incluso un pequeño cambio en un resultado puede tener un efecto dramático en la posición de un país. Asimismo, la posición de un país puede cambiar no porque hayan mejorado sus resultados, sino porque los de otros países se han deteriorado.

Objetivos alcanzables. Quizá lo más difícil del establecimiento de objetivos es que sean alcanzables. Los objetivos deben ser realmente alcanzables en un periodo específico y a la vez desafiantes para los educadores en todos los niveles del sistema escolar. Como ya se observó antes, las medidas educativas adquieren significado mediante comparaciones con un parámetro de referencia, por ejemplo, un grupo de jurisdicciones o un examen de tendencias en el tiempo. El valor de las evaluaciones internacionales como PIRLS, PISA y TIMSS para el establecimiento de objetivos es que proporcionan un contexto más amplio para discernir qué tan bien se están desempeñando los estudiantes de un país frente a los de otros países con un contexto socioeconómico comparable. Además, indican la magnitud de las desigualdades en el desempeño, las inequidades en la provisión y el grado en que los resultados de las escuelas difieren en distintos niveles socioeconómicos.

Cada escuela necesita recopilar datos regularmente sobre los resultados de prosperidad de cada uno de sus estudiantes. Los resultados de PISA pueden servir para determinar qué podría lograrse con una combinación adecuada de apoyo nacional y local, y la capacidad y voluntad del personal docente. Por ejemplo, los perfiles escolares de México que se presentan en las Figuras 12 y 13 indican el rango de puntuaciones de PISA y el rango de prevalencia de los estudiantes con puntuaciones bajas en todas las escuelas, en distintos niveles socioeconómicos. Si pasamos por alto algunos de los valores atípicos, los perfiles sugieren que las puntuaciones de PISA varían entre las escuelas en cada nivel socioeconómico en alrededor de 80 a 100 puntos y que la prevalencia de niños vulnerables tiene una variación aproximada de 40% a 60%. Los modelos lineales jerárquicos de los que se derivan los perfiles arrojan estimaciones de la variación en los resultados después de controlar el NSE de la escuela. El rango entre cuartiles es un método útil para determinar si una meta es alcanzable. En el caso de México, el rango entre cuartiles de puntuaciones PISA es de 36 puntos y el de prevalencia de estudiantes vulnerables, del 20%. Conforme a estos resultados, una meta alcanzable para una escuela podría ser aumentar las puntuaciones estandarizadas de la competencia lectora durante un periodo de cinco años en aproximadamente el 9% de una desviación estándar, y reducir los niveles promedio de vulnerabilidad en 5%. Estos objetivos representan alrededor de una cuarta parte de los rangos entre cuartiles.



Para dar seguimiento al progreso hacia estos objetivos es necesario recopilar datos sobre cada uno de los estudiantes a intervalos regulares, abarcando todos los grados que se atienden en la escuela, idealmente con indicadores que permitan vigilar el avance de cada niño en cuanto a los resultados de prosperidad. La información recabada a través de cuestionarios de contexto sobre los fundamentos para el éxito puede ayudar a las escuelas a dar seguimiento a sus avances en la provisión de un entorno propicio para la prosperidad de los estudiantes. La mayoría de las escuelas no pueden hacer esto solas. Por lo tanto, compete a los distritos escolares y a los gobiernos locales brindar las herramientas necesarias para que las escuelas den seguimiento a sus resultados. En muchos contextos, esto requiere cambiar de una evaluación para responsabilizar a las escuelas a una en la que los datos de las pruebas y los cuestionarios se usen como indicadores de tendencias para establecer objetivos y formular planes de mejoramiento escolar.

El modelo de prosperidad educativa exige que los administradores de la educación a nivel de ministerio establezcan objetivos basados en una serie de evaluaciones que abarquen el periodo de vida desde el nacimiento hasta la adolescencia tardía. También exige la medición de los fundamentos para el éxito en cada fase del desarrollo. Es necesario evaluar la fortaleza del sistema escolar utilizando una amplia variedad de indicadores y estableciendo los objetivos que corresponden. El seguimiento a este nivel de detalle requiere que se realicen evaluaciones periódicas, cada año o por lo menos cada dos años. Estas evaluaciones pueden basarse en muestras aleatorias de escuelas y estudiantes.

Las evaluaciones internacionales no son útiles para establecer objetivos, porque son demasiado infrecuentes y los resultados pueden variar de un ciclo a otro en formas que no reflejan los cambios reales en la estructura del sistema escolar o en la provisión de una educación de calidad. Por ejemplo, las puntuaciones en lectura PISA para México bajaron en 22 puntos, de 422 a 400, entre 2000 y 2003. Las puntuaciones en ciencias bajaron de 422 a 405 y las puntuaciones en matemáticas descendieron solo dos puntos, de 387 a 385. Vidal, Díaz y Jarquín (2004) sugirieron que la baja de las puntuaciones en lectura y en ciencias podrían atribuirse en parte a un aumento en la matrícula de estudiantes inscritos al menos en secundaria, del 52% en 2000 al 58% en 2003, pero esto no explicaría los resultados en el área de matemáticas. Por otra parte, el nivel socioeconómico promedio de la muestra PISA decreció ligeramente, de -1.17 a -1.11. Otra explicación plausible es que los cambios son principalmente atribuibles a un error de muestreo y medición. En 2009, las puntuaciones en lectura y ciencias en México habían regresado a sus niveles de 2000, en tanto que las del área de matemáticas habían subido casi al mismo nivel de las áreas de lectura y ciencias. Estos resultados acentúan el hecho de que las evaluaciones internacionales no son útiles para establecer objetivos. Si un país se propusiera elevar el desempeño académico en un 9% de una desviación estándar y reducir en un 5% los niveles de vulnerabilidad promedio, las fluctuaciones atribuibles a un error de medición y muestreo serían demasiado grandes como para representar un indicador fiable de los cambios en el desempeño del sistema. Sin embargo, las evaluaciones internacionales pueden respaldar la credibilidad de los resultados de las evaluaciones nacionales, y su validez puede corroborarse mediante estudios que las vinculen entre sí (Singer y Braun, 2018).



Objetivos pertinentes. El marco de prosperidad educativa identifica cuatro de cinco resultados de prosperidad para cada una de las seis fases de desarrollo. En este documento no se cuestiona su pertinencia, sino que ésta se examina de dos maneras. Primero, el objetivo debe considerarse importante para los actores de todos los niveles del sistema escolar, desde el Ministro y su personal, pasando por los administradores regionales, hasta los maestros, los padres y los estudiantes. Tal alineación del objetivo entre todas las partes interesadas del sistema escolar es decisiva para alcanzar el éxito. Segundo, el objetivo puede considerarse como un medio de comunicación. El mejoramiento de un sistema escolar requiere fortalecer los fundamentos para el éxito, lo que en la mayoría de los casos se facilita cuando se incrementa el financiamiento. Una declaración pública de un pequeño número de objetivos puede ayudar a establecer un “marco de entendimiento” de su importancia para una amplia variedad de partes interesadas, incluidos políticos, líderes comunitarios y donantes. La “pertinencia” de los resultados de prosperidad puede apuntarse haciendo referencia a sus beneficios económicos, sanitarios y sociales. Por ejemplo, las mejoras en la capacidad para leer y escribir se relacionan con los ingresos y la recaudación de impuestos, y con la reducción de las tasas de delincuencia, desempleo, dependencia del sistema de bienestar social, y costos de atención a la salud (Hanushek y Woessmann, 2015; Ross y Wu, 1995).

Objetivos acotados en el tiempo. Los estudios internacionales a gran escala como PIRLS y PISA tienen un ciclo relativamente largo. Además, los datos no se dan a conocer sino hasta, por lo menos, un año después de su recopilación. Estas evaluaciones son útiles para la planificación a largo plazo; sin embargo, para vigilar el progreso, los sistemas escolares también requieren obtener datos de seguimiento anualmente o incluso con mayor frecuencia. Esto es especialmente pertinente para monitorear las habilidades de lectura y escritura durante la primaria baja.

Un ejemplo de programa de seguimiento

La **Figura 22** presenta un ejemplo de un programa de seguimiento basado en la prosperidad educativa. Se trata de una versión revisada de un sistema de evaluación dirigida a poblaciones indígenas de Canadá (Willms, 2009b). Las flechas representan evaluaciones basadas en censos, y los rectángulos sólidos indican evaluaciones basadas en muestras. En la implementación de este diseño en varias jurisdicciones canadienses se utilizó la Evaluación Infantil Temprana para los niños de 4 y 5 años (flechas azules), la evaluación Confident Learners para preprimaria y primaria baja (flechas anaranjadas, salvo la evaluación de habilidad matemática) y los estudios OurSCHOOL para varios de los resultados de prosperidad y los fundamentos para el éxito (flechas moradas). Los datos de las mediciones basadas en muestras (rectángulos anaranjados y verdes) se recopilaron mediante evaluaciones provinciales. El seguimiento de la transición oportuna de los estudiantes a través de su trayectoria escolar corrió a cargo de las jurisdicciones escolares locales. Este diseño proporciona datos suficientes para monitorear los resultados de prosperidad y los fundamentos para el éxito que se bosquejan en este informe.

Nuevos rumbos



El marco de prosperidad educativa exige que se modifiquen las maneras en que los países, las jurisdicciones y las escuelas monitorean sus sistemas escolares. También exige nuevas maneras de usar los datos de seguimiento para formular la política educativa. Muchos países aplican pruebas de desempeño estudiantil en ciertos grados de primaria y secundaria. Además, algunas jurisdicciones solicitan a los estudiantes, maestros, padres y administradores de escuelas información sobre varios aspectos “contextuales” del sistema escolar. Por otro lado, muchos países participan en estudios internacionales a gran escala como PIRLS, TIMSS y PISA.

El desarrollo de la capacidad lectora de los niños al final de la escuela secundaria es el resultado acumulativo de su entorno durante el periodo prenatal y sus experiencias de aprendizaje desde el nacimiento. Los resultados de una evaluación infantil temprana nacional realizada en Uruguay en 2017 indican que aun antes de los que los niños aprendan formalmente a leer, sus habilidades lingüísticas y cognitivas varían sustancialmente. En este sentido, por lo menos una cuarta parte de los niños de 5 años en la fase de preparación para ingresar al primer grado estaban retrasados uno o varios años con respecto a sus pares de la misma edad.



Los niños que ingresan al primer grado sin la preparación básica para aprender a decodificar palabras están en mayor riesgo de tener una baja competencia lectora al final de la primaria baja. Además, la mayor parte del crecimiento de la competencia lectora ocurre durante los primeros años de la escuela primaria. Por ejemplo, los resultados de la evaluación nacional australiana de competencia lectora, denominada NAPLAN, revelan que esta capacidad varía de manera considerable entre los niños al final del grado 3, y que después la tasa de crecimiento se reduce anualmente hasta el grado 9. Los resultados de PISA muestran que después del grado 9 la tasa de crecimiento anual es de alrededor del 14% de una desviación estándar.

Los dos resultados antes mencionados, que las habilidades de los niños varían considerablemente cuando inician su educación formal y que aprender a leer durante la primaria es un factor crítico de su éxito a largo plazo, subrayan la importancia de cambiar el énfasis del seguimiento educativo a los primeros años de la infancia. Por lo tanto, la prosperidad educativa pone el acento en la recopilación de datos sobre las habilidades de los niños antes o poco después de su ingreso a la escuela, a intervalos regulares mediante evaluaciones individuales durante los tres primeros grados de escuela y a los 8 o 9 años, cuando se espera que hagan la transición de aprender a leer a leer para aprender. Es necesario hacer este cambio de énfasis en el seguimiento en todos los países de ingreso bajo, medio y alto. En su análisis de los resultados educativos en Sudáfrica, van der Berg, Spaull, Wills, Gustafsson y Kotzé (2016) afirman que “aprender a leer para comprender el significado y por placer en la fase básica es el objetivo más importante de la educación primaria” (p. 15).

Tradicionalmente, la mayoría de los sistemas de seguimiento han puesto énfasis en las aptitudes de los estudiantes para la lectura, las matemáticas y las ciencias. La prosperidad educativa requiere el monitoreo de una gama de resultados más amplia, que incluye la transición oportuna, la participación, la salud y el bienestar de los estudiantes. La medición de la transición oportuna de los estudiantes de los países de ingreso bajo y medio y de los estudiantes de las zonas de bajo ingreso de los países de alto ingreso tienen que abarcar la asistencia a la escuela y los niveles de escolaridad sucesivos completados por los estudiantes.

Para el seguimiento de los factores contextuales, muchos sistemas escolares han puesto énfasis en la estimación de las diferencias en el rendimiento académico y el uso de los resultados para responsabilizar a las escuelas. En numerosos casos, se recopilan datos sobre una serie de factores y características de los estudiantes para que los analistas puedan calcular el “valor agregado” por las escuelas después de ajustarlo a los antecedentes familiares de los estudiantes. Además, desde hace largo tiempo se ha realizado una “búsqueda de efectos escolares” –de los factores de la escuela y del aula asociados al valor agregado. Un principio de la prosperidad educativa es que ya sabemos qué factores contribuyen al éxito de los estudiantes, con base en más de treinta años de investigación sobre la eficacia de las escuelas. Éstos son los fundamentos para el éxito. La prosperidad educativa requiere que se cambie de la recopilación de datos sobre una amplia variedad de correlaciones potenciales del desempeño estudiantil a la recopilación de datos detallados sobre un conjunto pequeño de factores fundamentales. Estos últimos pueden usarse para evaluar la equidad de la provisión y establecer objetivos de mejoramiento del sistema.



En general, dichos cambios de énfasis se pueden representar por la recopilación de “indicadores prospectivos” para caracterizar el sistema escolar en vez de “indicadores retrospectivos”, que se usan principalmente para la rendición de cuentas. Los indicadores prospectivos pueden servir a nivel nacional y regional para formular políticas destinadas a mejorar los resultados, reducir las desigualdades en el desempeño de los estudiantes, asegurar la equidad de la provisión, asignar recursos y evaluar intervenciones. A nivel de escuela y de aula pueden ser útiles para orientar la práctica en clase, identificar a los estudiantes vulnerables, establecer objetivos de enseñanza y motivar la participación de los padres en distintas formas.

La prosperidad educativa es un modelo simple con un pequeño número de factores fundamentales. Como tal, proporciona un marco útil para establecer una visión de mejoramiento del sistema que puede ser compartida por las partes interesadas de todos los niveles del sistema escolar: estudiantes, padres, personal docente, directores y administradores de escuelas. El sistema mejorará si todos los participantes se empeñan incansablemente en construir cimientos sólidos con un único objetivo: brindar a todos los niños y todas las niñas la oportunidad de prosperar.



Referencias

- Adlof, S. M., H. W. Catts y J. Lee (2010). "Kindergarten predictors of second versus 8th grade reading comprehension impairments". *Journal of Learning Disabilities*, 43(4), 332-345.
- Alexander, K. L. y B. K. Eckland (1975). "Contextual effects in the high school attainment process". *American Sociological Review*, 4, 402-416.
- Alexander, K. L., J. Fennessey, E. L. McDill y R. J. D'Amico (1979). "School SES influences—Composition or context?". *Sociology of Education*, 52, 222-237.
- Anderson, L. W. (2004). *Increasing Teacher Effectiveness* (2nd ed). París: Instituto Internacional de Planeamiento de la Educación de la UNESCO (IIEP).
- Australian Curriculum, Assessment and Reporting Authority (2017). *NAPLAN Achievement in Reading, Writing, Language Conventions and Numeracy: National Report for 2017*. Sydney: Australian Curriculum, Assessment and Reporting Authority.
- Backhoff, E., A. Bouzas, E. Hernández y M. García (2007). *Aprendizaje y Desigualdad Social en México: Implicaciones de Política Educativa en el Nivel Básico*. México: Instituto Nacional para la Evaluación de la Educación. ISBN: 968 -5924-17-1
- Barnett, S., S. Ayers y J. Francis (2015). *Comprehensive Measures of Child Outcomes in Early Years: Report to the OECD*. París: OCDE.
- Ben-Arieh, A. e I. Frønes (2007). "Indicators of children's well being: What should be measured and why?". *Social Indicators Research*, 84(3), 249-250.
- Beswick, J.F., E. A. Sloat y J. D. Willms (2008). "Four educational myths that stymie social justice". *The Educational Forum*, 72(2), 115-128.
- Bhattacharjee, Y. (2015). "Baby brains: The first year". *National Geographic*, enero de 2015. Recuperado de: <http://ngm.nationalgeographic.com/2015/01/baby-brains/bhattacharjee-text>.
- Black, R.E., H. A. Lindsay, A. Zulfiqar, L. E. Bhutta, M. Caulfield, E. Majid, Colin Mathers, et al. (2008). "Maternal and child undernutrition: Global and regional exposures and health consequences". *Lancet* 371 (9608): 243-60.
- Black, M. M., P. Susan, L. C. H. Walker, C. T. Fernald, A. M. Andersen, C. L. D. DiGirolamo, C. McCoy, et al. (2017). "Early childhood development coming of age: Science through the life course". *Lancet*, 389(10064), 77-90.
- Boyce, W. T. y M. S. Kobor (2015). "Development and the epigenome: The 'synapse' of gene-environment interplay". *Developmental Science*, 18(1), 1-23.



- Boyce, W. T., M. B. Sokolowski y G. E. Robinson (2012). "Toward a new biology of social adversity". *PNAS*, 109(2), 17143-48.
- Bryk, A. S. y M. E. Driscoll (1988). "The high school community: Contextual influences and consequences for students and teachers". Madison: National Center on Effective Secondary Schools, University of Wisconsin.
- Caravolas, M., A. Lervåg, S. Defior, S. A. Málková y C. Hulme (2013). "Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies". *Psychological Science*, 24(8), 1398-1407.
- Center on the Developing Child (2007). *In Brief: The Science of Early Childhood Development*. Cambridge, MA: Center on the Developing Child. Recuperado de <http://developingchild.harvard.edu/resources/inbrief-science-of-eecd>.
- Creemers, B. P. M. y L. Kyriakides (2006). "Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model". *School Effectiveness and School Improvement*, 17, 347-366.
- Cynader, M. S. y B. J. Frost (1999). "Mechanisms of brain development: Neuronal sculpting by the physical and social environment". In D. Keating, C. Hertzman (Eds.), *Developmental Health and the Wealth of Nations* (pp. 153-84). Nueva York: Guilford.
- Deary, I. J., S. Strand, P. Smith y C. Fernandes (2007). "Intelligence and educational achievement". *Intelligence*, 35, 13-21.
- Doran, G. T. (1981). "There's a S.M.A.R.T. way to write management's goals and objectives". *Management Review. AMA FORUM*, 70(11), 35-36.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428-1446.
- Earl, S., F. Carden y T. Smutylo (2001). "Outcome mapping: Building learning and reflection into development programs". Ottawa: International Development Research Centre.
- Francis, D. J., S. E. Shaywitz, K. K. Stuebing, B. A. Shaywitz y J. M. Fletcher (1996). "Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis". *Journal of Educational Psychology*, 88, 3-17.
- Gonzales, E. (2016). *Calculating Standard Errors in International Large-Scale Studies*. Princeton: Educational Testing Service.
- Good, R.H., III, D. C. Simmons y E. J. Kame'enui (2001). "The importance of decision-making utility of a continuum of fluency-based indicators of foundational reading skills for 3rd grade high-stakes outcomes". *Scientific Studies of Reading*, 5, 257-288.



- Gough, P. B. y W. E. Tunmer (1986). "Decoding, reading and reading disability". *Remedial and Special Education*, 7(1), 6-10.
- Grek, S. (2009). "Governing by numbers: the PISA 'effect' in Europe". *Journal of Education Policy*, 24(1), 23-37. DOI: 10.1080/02680930802412669
- Hanushek, E. A. y L. Woessmann (2015). *The Knowledge Capital of Nations: Education and the Economics of Growth*. Cambridge y Londres: MIT Press.
- Hattie, J. (2009). *Visible Learning: A Synthesis of over 800 Meta-Analyses relating to Achievement*. Nueva York: Routledge.
- Hertzman, C. (1999). "The biological embedding of early experiences and its effects on health in adulthood". *Annals of the New York Academy of Sciences*, 896, 85-85.
- Heugh, K. A. (2013). "Multilingual education policy in South Africa constrained by theoretical and historical disconnections". *Annual Review of Applied Linguistics*, 33, 215.
- Kagan, S. L., E. Moore y S. Bredekamp (Eds.). (1995). *Reconsidering Children's Early Learning and Development: Toward Shared Beliefs and Vocabulary*. Washington, DC: National Education Goals Panel.
- Knudsen, E. I. (2004). "Sensitive periods in the development of the brain and behavior". *Journal of Cognitive Neuroscience*, 16(8), 1412-1425.
- Knudsen, E. I., J. J. Heckman, J. L. Cameron y P. Shonkoff (2006). "Economic, neurobiological and behavioral perspectives on building America's future workforce". *Proceedings of the National Academy of Sciences*, 103(27), 10155-10162.
- Kraemer, H. C., A. E. Kazdin, D. R. Offord, R. C. Kessler, P. S. Jensen y D. J. Kupfer (1997). "Coming to terms with the terms of risk". *Archives of General Psychiatry*, 54, 337-343.
- Kyriakides, L., C. Christoforou y Y. Charalambous (2013). "What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching". *Teaching and Teacher Education*, 36, 143-152.
- Lee, V. E. y J.B. Smith (1993). "Effects of school restructuring on the achievement and engagement of middle-grade students". *Sociology of Education*, 66 (julio) 64-187.
- Lee, V. E. y J.B. Smith (1995). "Effects of high school restructuring and size on early gains in achievement and engagement". *Sociology of Education*, 68(4), 241-270.
- Leppänen, U., P. Niemi, K. Aunola y J.-E. Nurmi (2004). "Development of reading skills among preschool and primary school pupils". *Reading Research Quarterly*, 39, 72-93.
- Levin, H. (2009). "The economic payoff to investing in educational justice". *Educational Researcher*, 28(1), 5-20.



- Lewin, K. (2015). *Educational Access, Equity and Development: Planning to Make Rights Realities*. París: UNESCO International Institute for Educational Planning (IIEP)..
- Lopez, A. Y. (2016). "Steps Towards a National Implementation of the Early Years Evaluation in Uruguay: Actors and Actions". Paper presented at the Inaugural Symposium of the Comparative and International Education Society, Scottsdale, AZ.
- Lorenz, M. O. (1905). "Methods of measuring the concentration of wealth". *Publications of the American Statistical Association*, 9(70), 209-219.
- Manyike, T. V. (2012). "A comparison of reading and writing proficiency in home language among Xitsonga speaking learners in South African primary township schools". *International Journal of Education Science*, 4(2), 143-152.
- Manyike, T. V. (2013). "Bilingual literacy or substantive bilingualism? L1 and L2 reading and writing performance among Grade 7 learners in three township schools Gauteng Province, South Africa". *Africa Education Review*, 10(2), 187-203. DOI: [10.1080/18146627.2013.812271](https://doi.org/10.1080/18146627.2013.812271).
- Martínez, F. y M. A. Díaz (2016). *México en PISA 2015*. México: Instituto Nacional para la Evaluación de la Educación. Recuperado de: <http://publicaciones.inee.edu.mx/buscadorPub/P1/D/316/P1D316.pdf>
- Matts, J. y J. Lachin (1988). "Properties of permuted-block randomization in clinical trials". *Control Clinical Trials*, 9, 327-344.
- McEwen, B. S. y H. M. Schmeck Jr. (1994). *The Hostage Brain*. Nueva York, NY, EUA: Rockefeller University Press.
- McClelland, M. M., F. J. Morrison y D. L. Holmes (2000). "Children at risk for early academic problems: The role of learning-related social skills". *Early Childhood Research Quarterly*, 15(3), 307-329.
- McClelland, M. M., C. C. Ponitz, E. E. Messersmith y S. Tominey (2010). "Self-regulation: The integration of cognition and emotion". In R. Lerner (Series Ed.) y W. Overton (Vol. Ed.), *Handbook of lifespan human development: Vol. 1. Cognition, biology and methods* (pp. 509-553). Hoboken, NJ: Wiley.
- Mislevy, R. J., A. E. Beaton, B. Kaplan y K. M. Sheehan (1992). "Estimation population characteristics from sparse matrix samples of item responses". *Journal of Educational Measurement*, 29, 133-161.
- Mueller, W. M. y T. L. Parcel (1981). "Measures of socioeconomic status: Alternatives and recommendations". *Child Development*, 52, 13-30.
- Murillo, F. J. y M. Román (2011). "School infrastructure and resources do matter: Analysis of the incidence of school resources on the performance of Latin American students". *School Effectiveness and School Improvement*, 22(1), 29-50.
- Nation, K. y M. J. Snowling (2004). "Beyond phonological skills: Broader language skills contribute to the development of reading". *Journal of Research in Reading*, Vol 27(4), 342-356.



- National Early Literacy Panel (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy. Recuperado de <http://lincs.ed.gov/publications/pdf/NELPReport09.pdf>
- Nonoyama-Tarumi, Y. y J.D. Willms (2010). "The relative and absolute risks of disadvantaged family background and low levels of school resources on student literacy". *Economics of Education Review*, 29(2), 214-224.
- OECD (2001). *Knowledge and Skills for Life – First Results from PISA 2000*. París: OCDE.
- OECD (2017). *PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science, Preliminary Version* Paris: OECD Publishing.
- Perfetti, C. A., N. Landi y J. Oakhill (2005). "The acquisition of reading comprehension skill". In M. J. Snowling y C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 227-247). Oxford, RU: Blackwell.
- Polo, D. S. Z., N. P. A. Araujo y J. C. R. Salceda (2017). "La concepción simple de la lectura en alumnos de 4º de primaria de una escuela fiscal de Quito". *Alteridad*, 12(1), pp. 228-235.
- Raudenbush, S. W. y J. D. Willms (1995). "The estimation of school effects". *Journal of Educational and Behavioural Statistics*. 20(4), 307-335.
- Raver, C. C., S. M. Jones, C. Li-Grining, F. Zhai, K. Bub y E. Pressler (2011). "The Chicago School Readiness Project's impact on low income preschoolers' preacademic skills: Self-regulation as a mediating mechanism". *Child Development*, 82, 362-378. doi:10.1111/j.1467-8624.2010.01561.x
- Rhode Island Kids Count (2005). *Getting Ready: Findings from the National School Readiness Indicators Initiative, a 17 State Partnership*. Providence, RI: Rhode Island Kids Count.
- Riehl, C. J. (2000). "The principal's role in creating inclusive schools for diverse students: A review of normative, empirical and critical literature on the practice of educational administration". *Review of Educational Research*, 70(1), 55-81.
- Ripoll, J. C., G. Aguado y A. P. Castilla-Earls (2014). "The simple view of reading in elementary school: A systematic review". *Revista de Logopedia, Foniatría y Audiología*, 34, 17-31.
- Robertson, D. y J. Symons (1996). *Do Peer Groups Matter? Peer Group versus Schooling Effects on Academic Attainment*. Londres: London School of Economics, Centre for Economic Performance.
- Rose, J. (2006). *Independent Review of the Teaching of Early Reading*. London, HMSO: Department for Education and Skills.
- Rosenshine, B. (2010). *Principles of Instruction*. International Academy of Education. Ginebra: Oficina Internacional de Educación de la UNESCO (OIE).



- Ross, C. E. y C. Wu (1995). "The links between education and health". *American Sociological Review*, 60 (5), 719-745.
- Rowan, R., S. W. Raudenbush y S. J. Kang (1991). "School climate in secondary schools: A multilevel analysis". In S.W. Raudenbush y J.D. Willms (Eds.), *Pupils, Classrooms and Schools: International Studies of Schooling from a Multilevel Perspective*. Nueva York: Academic Press.
- Rust, K. (1985). "Variance estimation for complex estimators in sample surveys". *Journal of Official Statistics*, 1(4), 381-397.
- Särndal, C. E., B. Swensson y J. Wretman (1992). *Model Assisted Survey Sampling*. Nueva York: Springer-Verlag.
- Scarborough, H. S. (1989). "Prediction of reading disability from familial and individual differences". *Journal of Educational Psychology*, 81(1), 101-108.
- Scarborough, H. S. (2001). "Connecting early language and literacy to later reading (dis)abilities: Evidence, theory and practice". In S. B. Neuman y D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97-110). Nueva York: Guilford Press.
- Schatschneider, C., J. M. Fletcher, D. J. Francis, C. D. Carlson y B. R. Foorman (2004). "Kindergarten prediction of reading skills: A longitudinal comparative analysis". *Journal of Educational Psychology*, 96(2), 265-282.
- Shonkoff, J. P. y D. A. Phillips (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC: National Academy Press.
- Simola, H. (2005). "The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education". *Comparative Education*, 41(4), 455-70.
- Singer, J. D. y H. I. Braun (2018). "Testing international education assessments". *Science*, 360(6384), 38-40. <http://science.sciencemag.org/content/360/6384/38>
- Snow, C. E., M. S. Burns y P. Griffin (1998). *Preventing Reading Difficulties in Young Children*. Washington DC: National Academy Press.
- Storch, S. and G. Whitehurst (2002). "Oral language and code-related precursors to reading: Evidence from a longitudinal structural model". *Developmental Psychology*, 38(6), 934-947. doi:10.1037//0012-1649.38.6.934
- The Learning Bar Inc. (2009). OurSCHOOL Student Survey [Measurement instrument]. Recuperado de <http://www.thelearningbar.com/>
- The Learning Bar Inc. (2011). Early Years Evaluation [Measurement instrument]. Recuperado de <http://www.thelearningbar.com/>



- The Learning Bar Inc. (2016). Confident Learners [Measurement instrument]. Recuperado de <http://www.thelearningbar.com/>
- Torgesen, J., S. Otaiba y M. Grek (2005). "Assessment and instruction for phonemic awareness and word recognition skills". In Catts, H. y Kamhi, A. (Eds.), *Reading and Language Disabilities* (pp. 112-145). Boston, MA: Allyn y Bacon.
- Tramonte, L. y J. D. Willms (2018). *New Measures for Comparative Studies of Low- and Middle-Income Countries*. Manuscrito presentado para publicación.
- Trzesniewski, K. H., T. E. Moffitt, A. Caspi, A. Taylor y B. Maughan, B. (2006). "Revisiting the association between reading ability and antisocial behavior: New evidence from a longitudinal behavior genetics study". *Child Development*, 77, 72-88.
- Naciones Unidas (2015). *Transformar nuestro mundo: la Agenda 2030 para el Desarrollo Sostenible*. <https://sustainabledevelopment.un.org/post2015/transformingourworld>
- UNESCO (2005). *Guidelines for Inclusion: Ensuring Access to Education for All*. París: UNESCO
- Instituto de Estadística de la UNESCO (UIS) (2017). *The Quality Factor: Strengthening National Data to Monitor Sustainable Development Goal 4*. Montreal: Instituto de Estadística de la UNESCO.
- UNICEF (2012). *School Readiness: A Conceptual Framework*. Nueva York, NY: Fondo de las Naciones Unidas para la Infancia.
- van der Berg, S., N. Spaull, G. Wills, M. Gustafsson y J. Kotzé (2016). "Identifying binding constraints in education: Synthesis report for the Programme to support Pro-poor Policy development (PsPPd)". University of Stellenbosch: Department of Economics.
- Vellutino, F. R. y D. M. Scanlon (1987). « Phonological coding, phonological awareness and reading ability: Evidence from a longitudinal and experimental study". *Merrill-Palmer Quarterly*, 33(3), 321-363.
- Verhoeven, L., J. van Leeuwe y A. Vermeer (2011). "Vocabulary growth and reading development across the school years". *Scientific Studies of Reading*, 1, 8-25.
- Vidal, R., M. A. Díaz y H. Jarquín (2004). *Resultados de las Pruebas PISA 2000 y 2003 en México: Habilidades para la Vida en Estudiantes de 15 Años*. México: Instituto Nacional para la Evaluación de la Educación. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P1/D/202/P1D202.pdf>
- Warwick, L. (2005). "Words to grow on". *Bulletin of the Centre of Excellence for Early Childhood Development*, 4(1), 2-4.
- Willms, J. D. (1986). "Social class segregation and its relationship to pupils' examination results in Scotland". *American Sociological Review*, 51, 224-241.



- Willms, J. D. y S. W. Raudenbush (1989). "A longitudinal hierarchical linear model for estimating school effects and their stability". *Journal of Educational Measurement*, 26(3), 209-232.
- Willms, J. D. y M. A. Somers (2001). "Family, classroom and school effects on children's educational outcomes in Latin America". *International Journal of School Effectiveness and Improvement*, 12(4), 409-445.
- Willms, J. D. (2011). "Measures of educational equality and equity: A methodological note for the INES Network for the Collection and the Adjudication of System-Level Descriptive Information on Educational Structures, Policies and Practices (NESLI)". Documento presentado en la reunión de la red NESLI de la OCDE celebrada en marzo de 2011 en los Países Bajos.
- Willms, J.D. (2003a). *Ten Hypotheses about Socioeconomic Gradients and Community Differences in Children's Developmental Outcomes*. Ottawa, ON: Applied Research Branch of Human Resources Development Canada.
- Willms, J.D. (2003b). "Literacy proficiency of youth: Evidence of converging socioeconomic gradients". *International Journal of Educational Research*, 39(3), 247-252.
- Willms, J. D. (2006). *Learning Divides: Ten Policy Questions about the Performance and Equity of Schools and Schooling Systems*. Montreal: Instituto de Estadística de la UNESCO.
- Willms, J. D. (2008). "The case for universal French instruction". *Policy Options*, 29(7), 91-96.
- Willms, J. D. (2009a, octubre). "Classroom diversity and Inclusion: The Educational Advantage". Presentación plenaria en el congreso Volviendo a Salamanca – Conferencia Mundial sobre Educación Inclusiva. Salamanca, España.
- Willms, J. D. (2009b). "Feasibility study for a First Nations assessment system". Ottawa: Indian and Northern Affairs Canada.
- Willms, J. D. (2010). "School composition and contextual effects on student outcomes".
- Willms, J. D., L. Tramonte, J. Duarte y S. Bos (2012). *Assessing Educational Equality and Equity with Large-Scale Assessment Data: Brazil as a Case Study*. Washington: Banco Interamericano de Desarrollo.
- Willms, J. D. y L. Tramonte (2018). *The Measurement and Use of Socioeconomic Status in Educational Research*. Manuscrito presentado para publicación.
- Willms, J. D. (2018a). *Educational Prosperity*. Fredericton, NB: The Learning Bar Inc.
- Willms, J. D. (2018b). "Educational Prosperity: An assessment strategy for supporting student learning in low-income countries". In D. A. Wagner, S. Wolf y R. F. Boruch (Eds.), *Learning at the Bottom of the Pyramid: Science, Measurement and Policy in Low-Income Countries*. París: UNESCO-IIEP.



Banco Mundial (2018a). Datos sobre las cuentas nacionales del Banco Mundial y archivos de datos sobre las cuentas nacionales de la OCDE. Washington, D. C.: Banco Mundial. Recuperado de <https://data.worldbank.org/indicator/NY.GNP.MKTP.KD>

Banco Mundial (2018b). *Informe sobre el desarrollo mundial 2018: Learning to Realize Education's Promise*. Washington, DC: Banco Mundial. DOI:10.1596/978-1-4648-1096-1

Organización Mundial de la Salud (OMS) (2010). "Disorders related to short gestation and low birth weight, not elsewhere classified". *International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) version for 2010*. Recuperado de: <http://apps.who.int/classifications/icd10/browse/2010/en#/P07>

Zimmer, R. W. y E. F. Toma (1997). *Peer Effects in Private and Public Schools: A Cross-Country Empirical Analysis*. Lexington: University of Kentucky.



Apéndice 1. Tasas de crecimiento de la competencia lectora, PISA 2000-2015

	Nivel de referencia (2000)	(ES)	Crecimiento anual	(ES)	NSE	(ES)
Australia	517.1	(2.4)	-1.49	(0.24)	43.8	(1.0)
Austria	500.1	(3.9)	-1.79	(0.44)	46.0	(1.8)
Bélgica	507.2	(3.0)	-0.81	(0.32)	48.2	(1.1)
Canadá	520.0	(2.8)	-0.62	(0.25)	33.9	(1.5)
República Checa	488.5	(4.4)	0.37	(0.39)	49.0	(1.6)
Dinamarca	491.4	(3.0)	-0.74	(0.27)	37.3	(1.7)
Finlandia	548.3	(2.5)	-2.45	(0.30)	31.7	(1.6)
Francia	521.2	(8.8)	-2.00	(0.84)	50.7	(3.4)
Alemania	490.0	(5.5)	0.80	(0.70)	45.4	(3.2)
Grecia	478.1	(4.1)	-0.16	(0.31)	34.4	(0.6)
Hungría	491.4	(5.2)	-0.12	(0.52)	47.9	(2.1)
Islandia	487.6	(3.7)	-1.87	(0.33)	26.7	(0.9)
Irlanda	524.0	(3.6)	-1.00	(0.40)	38.0	(0.6)
Italia	483.8	(6.2)	0.30	(0.48)	32.1	(1.0)
Japón	504.6	(2.2)	2.20	(0.72)	37.9	(3.6)
Corea	530.1	(7.2)	1.04	(0.89)	32.0	(4.1)
Luxemburgo	465.6	(2.9)	1.00	(0.19)	40.4	(1.0)
México	456.5	(37.7)	-1.16	(2.54)	26.6	(13.0)
Países Bajos	520.4	(6.2)	-1.78	(0.58)	38.9	(1.1)
Noruega	486.1	(6.1)	-0.07	(0.69)	37.3	(1.9)
Nueva Zelanda	523.8	(3.0)	-1.10	(0.26)	47.7	(1.5)
Polonia	496.6	(4.0)	2.18	(0.51)	40.2	(1.4)
Portugal	491.3	(2.4)	1.14	(0.19)	31.5	(1.7)
República Eslovaca	478.1	(3.6)	-0.99	(0.61)	47.7	(2.3)
España	486.8	(13.0)	0.96	(1.27)	29.5	(1.0)
Suecia	506.1	(2.7)	-1.65	(0.51)	39.8	(1.0)
Suiza	504.8	(3.0)	-0.92	(0.39)	42.1	(2.1)
Reino Unido	511.6	(11.6)	-1.80	(1.95)	43.0	(2.0)
Estados Unidos	494.6	(6.7)	-0.20	(0.55)	40.5	(3.5)



Apéndice 2. Estadísticas clave para formular la política educativa

Estimación de estadísticas descriptivas en estudios internacionales a gran escala

Ponderaciones replicadas. Los estudios internacionales usan una o dos técnicas para considerar el diseño de la muestra. En PIRLS y TIMSS se usa la técnica *jackknife*; y en PISA, la de replicación repetida balanceada (BRR, por sus siglas en inglés) (véase Rust, 1985). El conjunto de datos de PISA 2015, por ejemplo, es una ponderación general del diseño del estudio y un conjunto de 80 ponderaciones BRR. Las medidas estadísticas como la media y la desviación estándar se pueden calcular con la ponderación general del diseño; pero para obtener estimaciones adecuadas del error estándar de una estadística, hay que usar las ponderaciones BRR. Cuando la estadística implica una puntuación de una prueba, las ponderaciones BRR se usan junto con los valores plausibles.

Valores plausibles. El contenido de las pruebas para estudios internacionales se organiza en varios folletos, y a cada estudiante se le da uno. En PISA 2015, por ejemplo, las pruebas constaron de 810 minutos de ítems de lectura, ciencias, matemáticas y resolución colaborativa de problemas, y cada estudiante completó únicamente 120 minutos del contenido de las pruebas. Cada prueba está concebida de manera que incluya una dimensión de la competencia asociada con todo el conjunto de ítems. Las técnicas estadísticas que se utilizan en PISA proporcionan estimaciones de las distribuciones de la competencia en las pruebas, así como la probabilidad de una respuesta correcta de cada alumno para cada ítem, si han completado todos los ítems (Mislevy, Beaton, Kaplan, Sheehan, 1992). El conjunto de datos del estudiante incluye 10 valores plausibles. Estos valores se seleccionan al azar entre la distribución estimada de habilidades de los estudiantes con un patrón de respuesta y antecedentes familiares similares. En esencia, la técnica arroja 10 estimaciones o *valores plausibles* que indican qué tan bien podría haberse desempeñado cada estudiante si hubiera respondido toda la prueba.

Las estadísticas como la puntuación media, la desviación estándar y los coeficientes de regresión se calculan en la forma habitual, utilizando la ponderación general del diseño. El error estándar de la estadística de interés, si no implica valores plausibles, viene dado por dos puntos:

$$SE_{\varepsilon} = \sqrt{\frac{\sum_{r=1}^R (\varepsilon_r - \varepsilon_0)^2}{R(1 - 0.5)^2}}$$

donde ε_0 es la estadística de interés, calculada utilizando la ponderación de la muestra completa, ε_r es la estadística de interés, calculada utilizando la ponderación replicada, r , y R es el número de replications, que para PISA es 80.

El error estándar de la media, si implica valores plausibles, viene dado por:



$$SE_{\varepsilon} = \sqrt{\left[\sum_{p=1}^P \left(\frac{\sum_{r=1}^R (\varepsilon_{r,p} - \varepsilon_{0,p})^2}{R(1-0.5)^2} \right) * \frac{1}{P} \right] + \left[\left(1 + \frac{1}{P} \right) \frac{\sum_{p=1}^P (\varepsilon_{0,p} - \bar{\varepsilon}_{0,P})^2}{P-1} \right]}$$

donde $\varepsilon_{r,p}$ es la estadística de interés, calculada utilizando el valor plausible, p , y la ponderación replicada, r , y P es el número de valores plausibles, que para PISA es 10.

Estas fórmulas parecen más bien complicadas para los lectores no especializados en estadística. No obstante, si construimos una matriz de 80 x 1 de las ponderaciones replicadas para las estadísticas que no implican valores plausibles, o una matriz de 80 x 10 (80 BRR x 10 PV) para las estadísticas que implican valores plausibles, se puede aplicar una fórmula simple. Este método se puede utilizar con cualquier estadística, por ejemplo, una estimación de asimetría, un índice de segregación o coeficientes de regresión.

Gonzales (2016) escribió un documento útil que establece las fórmulas para la mayoría de los principales estudios internacionales.

[1] Media. La fórmula para calcular la media es: $\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$, donde n es el tamaño de la muestra y Y_i es la puntuación del estudiante i . La media se calcula utilizando la ponderación general del diseño y el error estándar se calcula utilizando una de las dos fórmulas anteriores, dependiendo de si la media corresponde o no a la puntuación de una prueba.

[2] Desviación estándar. La fórmula para calcular la desviación estándar es: $SD = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}}$, donde n es el tamaño de la muestra, y_i es la puntuación del estudiante i , y \bar{y} es la media. La desviación estándar se calcula utilizando la ponderación general del diseño y el error estándar se calcula utilizando una de las dos fórmulas anteriores, dependiendo de si la desviación estándar corresponde o no a la puntuación de una prueba.

[3] Asimetría. La fórmula para calcular la asimetría es: $G = \frac{\sqrt{n(n-1)}}{n-2} \left[\frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^3}{\left(\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right)^{\frac{3}{2}}} \right]$

donde n es el tamaño de la muestra, y_i es la puntuación para el estudiante i , y \bar{y} es la media. La asimetría se calcula utilizando la ponderación general del diseño y el error estándar se calcula utilizando una de las dos fórmulas anteriores, dependiendo de si la asimetría corresponde o no a la puntuación de una prueba.

[4] Porcentaje vulnerable. La fórmula para calcular el porcentaje vulnerable es: $P = \frac{\sum_{i=1}^n D_i}{n}$, donde n es el tamaño de la muestra y D_i es la puntuación dicotómica (0 ó 1) del estudiante i . El porcentaje vulnerable se calcula utilizando la ponderación general del diseño y el error estándar se calcula utilizando una de las dos



fórmulas anteriores, dependiendo de si el porcentaje variable corresponde o no a la puntuación de una prueba.

Gradientes socioeconómicos

Los gradientes socioeconómicos se calculan con un modelo de regresión de MCO que incluye SES y SES²:

$$Y_i = \beta_0 + \beta_1 SES_i + \beta_2 SES_i^2 + r_i$$

donde Y_i es la puntuación en lectura del estudiante i y SES_i es el NSE del estudiante i . Los parámetros, r_i , son los residuos del nivel del estudiante, es decir, la desviación de sus puntuaciones con respecto a la línea de regresión. La siguiente ecuación sirve para calcular cuatro estadísticas:

[5] Nivel del gradiente: La intersección β_0 , es la puntuación prevista de un estudiante con una puntuación de 0 para el NSE. La variable SES se puede “centrar” en un valor específico, pero generalmente conviene centrarla en el promedio del país o, en el caso de los análisis PISA, en el promedio de la OCDE.

[6] Pendiente del gradiente. El coeficiente β_1 , es la pendiente del gradiente socioeconómico. La hipótesis de gradiente es: $H_0: \beta_1 = 0$; $H_1: \beta_1 \neq 0$. La significación estadística de β_1 depende de la magnitud del error estándar y se evalúa mediante una prueba t con n-1 grados de libertad.

[7] Rendimientos decrecientes. El coeficiente β_2 , es la pendiente de SES cuadrado. Cuando β_2 resulta negativo, hay menores efectos en los resultados que se asocian con SES en niveles socioeconómicos más altos. La hipótesis de rendimientos decrecientes es: $H_0: \beta_2 = 0$; $H_1: \beta_2 \neq 0$. La significación estadística de β_2 depende de la magnitud del error estándar y se evalúa mediante una prueba t con n-1 grados de libertad.

[8] Fuerza del gradiente. La fuerza del gradiente, denotada por R^2 , es la proporción de varianza en la medida de resultados atribuibles al nivel socioeconómico. Es la diferencia entre la varianza de Y_i y la de los residuos, r_i , expresada como una fracción de la varianza de Y_i .

Para obtener el error estándar de estas estadísticas, hay que hacer el análisis de regresión 80 veces (uno por cada BRR) cuando el resultado es una variable que no corresponde a la puntuación de una prueba. Si el resultado es la puntuación de una prueba, hay que hacer el análisis de regresión 800 veces, una por cada BRR y cada valor plausible. Como ya se mencionó, se necesita una matriz de 80 x 1 o de 80 x 10 para cada una de las estadísticas.

Índices de inclusión

[9] Índice de inclusión vertical. La inclusión vertical es la proporción de la variación en la puntuación de los resultados *dentro* de las escuelas.

$$V = \frac{\sigma^2}{\tau}$$



La inclusión vertical se calcula ajustando un modelo de regresión lineal (HLM) jerárquica “nulo” a los datos de resultados:

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim NID(0, \sigma^2)$$

$$\beta_{0j} = \gamma_{00} + U_{0j} \quad U_{0j} \sim NID(0, \tau)$$

En el HLM, la varianza de los términos del error del nivel 1, que es la varianza *dentro* de las escuelas, se llama σ^2 . la varianza de los términos del error del nivel 2, que es la varianza *entre* escuelas, se llama τ .

$$Var(\varepsilon_{ij}) = \sigma^2 \quad Var(U_{0j}) = \tau$$

[10] Índice de inclusión horizontal. La inclusión horizontal es la proporción de la variación en el nivel socioeconómico *dentro* de las escuelas:

$$H = \frac{\sigma^2}{\tau}$$

Se calcula con un modelo nulo en el que el NSE es la variable de resultados.

Modelo de gradientes multinivel

El modelo de gradientes multinivel básico tiene dos niveles:

$$(1) Y_{ij} = \beta_{0j} + \beta_{1j}SES_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim NID(0, \sigma^2)$$

$$(2) \beta_{0j} = \gamma_{00} + U_{0j} \quad U_{0j} \sim NID(0, \tau_0)$$

$$\beta_{1j} = \gamma_{10} + U_{1j} \quad U_{1j} \sim NID(0, \tau_1)$$

donde Y_{ij} es la puntuación de los resultados del estudiante i en la escuela j , SES_{ij} es la puntuación del nivel socioeconómico del estudiante i en la escuela j , y ε_{ij} son los residuos que supuestamente tienen una distribución normal con una media de cero y una varianza de σ^2 . Los niveles de los gradientes son β_{0j} , esto es, las puntuaciones previstas de los estudiantes con un nivel socioeconómico de 0. Por lo tanto, se les denomina “media ajustada al nivel socioeconómico”.

[11] Gradiente intraescolar. Las pendientes del gradiente intraescolar son β_{1j} . Se modelizan en el nivel 2 como una pendiente promedio, γ_{10} , más un residuo de la pendiente promedio U_{1j} .

[12] Índice de gradientes convergentes. La varianza de la media ajustada al nivel socioeconómico, β_{0j} , es la varianza de U_{0j} , que es τ_0 . De manera similar, la varianza de los gradientes intraescolares, β_{1j} , es la varianza de U_{1j} , que es τ_1 .



El HLM proporciona una estimación de la covarianza de U_{0j} y U_{1j} . La correlación de la media ajustada al nivel socioeconómico y las pendientes, que aquí se denomina índice de gradientes convergentes, está dada por:

$$CGI = \frac{Cov(U_0, U_1)}{\tau_0 \tau_1}$$

[13] Gradiente interescolar. La gradiente interescolar, que es una regresión de la media de resultados escolares en el nivel socioeconómico medio escolar, también se puede calcular con un modelo de dos niveles:

$$(1) Y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim NID(0, \sigma^2)$$

$$(2) \beta_{0j} = \gamma_{00} + \gamma_{01} \overline{SES}_j + U_{0j} \quad U_{0j} \sim NID(0, \tau_0)$$

La pendiente interescolar es γ_{01} .

[14] Efecto de composición asociado al nivel socioeconómico medio. El efecto de la composición de la escuela asociado al NSE medio se calcula agregando este nivel a la ecuación para la media escolar ajustada en el nivel 2:

$$(1) Y_{ij} = \beta_{0j} + \beta_{1j} SES_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim NID(0, \sigma^2)$$

$$(2) \beta_{0j} = \gamma_{00} + \gamma_{01} \overline{SES}_j + U_{0j} \quad U_{0j} \sim NID(0, \tau_0)$$

$$\beta_{1j} = \gamma_{10} + U_{1j} \quad U_{1j} \sim NID(0, \tau_1)$$

El cálculo de γ_{01} es el efecto de la composición de la escuela asociado al nivel socioeconómico medio.

[15] Efecto de composición de la escuela asociado a la variabilidad del estudiantado. El efecto de composición de la escuela asociado a la variabilidad del estudiantado se estima de la misma manera que el efecto de composición asociado al NSE medio: se agrega la desviación estándar del NSE a la ecuación del nivel 2. El efecto de la composición de la escuela asociado a la variabilidad del estudiantado es γ_{01} .