

Las evaluaciones estandarizadas del aprendizaje y la mejora de la calidad educativa

Karla María Díaz López*

Cecilia Osuna Lever**

Recibido: 15/04/2016

Aceptado: 30/06/2016



Resumen

Una característica de los actuales sistemas educativos es la importancia que otorgan a la evaluación como vía para determinar la calidad de la educación. El presente trabajo comprende una reflexión teórica centrada en describir y analizar el alcance e impacto de las evaluaciones estandarizadas del aprendizaje, tanto en ámbito internacional como latinoamericano, haciendo puntual referencia al caso de la evaluación del aprendizaje en México. Asimismo, se analiza la contribución de este tipo de evaluaciones como insumo para la mejora de la calidad educativa lo que finalmente deriva en el planteamiento de una serie de desafíos como respuesta tentativa a la interrogante, ¿la implementación y resultados de evaluaciones del aprendizaje contribuyen en la mejora de la calidad educativa?

Palabras clave

Evaluación del aprendizaje, rendimiento académico, calidad educativa, pruebas estandarizadas.

The standardized assessment learning and the improvement of the educational quality

Abstract

A feature of the present education system is the importance given to evaluation as a way to determine the quality of education. This work includes a theoretical reflection focused on describing and analyzing the scope and impact of standardized assessments of learning, both internationally and Latin American, making timely reference to the case of the evaluation of learning in Mexico. In addition, the contribution of these assessments as input for improving educational quality analyzes which ultimately leads to the approach of a series of challenges as tentative answer to the question: Is the implementation and results of learning assessments contribute to the improving educational quality?

Key words

Assessment of learning, academic performance, educational quality, standardized tests.

* Centro de Enseñanza Técnica y Superior CETYS Universidad, México. E-mail: Karla.diaz@cetys.mx

** Centro de Enseñanza Técnica y Superior CETYS Universidad, México. E-mail: cecilia.osuna@cetys.mx

1. Introducción

A principio de la década de los noventa del siglo pasado, a nivel internacional, se presentaron diversos cambios en materia de política educativa; por ejemplo, el énfasis dejó de ser la cobertura y expansión y se comienza a otorgar mayor importancia al funcionamiento del sistema. Se pasó de un enfoque centrado en la «cantidad» hacia uno cuyo centro sería elevar la «calidad» de procesos y resultados (Casassus, 2007). Por tanto, las reformas educativas que proliferaron en dicha década atendieron algunos aspectos asociados a la «calidad educativa» tales como, la relación entre costes y beneficios alcanzados, la adecuación a estándares preestablecidos, la consecución de objetivos, la descentralización y evaluación de la calidad de los aprendizajes (Morillejo, Reboloso, Pozo & Fernández, 1999). Así, en nuestros días, uno de los elementos que caracterizan a los sistemas educativos es la importancia que otorgan a la evaluación como vía para determinar la calidad de la educación, por lo que ésta trascendió los discursos y prácticas educativas y pasó a formar parte de la agenda pública (Van Dijk, 2009).

En el presente artículo nos centramos en describir y analizar algunas cuestiones particulares, tales como, el alcance e impacto de las evaluaciones estandarizadas del aprendizaje, tanto en ámbito internacional, latinoamericano y en particular el caso de México. Asimismo, se analiza la contribución de este tipo de evaluaciones como insumo para la mejora de la calidad educativa lo que finalmente conlleva al planteamiento de una serie de desafíos que podrían dar respuesta a la interrogante, ¿la implementación y resultados de evaluaciones del aprendizaje contribuyen en la mejora de la calidad educativa?

En Estados Unidos, tres acontecimientos históricos marcaron la pauta para el inicio de las actividades de evaluación a gran escala en el ámbito educativo. En primera instancia, en 1958 se creó la International Association for the Evaluation of Educational Achievement (IEA), hoy con sede en Amsterdam, Holanda, institución dedicada a promover y realizar estudios internacionales de evaluación educativa. En segunda instancia, en 1965 se aprobó el Primary and Secondary Education Act, proyecto que tenía como objetivo asegurar la evaluación de los programas educativos puestos en práctica (Degenhart, 1990).

Posteriormente, en 1968, el *Informe Coleman* generó una serie de debates y expectativas entre los políticos y la sociedad respecto al desempeño de las instituciones escolares (Tiana, 1996). A estos acontecimientos, se suma el proyecto denominado Indicadores Internacionales de la Educación (INEES) instrumentado, a finales de los años ochenta del siglo pasado, por la Organización para la Cooperación y el Desarrollo Económico (OCDE); que a la fecha continúa vigente, y cuyo objetivo es generar información que permita comparar el funcionamiento, la eficacia y evolución de los sistemas educativos. Estas acciones detonaron el interés en la evaluación de los sistemas educativos en el mundo y en los países latinoamericanos, donde las prácticas evaluativas dejaron de concebirse como una temática de indagación propia de investigadores y académicos, convirtiéndose en un asunto de interés político y social (Tiana, 1996).

En las últimas dos décadas se ha suscitado un creciente interés en la evaluación estandarizada del aprendizaje, misma que ha trascendido el ámbito educativo, convirtiéndose en un asunto de alcances políticos, económicos e incluso sociales. El principal indicador del aprendizaje de los estudiantes es el rendimiento académico, y éste, a su vez, se concibe como un indicador del funcionamiento de los sistemas educativos, lo que ha propiciado un desarrollo sin precedentes de evaluaciones estandarizadas a gran escala (González,

Caso, Díaz & López–Ortega, 2012; Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación [LLECE], 2015). Así, en el ámbito internacional, diversas organizaciones instrumentan evaluaciones, que involucran la aplicación de pruebas objetivas, enfocadas en la medición de conocimientos, habilidades, dominios específicos y competencias de estudiantes, destacándose las siguientes:

- El Estudio Internacional de Tendencias en Matemáticas y Ciencias (TIMSS, por sus siglas en inglés), se centra en evaluar conocimientos en matemáticas y ciencias de los estudiantes inscritos en los grados cuarto y octavo de la educación básica; mientras que el Estudio Internacional de Progreso en Comprensión Lectora (PIRLS, por sus siglas en inglés) mide las tendencias en el rendimiento en comprensión lectora de los estudiantes del cuarto curso de educación primaria, vale referir que en estos estudios han participado más de 50 países. Ambos proyectos son instrumentados por la Asociación Internacional para la Evaluación del Logro Educativo (IEA, por sus siglas en inglés).¹
- El Programa para la Evaluación Internacional de Estudiantes (PISA, por sus siglas en inglés), que instrumenta la Organización para la Cooperación y el Desarrollo Económico (OCDE),² se centra en identificar el nivel de habilidades necesarias que han adquirido los estudiantes para participar plenamente en la sociedad, centrándose en dominios claves como Lectura, Ciencias y Matemáticas. Mide si los estudiantes tienen la capacidad de reproducir lo que han aprendido, de transferir sus conocimientos y aplicarlos en nuevos contextos académicos y no académicos, de identificar si son capaces de analizar, razonar y comunicar sus ideas efectivamente, y si tienen la capacidad de seguir aprendiendo durante toda la vida. Para PISA, esos dominios están definidos como competencia (literacy) científica, lectora o matemática.
- En Latinoamérica, destacan los Estudios Comparativos y Explicativos³ realizados por el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), el cual se fundó en 1994 como una red de unidades de medición y evaluación de la calidad de los sistemas educativos de la región, coordinado técnicamente por un equipo de la Oficina Regional de Educación para América Latina y el Caribe (OREALC/UNESCO, Santiago). En estos estudios se aplican pruebas estandarizadas de evaluación de aprendizajes en distintas áreas a estudiantes de educación primaria. Conviene referir que LLECE produce información y conocimiento que retroalimentan la política educativa, por medio de sus reportes de estudios que reflejan el estado de la calidad de la educación en la región. En estos estudios participan los siguientes países: Argentina, Brasil, Chile, Colombia, Costa Rica, Ecuador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Perú, República Dominicana y Uruguay.

1 TIMSS se aplicó por primera vez en 1995, a partir de entonces se realiza cada cuatro años, participando en la prueba de 2007 un total de 59 naciones; en tanto que PIRLS 2011 ha sido el tercer estudio internacional, tras PIRLS 2001 y 2006.

2 En cada aplicación PISA se ha concentrado en evaluar una temática concreta: la lectura (en 2000), las matemáticas (en 2003) y las ciencias (en 2006); siendo la resolución de problemas un área temática especial en PISA 2003. El programa llevó a cabo una segunda fase de evaluaciones en el 2009 (lectura), 2012 (matemáticas) y 2015 (ciencias).

3 En 1997 se aplicó el primer estudio, denominado PERCE, en el cual participaron 13 países. Esta pesquisa evaluó matemática y lectura en tercer y cuarto grados de educación primaria. El segundo estudio (SERCE) se aplicó nueve años más tarde, en 2006, y evaluó las áreas de Matemáticas y Lectura en tercer y sexto grado, y Ciencias Naturales en sexto grado. El tercer estudio se aplicó en 2013, evaluando las mismas áreas y grados que el SERCE.

A su vez, cabe mencionar la labor realizada por otras instancias enfocadas en la evaluación educativa de la región, así, destacan las siguientes instancias:

- El Grupo de Trabajo Estándares y Evaluaciones (GTEE) conforma uno de los grupos del Programa de Promoción de la Reforma Educativa en América (PREAL). El GTEE es una red regional de especialistas que se constituyó en 1998 con el fin de promover los avances en materia de evaluación educativa, para proponer recomendaciones a nivel regional, nacional o local.
- El Instituto de Evaluación y Asesoramiento Educativo (IDEA) es una red de trabajo e investigación en la que participan España, Chile, México, Brasil y Puerto Rico. Se creó en 1999 con el propósito de colaborar con las instituciones educativas de los países participantes en el conocimiento y la mejora de la calidad de los procesos de enseñanza–aprendizaje, para lo cual realiza evaluaciones externas e independientes.

Asimismo, cabe mencionar que desde la década de los noventa del siglo pasado, algunos países de la región como Chile, Argentina, Brasil, Bolivia y México cuentan con pruebas estandarizadas para medir el aprendizaje de los estudiantes. En el siguiente apartado se describe el caso de México.

2. Las evaluaciones del aprendizaje en México

Con la federalización de la educación básica en 1992, el sistema educativo nacional se modificó, lo que implicó realizar evaluaciones en forma conjunta con los estados y el gobierno federal. Para ello, se estableció un órgano central de evaluación y un Comité Técnico Interestatal de Evaluación, y se impulsó la creación de unidades de evaluación. El propósito fue capacitar y profesionalizar las actividades de evaluación en temas particulares como la construcción de reactivos, la identificación de factores de contexto y el análisis de la información (Álvarez, Guerrero & De la Cera, 1994). Para ello, se conformó un banco nacional automatizado de reactivos, con lo que se pretendía que cada estado elaboraría sus propias pruebas.

En 1992, en la *Ley General de Educación* se concibió la evaluación como una función sustantiva y estratégica para elevar la calidad de la educación (Zorrilla, 2003). En consecuencia, en la década de los noventa, la Secretaría de Educación Pública (SEP) se encargó del diseño y la coordinación de múltiples evaluaciones a nivel nacional. Se inicia de este modo, el desarrollo de un amplio conjunto de instrumentos de evaluación de aprendizajes, con distintos propósitos, particularmente en educación básica, en las que destacan las pruebas denominadas Estándares Nacionales (aplicadas entre 1998 y 2003), así como el Factor de Aprovechamiento Escolar, enmarcado en el Programa de Carrera Magisterial, en 1994. A estos esfuerzos se sumó, ese mismo año, la creación del Centro Nacional de Evaluación para la Educación Superior (CENEVAL). Con estas acciones, se iniciaron en México evaluaciones regulares de los conocimientos de los estudiantes. Los datos recogidos por las evaluaciones nacionales comenzaron a hacerse públicos en el año 2000; no obstante, la evaluación del aprendizaje se instrumentó casi tres décadas después que en los países más desarrollados, y ésta se empezó a realizar sin que sus resultados se consideraran para la toma de decisiones de política pública.

En el Acuerdo de la Alianza por la Calidad de la Educación 2008–2009 se estipuló la necesidad de articular el sistema nacional de evaluación conjuntando las instancias, procesos y procedimientos existentes. Ello derivó en una expansión acelerada de distintas iniciati-

vas en esta materia. En 2002, se creó en México, el Instituto Nacional para la Evaluación de la Educación (INEE), con fundamento en la mejora la educación a través de evaluaciones de la calidad del sistema educativo nacional y de los factores que en ella inciden. Este organismo fue creado con la consigna de difundir, de manera transparente y oportuna, los resultados de las evaluaciones, y con base en éstos, apoyar la toma de decisiones para la mejora pedagógica en las escuelas; todo ello con el compromiso de rendir cuentas a la sociedad sobre el rendimiento de los estudiantes.

Durante los últimos quince años se usaron, principalmente, dos instrumentos para evaluar el aprendizaje alcanzado por los alumnos en las escuelas mexicanas de educación básica y media superior: los Exámenes Nacionales del Logro Académico en Centros Escolares (ENLACE) propuesto por la SEP y los Exámenes de la Calidad y el Logro Educativo (EXCALE) instrumentado por el INEE, que suscitaron discusiones respecto a su utilidad para promover la mejora de la calidad educativa, tanto en los medios especializados como entre el público en general (Martínez Rizo, 2015). Debido a su importancia e impacto, en las siguientes líneas, se describen brevemente.

Los Exámenes para la Calidad y Logro Educativos (EXCALE) consisten en una evaluación nacional que se aplicó por primera vez en 2005,⁴ a una muestra representativa de estudiantes de escuelas públicas y privadas, de educación básica de todo el país. Su plan de evaluación se basa en un programa cuatrianual, es decir, un mismo grado se evaluará cada cuatro años. Estos exámenes miden el logro educativo de estudiantes en las asignaturas de Español, Matemáticas y áreas curriculares relacionadas con Ciencias Naturales y Ciencias Sociales. Por tanto, la finalidad fue aportar información para conocer en qué medida se estaba cumpliendo con los propósitos educativos que se formalizan en el currículo nacional. La selección de contenidos a evaluar se ha definido conforme al currículo nacional y el grado escolar. Los grados que se evalúan son, tercero de preescolar, tercero de primaria, sexto de primaria, tercero de secundaria y tercero de bachillerato. Con base en los resultados de esta evaluación, el Instituto señaló que el sistema educativo mexicano debe enfrentar desafíos importantes, como elevar los bajos niveles de rendimiento académico exhibidos por los estudiantes, así como lograr la equidad en la educación básica (Instituto Nacional para la Evaluación de la Educación [INEE], 2006). Conviene señalar que, adicionalmente, se elaboró el Modelo de Logro Educativo y, con base en éste, se construyeron los cuestionarios de contexto para estudiantes y docentes de tercero de primaria, esta vez con la contribución de un grupo de expertos en evaluación y en educación. Igualmente se continuó con la elaboración de manuales técnicos y se iniciaron los marcos de referencia de los EXCALE y cuestionarios de contexto; así como el desarrollo del sistema informático para la administración de los reactivos de la prueba.

Por su parte, en el año 2006, se aplicó por primera vez la Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE), considerada hasta 2013, como la evaluación más importante en el país, ya que se aplicó anualmente en toda la población estudiantil de primaria, secundaria y media superior, evaluando año con año, las asignaturas de Español y Matemáticas. En el caso de secundaria, se selecciona de manera rotativa, una materia adicional tal como Historia, Geografía, Cívica y Ética. ENLACE tuvo como propósito

4 En 2005 se aplicaron los EXCALE-06 y los EXCALE-09 y se definieron los niveles de logro para ambas. También se inició la construcción de los EXCALE-03 (3º de primaria) de Español, Matemáticas, Ciencias Naturales y Ciencias Sociales.

informar a padres de familia, docentes, directores y sociedad en general, acerca del logro académico de cada estudiante (Caso, I. González, Contreras & J. M. González, 2007). En 2014 la SEP determinó cancelar la aplicación de la prueba en Educación Básica y, por su parte, el INEE utilizó los resultados del estudio de validez de dicha prueba para elaborar los lineamientos que correspondan al diseño de una segunda generación de instrumentos para medir y evaluar el logro de aprendizajes de los alumnos de Educación Básica, que deberán aplicarse a partir de 2015 (Hernández, 2013).

Como se establece en el decreto publicado en el *Diario Oficial de la Federación*, a partir del 26 de febrero de 2013, el Instituto Nacional para la Evaluación de la Educación, se convirtió en un organismo público autónomo, con personalidad jurídica y patrimonio propio.⁵ En esta nueva etapa, el Instituto comprometió, como principal tarea, evaluar la calidad, el desempeño y los resultados del Sistema Educativo Nacional en la educación preescolar, primaria, secundaria y media superior. Para cumplir con ella debería: a) Diseñar y realizar las mediciones que correspondan a componentes, procesos o resultados del sistema; b) Expedir los lineamientos a los que se sujetarán las autoridades educativas federal y locales para llevar a cabo las funciones de evaluación que les correspondan y, c) Generar y difundir información para, con base en ésta, emitir directrices que sean relevantes para contribuir a las decisiones tendientes a mejorar la calidad de la educación y su equidad como factor esencial en la búsqueda de la igualdad social (INEE, 2015).

En materia de evaluación del aprendizaje, en 2015, el instituto desarrolló una nueva generación de pruebas: el Plan Nacional para la Evaluación de los Aprendizajes (PLANEA), Evaluaciones de Logro referidas al Sistema Educativo Nacional, en las que participan sexto grado de primaria y tercer grado de secundaria. En estas pruebas se valoran aspectos relacionados con los aprendizajes clave del currículo: los campos formativos de Lenguaje y Comunicación y Matemáticas,⁶ y también habilidades socioafectivas.⁷ Las pruebas de logro tienen la intención de conocer, en qué medida, los estudiantes alcanzan un conjunto de aprendizajes que se espera hayan adquirido al finalizar un ciclo o nivel escolar con el propósito de contar con un diagnóstico sobre el sistema educativo, en su conjunto, y de sus centros educativos en particular.

La muestra de participantes fue diseñada para obtener resultados a nivel nacional, por tipo de escuela y por entidad federativa. Además, se aplican cuestionarios de contexto para directivos, docentes y estudiantes, con el propósito de conocer las condiciones personales, familiares y escolares en que se desarrollan los aprendizajes. Junto con las pruebas PLANEA se emplean los cuestionarios de contexto para conocer el entorno personal, familiar y escolar en que se desarrollan los aprendizajes⁸ (INEE, 2016).

5 El INEE, antes de constituirse como organismo autónomo, operó en una primera etapa, del 8 de agosto de 2002 al 15 de mayo de 2012, como un descentralizado de la SEP y, en una segunda, del 16 de mayo de 2012 al 25 de febrero de 2013, como un descentralizado no sectorizado.

6 Los campos de formación relacionados con Lenguaje y Comunicación y Matemáticas integran herramientas esenciales para el desarrollo del aprendizaje de otras áreas del conocimiento; por ello, serían buenos indicadores de los resultados educativos.

7 La evaluación de las habilidades socioafectivas es clave para conocer en qué medida se favorecen los aprendizajes que intervienen en la convivencia escolar. Sus resultados permiten obtener un diagnóstico sobre la adquisición de dichos saberes, pero no sobre conductas específicas.

8 En 2015, las modalidades de ELSEN y ELCE, para 6to de primaria y 3ro de secundaria, utilizaron cuestionarios para el alumno, el grupo escolar y directores.

Los resultados nacionales⁹ más recientes del logro en las materias instrumentales (Lenguaje y Comunicación y Matemáticas) de estudiantes de 6to de primaria y 3ro de secundaria son justamente lo reportados en esta evaluación durante el 2015. Éstos se presentan en la Tabla 1 y Tabla 2.

Tabla 1: Resultados nacionales en Lenguaje y Comunicación y matemáticas en 6to de primaria

Nivel e interpretación	Porcentaje de estudiantes	
	L y C	MAT
Nivel I		
L y C: Son capaces de seleccionar información sencilla que se encuentra explícitamente en textos descriptivos. MAT: Resuelven problemas que implican comparar o realizar cálculos con números naturales.	49,5 %	60,6%
Nivel II		
L y C: Comprenden la información contenida en: textos expositivos y literarios. MAT: Resuelven problemas con números decimales y ecuaciones lineales sencillas.	33,2%	24,0%
Nivel III		
L y C: Realizan inferencias, como interpretar el sentido de una metáfora en una fábula. MAT: Resuelven problemas con números fraccionarios, con signo o potencias de números naturales. Suman o restan expresiones algebraicas.	14,6%	7,5%
Nivel IV		
L y C: Demuestran comprender textos argumentativos como artículos de opinión. MAT: Multiplican expresiones algebraicas; resuelven problemas que implican números fraccionarios y decimales (combinados); resuelven problemas que implican sistemas de ecuaciones y calculan el área de sectores circulares, y el volumen de cuerpos redondos.	2,6%	2,6%

Nota: L y C (Lenguaje y Comunicación), MAT (Matemáticas).

9 Los resultados corresponden a la aplicación en sexto de primaria y tercero de secundaria, en junio de 2015 (fin de ciclo 2014–2015). La muestra del INEE constó de 104204 estudiantes de 3446 escuelas de sexto de primaria y 144517 estudiantes de 3529 escuelas de tercero de secundaria. Los resultados nacionales no incluyen a Oaxaca, e incluyen datos parciales de Chiapas, Michoacán y Guerrero. Los resultados de PLANEA no son comparables de manera directa con EXCALE y ENLACE.

Tabla 2: Resultados nacionales en Lenguaje y Comunicación y matemáticas en 3ro de secundaria

Nivel e interpretación	Porcentaje de estudiantes	
	L y C	MAT
Nivel I		
L y C: Identifican definiciones y explicaciones en artículos de divulgación científica. MAT: Resuelven problemas que implican comparar o realizar cálculos con números naturales.	29,4%%	65,4%
Nivel II		
L y C: Reconocen la trama y el conflicto en un cuento. Identifican el propósito, el tema, la opinión y las evidencias en textos argumentativos. MAT: Resuelven problemas con números decimales y ecuaciones lineales sencillas.	46%.	24%
Nivel III		
L y C: Interpretan hechos, identifican valores y comparan el tratamiento de un mismo tema en dos relatos comparan géneros periódicos. MAT: Resuelven problemas con números fraccionarios, con signo o potencias de números naturales. Suman o restan expresiones algebraicas.	18,4%	7,5%
Nivel IV		
L y C: Identifican secuencias argumentativas y valoran sus fundamentos en un ensayo, un artículo de opinión y un debate. MAT: Multiplican expresiones algebraicas. Resuelven problemas que implican números fraccionarios y decimales (combinados). Resuelven problemas que implican sistemas de ecuaciones. Calculan el área de sectores circulares, y el volumen de cuerpos redondos.	6,1%	3,1%

Nota: L y C (Lenguaje y Comunicación), MAT (Matemáticas).

Como se observa en Lenguaje y Comunicación, casi la mitad de los estudiantes de 6to de primaria presenta deficiencias en el desarrollo de conocimientos y habilidades y más de la mitad en Matemáticas, ya que sus resultados se ubican en el Nivel I. Solamente un 2.6% domina los conocimientos y habilidades tanto en Lenguaje y Comunicación como en Matemáticas.

Por otra parte, cerca de la mitad de los estudiantes de 3ro de secundaria cuenta con un dominio deficiente de conocimiento en Lenguaje y Comunicación (Nivel II), mientras que en Matemáticas, más de la mitad obtuvo deficiencias en el desarrollo de conocimientos y habilidades (Nivel I). En ambas asignaturas los porcentajes menores se encuentran en el Nivel IV que implica el dominio de conocimientos y habilidades 6,1% y 3,1%, respectivamente.

Los resultados que obtienen los estudiantes mexicanos, tanto en primaria como secundaria, dan cuenta del problema que el sector educativo enfrenta para garantizar el acceso al aprendizaje y, por ende, a una educación de calidad (Instituto Mexicano para la Competitividad [IMCO], 2015). En México, se cuenta con más de diez años de experiencia en la aplicación de pruebas estandarizadas que dan cuenta, una y otra vez, que una proporción significativa de los estudiantes no está aprendiendo lo suficiente (Santibáñez, 2015).

En cuanto a los resultados de estudiantes mexicanos en PISA 2012, en Competencia Matemática, el 55% de los estudiantes se encuentra en los niveles inferiores (1 y Debajo del nivel 1), el 41% en los niveles intermedios (2 y 3) y sólo el 4% en los niveles altos. En lo que respecta al desempeño en la Competencia Lectora la situación es similar, ya que el 44% se ubica por debajo del nivel 2, un 54% se encuentra en los niveles intermedios y sólo un 5% en los niveles superiores.¹⁰ Con ello el país se ubicó en el lugar 52 en Competencia Lectora y en el 53 en Competencia Matemática, y al igual que el resto de los países latinoamericanos, se ubica por debajo de la media de desempeño del resto de los países miembros de la OCDE (INEE, 2013a). Así, los resultados llevan a esta conclusión: la educación recibida no es suficiente para el dominio de conocimientos y habilidades de las asignaturas instrumentales evaluadas.

3. Contribución de las evaluaciones del aprendizaje en la mejora de la calidad educativa

Como hemos revisado, desde la década de los noventa del siglo pasado, en el contexto latinoamericano distintos organismos, tanto de alcance nacional como internacional, realizan evaluaciones con el propósito de conocer el funcionamiento de sus respectivos sistemas, así, hoy por hoy, contamos con una cuantía importante de estudios en materia de evaluación educativa, particularmente, de informes de resultados de las evaluaciones del aprendizaje.

Aunado a ello parece existir un consenso respecto a que el rendimiento¹¹ alcanzando por los estudiantes (en los dominios evaluados) representa un indicador de la calidad educativa. Por ejemplo, si realizamos una búsqueda en Google Académico fijando un intervalo específico entre 1980 y 2016 «resultados en pruebas estandarizadas» en 0,06 segundos tenemos 20400 resultados. Si optamos por «resultados del rendimiento académico en pruebas estandarizadas» en 0,10 segundos obtenemos 21100 resultados. Si somos más precisos y buscamos «pruebas estandarizadas en el mejoramiento de la educación» en

10 Las tareas más fáciles corresponden al nivel 1, mientras que las más difíciles se asocian a los niveles 4, 5 y 6. El nivel 2 se identifica como el nivel básico, en el cual los estudiantes empiezan a demostrar las competencias matemáticas que les permitirán participar de manera eficaz y productiva en situaciones de la vida real.

11 Conviene señalar que al concepto de medición de rendimiento acuñado por Glaser en 1963 subyace la noción de la adquisición continua de conocimientos, por lo que se parte de la carencia total de competencias, hasta el rendimiento perfecto, con lo cual se asume que el nivel de rendimiento de un individuo se ubica en algún punto de ese *continuum* según lo indicado por los comportamientos que exhiba durante un examen.

menos de 0,04 segundos obtenemos 14700 resultados.

La creciente incorporación de los países latinoamericanos a procesos de evaluación, a nivel regional e internacional, ha tenido consecuencias positivas, ya que ha coadyuvado a la construcción y acumulación de capacidad técnica (elaboración de pruebas, incorporación de preguntas de respuesta abierta, muestreo, control de las aplicaciones, análisis de datos y modos de difusión). A su vez, ha facilitado el intercambio y desarrollo de un lenguaje común entre países (Ravela *et al.*, 2008). Actualmente, casi en toda la región, se han desarrollado buenas experiencias nacionales de evaluación, centradas en medir el aprendizaje de los estudiantes (Ferrer, 2006; Murillo, 2011; Murillo & Román, 2008). En dichos estudios se da cuenta del rendimiento académico de estudiantes de educación primaria y secundaria en las asignaturas esenciales para el acceso al conocimiento, así como para el manejo y la apropiación de códigos culturales básicos que posibilitan la inclusión a la sociedad. Así pues, con base en lo descrito en los apartados anteriores, conviene plantearse y atender una interrogante: ¿la implementación y resultados de evaluaciones del aprendizaje contribuyen en la mejora de la calidad educativa?

Se ha pronunciado que para respaldar el financiamiento de programas, acciones y prácticas educativas necesarios para optimizar, principalmente, el rendimiento académico y la condición social de los estudiantes, los educadores y políticos necesitan los resultados de evaluaciones que tengan una base científica (Schneider, Carnoy, Kilpatrick, Schmidt & Shavelson, 2007). Países como Reino Unido, Estados Unidos, Alemania, China, Corea del Sur, India, Brasil y Chile han logrado, mediante los resultados derivados de evaluaciones estandarizadas, mejorar el rendimiento académico de los estudiantes y el funcionamiento de sus sistemas educativos (McKinsey & Company, 2007). Asimismo, el análisis de la calidad educativa debe incluir resultados basados en la investigación sobre el aprendizaje de los estudiantes ya que resulta de gran utilidad en procesos de toma de decisiones en aras de un sistema educativo más justos (Garbazo, 2007). A su vez, las pruebas estandarizadas de logros comprenden una evidencia necesaria para hacer interpretaciones referidas a normas, respecto a los conocimientos y habilidades de cada estudiante en relación a una muestra nacional de estudiantes de la misma edad o grado escolar. En este sentido, los resultados tienen una utilidad educativa considerable (Popham, 1999).

Si nos referimos, particularmente a PISA, esta evaluación expone a aquellos países que han alcanzado un buen rendimiento y, al mismo tiempo, un reparto equitativo de oportunidades de aprendizaje, ayudando así a establecer metas ambiciosas para otros países (INEE, 2013b). A su vez, otra importante contribución de PISA estriba en el conjunto de cuestionarios dirigidos a alumnos y directivos escolares, ya que la información derivada de éstos resulta útil para contemplar el proceso educativo, y no sólo como estrategia para revelar qué factores pueden ser explicativos de los resultados, sino también como fuente de información en sí misma (Ferrer, 2012). También, este mismo autor reconoce que otro aspecto, en el que contribuye esta evaluación, concierne al rigor metodológico acompañado de una perspectiva y actuación fuertemente independiente en su diseño y explotación de los datos (Ferrer, 2012).

Por su parte, el LLECE reconoce algunas contribuciones para la mejora de la calidad educativa derivadas de sus estudios regionales comparativos y explicativos. Particularmente, en su última aplicación TERCE reportó información sobre la eficacia, una dimensión clave de la calidad de la educación, y a su vez da cuenta de la brecha en el nivel de logro

de aprendizajes entre estudiantes provenientes de distintos contextos. También conviene referir que la UNESCO confía en que los hallazgos y las conclusiones de los estudios en cuestión serán útiles para que los países conozcan el nivel de sus logros de aprendizaje y en consecuencia formulen políticas que mejoren sus sistemas de educación (LLECE, 2015).

En el caso de México, al remitirnos a la contribución de las evaluaciones del aprendizaje en la mejora de la calidad educativa, en primera instancia conviene destacar las actividades actuales emprendidas por el INEE y la SEP, quienes reconocen que los objetivos principales del esfuerzo evaluativo emprendido a través de PLANEA, son: a) Conocer la medida en que los estudiantes logran el dominio de un conjunto de aprendizajes esenciales al término de los distintos niveles de la educación obligatoria, b) Ofrecer información contextualizada para la mejora de los procesos de enseñanza en los centros escolares, c) Informar a la sociedad sobre el estado que guarda la educación, en términos del logro de aprendizajes de los estudiantes y, d) Aportar a las autoridades educativas la información relevante y utilizable para el monitoreo, la planeación, programación y operación del sistema educativo y sus centros escolares.¹²

En segunda instancia, la publicación de los resultados de las evaluaciones, tanto nacionales, regionales como internacionales, ha permitido que los docentes de secundaria y de Educación Media Superior identifiquen cómo está el país en comparación con otros, y cómo está su entidad federativa. El instituto también se ha esforzado por elaborar material derivado de las pruebas a gran escala que le permita a los docentes hacer uso educativo de ellas.¹³ En éste se incluyen elementos concretos para facilitar el trabajo analítico, práctico y reflexivo de los estudiantes, así como propuestas didácticas para la educación secundaria. Cabe referir que el INEE pone a disposición de investigadores y el público en general, las bases de datos derivadas de la aplicación de las Evaluaciones de Logro referidas al Sistema Educativo Nacional (ELSEN), para que las consulten o incluso para realizar sus propios análisis.

Si bien las aportaciones descritas en párrafos anteriores no deben soslayarse, algunos investigadores e instancias han identificado ciertas problemáticas o limitaciones implicadas en la presunta contribución de la evaluación estandarizada del aprendizaje en la mejora de la calidad educativa. En las siguientes líneas se discuten algunos cuestionamientos.

Aún no se demuestra si la evaluaciones estandarizadas del aprendizaje realmente contribuye a mejorar y optimizar la calidad de la educación, lo que a su vez compromete reducir las inequidades existentes en los sistemas educativos (Murillo & Román, 2010). En algunos estudios se ha expuesto el escaso uso de la información brindada por los sistemas de evaluación (Ferrer, 2005; Montoya, Perusia & Vera, 2001; Tiana, 2002). Particularmente, los resultados de los estudios comparativos y explicativos que realiza el LLECE aportan información relevante que comprende insumos para el desarrollo de políticas en mate-

12 Los resultados de las evaluaciones PLANEA no deberán utilizarse para juzgar el desempeño de los docentes, realizar *rankings* de escuelas, justificar procesos punitivos u otros de control administrativo sobre estudiantes, docentes o escuelas.

13 El libro titulado *PISA para docentes*. También pueden consultar y usar, como un recurso complementario en la enseñanza, la serie *PISA en el aula* integrada por tres libros, uno por área de evaluación (Lectura, Ciencias y Matemáticas). Otro material dirigido a los docentes es la colección *La competencia lectora desde PISA* que integra tres cuadernos organizados, a partir de los tres formatos textuales que maneja PISA (continuo, discontinuo y mixto).

ria de evaluación para los países participantes, no obstante, en Argentina, Colombia y México la difusión de los informes de resultados, hasta ahora han tenido escaso impacto (Ferrer, 2006).

Si bien existe una creciente preocupación respecto a la divulgación y el uso de los resultados, parece que las acciones para lograr que los docentes comprendan y utilicen los resultados de las evaluaciones y las incorporen a la práctica y cultura escolar, son insuficientes. A su vez, la mayoría de los docentes en Latinoamérica, no comprende los aspectos técnicos involucrados en el diseño de las evaluaciones nacionales e internacionales, ni los tipos de resultados que las mismas arrojan; por tanto, difícilmente pueden utilizarlos productivamente (Ravela *et al.*, 2008). Co todo ello, pareciera que las conclusiones derivadas de la investigación no están teniendo repercusiones en la mejora, ni en las políticas educativas, ya que también en las acciones educativas obvian los modelos y hallazgos de estudios (Martínez Rizo, 2015).

Por otra parte, se ha expresado que las pruebas estandarizadas de logros académicos tienen la misión de medir algo distinto que cuán buena o cuán mala es una escuela, por tanto, no deben ser usadas para medir la calidad de la educación (Popham, 1999). Al parecer, la calidad educativa debería ser medida en términos de la contribución de la escuela en el desarrollo económico y social sostenible de las comunidades (Yepes, 2004). Aunado a ello, debe de disminuir el peso que tiene la rendición de cuentas en las evaluaciones estandarizadas y aumentar la importancia de su función pedagógica, es decir, aquella que es útil en el salón de clases para el docente y sus estudiantes (Gordon Commission, 2013).

Retomando la interrogante inicialmente expuesta, son indiscutibles los avances en cuanto a la implementación (hasta cierto punto sistemática) de evaluaciones centradas, principalmente, en medir el aprendizaje de estudiantes, por lo que se ha construido y acumulado capacidad técnica y metodológica en la elaboración de pruebas, muestreo, control de las aplicaciones y análisis de datos. No obstante, respecto el uso y contribución de los resultados de este tipo de evaluaciones para la mejora educativa comprende un terreno frágil, por ello conviene pronunciar algunos alcances y desafíos.

4. Consideraciones finales: alcances y desafíos

Es innegable que, en los últimos veinte años, los países latinoamericanos han conseguido avances significativos en materia de alfabetización y cobertura educativa, sin embargo, continúan pendientes importantes desafíos en materia de calidad y equidad (LLECE, 2015).

La evaluación tiene una indudable dimensión política, no es una operación puramente técnica. Al definir qué evaluamos y cómo lo hacemos, tomamos decisiones importantes sobre la relevancia de los aspectos y dimensiones que sometemos a estudio. Expresado en otras palabras, otorgamos relevancia a determinadas parcelas de la realidad misma que sometemos a estudio. Así, de poco sirven los resultados de PISA para tomar acciones de mejora si no consideramos el impacto que ejercen cuestiones como la formación y el trabajo de los docentes (Tiana, 2011).

En el caso de México, para que las evaluaciones nacionales cumplan con sus propósitos, es importante que tengan un perfil de bajo impacto que ayude a garantizar que la información que proporcionan, tenga la mayor validez y confiabilidad posible. De tal suerte

que orienten las estrategias necesarias para contribuir a la calidad de la educación y al máximo de los aprendizajes de los estudiantes, tal y como lo señala la Reforma Educativa en México (Congreso de la Unión, 2013).

Un desafío particular es que la difusión de los resultados requiere adecuarse a las necesidades y capacidades de los diferentes usuarios de la información. En este sentido también suele esperarse que otros actores, como las familias, emprendan acciones que contribuyan a la mejora de la calidad, o que docentes que ya hacían su mejor esfuerzo, reorienten sus prácticas en sentido más productivo (Contreras & Backhoff, 2014).

Sabemos que los bajos niveles de logro obtenidos por los estudiantes se deben a combinaciones de varios déficits que involucran principalmente al contexto escolar (considerando la preparación de los docentes) y al contexto socio-familiar. Si esto no se contempla, difícilmente, mejorará el logro de los estudiantes, aunque se les evalúe frecuentemente, por tanto, evaluar resulta necesario para mejorar, pero no precisa ni principalmente, mediante evaluaciones a gran escala (Martínez Rizo, 2015).

Evaluar la calidad de la educación exige un enfoque integral, acorde a la complejidad y finalidad del objeto de dicha evaluación, por lo que deseablemente la valoración de sus componentes, debe estar interrelacionada (la evaluación del aprendizaje, del desempeño docente, de la escuela, de los directivos, de la administración educativa y de los programas implementados). Lo único que parece indiscutible es que las tareas evaluativas tienen sentido, justamente, en la medida en que contribuyen en la mejora educativa.

Referencias bibliográficas

- Álvarez, G., Guerrero, C. & De la Cera, S. (1994). *Sistema Educativo Nacional de México*. Madrid: Secretaría de Educación Pública—Organización de Estados Iberoamericanos. Recuperado de <http://www.oei.es/quipu/mexico/mex15.pdf>.
- Casassus, J. (2007). El precio de la evaluación estandarizada: la pérdida de calidad y la segmentación social. *Revista Brasileira de Política e Administração da Educação*, 23(1), 71–79. Recuperado de <http://seer.ufrgs.br/index.php/rbpael/article/view/19014/11045>.
- Caso, J., González, I., Contreras, L. A. & González, J. M. (2007). *Antecedentes de la evaluación a gran escala en Baja California*. Ensenada: Unidad de Evaluación Educativa/Universidad Autónoma de Baja California. Recuperado de <http://uee.uabc.mx/uee/documentos/estudiosYproyectos/UEERT07-001.pdf>.
- Congreso de la Unión. (2013). Decreto por el que se reforman, adicionan y derogan diversas disposiciones de la Ley General de Educación. *Diario Oficial de la Federación*, México D.F.: Secretaría de Educación Pública. Recuperado de <http://www.presidencia.gob.mx/reformaeducativa/assets/downloads/reforma-la-leygeneral-de-educacion.pdf>.

- Contreras, S. & Backhoff, E. (2014). Tendencias en el aprendizaje de la educación en México: una comparación entre ENLACE, Excale y Pisa. *Nexos*. Recuperado de <http://www.nexos.com.mx/?p=22749>.
- Degenhart, R. E. (ed.) (1990). *Thirty years of international research. An annotated bibliography of IEA publications (1960–1990)*. La Haya: IEA Headquarters.
- Ferrer, F. (2012). PISA: Aportaciones e incidencia sobre las políticas educativas nacionales. *Revista Española de Educación Comparada*, (19), 11–16.
- Ferrer, G. (2005). *Estado de situación de los sistemas nacionales de evaluación de logros de aprendizajes en América Latina*. Recuperado de <http://www.preal.cl/GTEE/docr/Estado%20de%20Situacion.pdf>.
- Ferrer, G. (2006). *Sistemas de evaluación de aprendizajes en América Latina. Balance y desafíos*. Santiago, Chile: PREAL. Recuperado de http://www.enlaceacademico.org/uploads/media/Libro_completo.pdf.
- Garbazo, G. (2007). Factores asociados al rendimiento académico en estudiantes universitarios. Una reflexión desde la calidad de la educación superior pública en Costa Rica. *Revista de la Universidad de Costa Rica*, 31(1), 43–63.
- González, C., Caso, J. Díaz, K. & López–Ortega, M. (2012). Rendimiento académico y factores asociados: aportaciones de algunas evaluaciones a gran escala. *Bordón. Revista de Pedagogía*, 64(2), 51–68.
- Gordon Commission. (2013). *A public policy statement. The Gordon Commission on the future of assessment in education*. Recuperado de http://www.gordoncommission.org/rsc/pdfs/gordon_commission_public_policy_report.pdf.
- Hernández, L. (19 de diciembre de 2013). Prueba ENLACE ya no existirá para educación básica: SEP. *Excelsior*. Recuperado de <http://www.excelsior.com.mx/nacional/2013/12/19/934650>.
- Instituto Mexicano para la Competitividad (2015). *Resultados de la prueba PLANEA vía SEP*. Recuperado de http://imco.org.mx/banner_es/resultados-de-la-prueba-planea-via-sep/.
- Instituto Nacional para la Evaluación de la Educación. (2006). *Plan Maestro de Desarrollo 2007–2014*. Recuperado el 13 de junio de 2012 de http://www.inee.edu.mx/images/stories/Publicaciones/Documentos_oficiales/2006/Plan_maestro/Completo/planmaestro.pdf.
- Instituto Nacional para la Evaluación de la Educación (2013a). *Ley del Instituto Nacional para la Evaluación de la Educación*. México: Plan Nacional para la Evaluación de los Aprendizajes PLANEA. Documento interno de la Unidad de Evaluación del Sistema Educativo Nacional.

- Instituto Nacional para la Evaluación de la Educación (2013b). *México en PISA 2012*. México: INEE. Recuperado de http://www.sems.gob.mx/work/models/sems/Resource/11149/1/images/Mexico_PISA_2012_Informe.pdf.
- Instituto Nacional para la Evaluación de la Educación (2015). *¿Qué es el INEE?* Recuperado de (2006).<http://www.inee.edu.mx/index.php/acerca-del-inee>.
- Instituto Nacional para la Evaluación de la Educación (2016). *Plan Nacional para la Evaluación de los Aprendizajes (Planea. Resultados nacionales 2015. Sexto de primaria y tercero de secundaria. Lenguaje y comunicación y Matemáticas*. Recuperado de http://planea.sep.gob.mx/content/general/docs/2015/difusion_resultados/1_Resultados_nacionales_Planea_2015.pdf.
- Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (2015). *Informe de resultados TERCE. Antecedentes iniciales*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. Recuperado de <http://unesdoc.unesco.org/images/0024/002439/243980s.pdf>.
- McKinsey & Company. (2007). *¿Cómo se convierte un sistema educativo de bajo desempeño en uno bueno?* Recuperado de http://recursostic.educacion.es/blogs/europa/media/blogs/europa/informes/Resumen_ejecutivo_McKinsey2010.pdf.
- Martínez Rizo, F. (coord.) (2015). *Las pruebas ENLACE y Excale. Un estudio de validación*. México D.F.: Instituto Nacional para la Evaluación de la Educación.
- Montoya, S., Perusia, J. C. & Vera, M. A. (2001). *Evaluación de la calidad educativa: De los sistemas centrales al aula. Estudio del impacto de las políticas de devolución de información personalizada a las escuelas, en tres provincias de Argentina*. Buenos Aires: Centro de Investigaciones para la Equidad y Calidad Educativa.
- Morillejo, E., Reboloso, E., Pozo, C. & Fernández, B. (1999). Evaluación de la calidad en la educación superior. *Papeles del Psicólogo*, 74, 12–17.
- Murillo, F. J. (2011). Hacer de la educación un ámbito basado en evidencias científicas. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 9(3), 4–12. Recuperado de <http://www.rinace.net/reice/numeros/arts/vol9num3/editorial.pdf>.
- Murillo, F. J. & Román, M. (2008). Resultados de aprendizaje en América Latina a partir de las evaluaciones nacionales. *Revista Iberoamericana de Evaluación Educativa*, 1(1), 45–61.
- Murillo, F. J. & Román, M. (2010). Retos en la evaluación de la calidad de la educación en América Latina. *Revista Iberoamericana de Educación*, (53), 97–120.
- Popham, W. J. (1999). ¿Por qué las pruebas estandarizadas no miden la calidad educativa? Programa de Promoción de la Reforma Educativa en América Latina y el Caribe. Recuperado de [file:///C:/Users/Jenny/Downloads/pruebas_estandarizadas_no_miden_calidad_educativa_popham%20\(1\).pdf](file:///C:/Users/Jenny/Downloads/pruebas_estandarizadas_no_miden_calidad_educativa_popham%20(1).pdf).

- Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Martínez-Rizo, F.,... Wolff, L. (2008). Las evaluaciones que América Latina necesita. *Revista Iberoamericana de Evaluación Educativa*, 1(1), 51–63. Recuperado de <http://dialnet.unirioja.es/servlet/articulo?codigo=2602512>.
- Santibáñez, L. (2015). ¿Qué nos dice PLANEA? *Nexos*. Recuperado de <http://educacion.nexos.com.mx/?p=81>.
- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. & Shavelson, R. (2007). *Estimating Causal Effects Using Experimental and Observational Designs. A Think Tank White Paper*. The governing board of the AERA grants programs. Recuperado de http://69.8.231.237/uploadedFiles/Publications/Books/Estimating_Causal_Effects/ECE_Front-TOC.pdf.
- Tiana, A. (1996). La evaluación de los sistemas educativos. *Revista Iberoamericana de Educación*, 10(2) 37–54.
- Tiana, A. (2002). *Una mirada técnico-pedagógica acerca de las evaluaciones de la calidad de la educación. ¿Qué pretendemos evaluar, qué evaluamos y qué conclusiones podemos extraer de la evaluación?* Recuperado de http://www.oei.org.ar/noticias/Que_pretendemos_evaluar.pdf.
- Tiana, A. (2011). ¿Y después de PISA qué? *REICE: Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 9(1), 3–5.
- Van Dijk, S. (2009). Breve revisión histórica de la evaluación en México. *Revista de Investigación Educativa*, 7, 51–72. Recuperado de <http://www.educacion.ugto.mx/educatio/PDFs/educatio7/Van.pdf>.
- Yepes, R. (2004). Calidad educativa: más que resultados en pruebas estandarizadas. *Educación y Pedagogía*, 16(38), 75–89.
- Zorrilla, M. (2003). *La evaluación de la educación básica en México 1990–2000. Una mirada a contraluz*. Aguascalientes, México: Universidad Autónoma de Aguascalientes.