

## Capítulo 3

### La evaluación de los alumnos

*En México, el desempeño de los alumnos es evaluado mediante una amplia gama de instrumentos que van desde los exámenes nacionales estandarizados hasta la evaluación formativa y continua en el aula. Todos los alumnos son evaluados de manera continua durante todo el año escolar en cada asignatura o área del currículo. Las calificaciones empleadas para reportar el logro de los alumnos utilizan una escala de 5 a 10. Los criterios y los métodos de evaluación son definidos por cada maestro. También se cuenta con exámenes nacionales finales, aplicados externamente, al término de la primaria (Instrumento para el Diagnóstico de Alumnos de Nuevo Ingreso a Secundaria, IDANIS) y de la educación secundaria (Examen Nacional de Ingreso a la Educación Media Superior, EXANI I). Estas evaluaciones sirven para diagnóstico y selección (por la escuela en el siguiente nivel). A nivel nacional, existe también una cohorte completa de evaluaciones externas (Evaluación Nacional del Logro Académico en Centros Escolares, ENLACE) que se utiliza con fines de diagnóstico y mejora, pero que tiene un “alto impacto” para los maestros y las escuelas. En educación básica, ENLACE se administra anualmente a todos los alumnos de tercero de primaria a tercero de secundaria, en español y matemáticas y una tercera asignatura que varía cada año.*

*Una ventaja importante es que en México la evaluación es vista como parte de la función profesional de los maestros. Otros aspectos positivos incluyen la introducción de un nuevo marco global para la evaluación dentro del aula; los avances en la alineación de las calificaciones con los resultados de aprendizaje esperados; la debida atención para reducir la reprobación; la promoción de la participación de los padres de familia en el aprendizaje de sus hijos; y la capacidad para implementar evaluaciones a gran escala. Sin embargo, existen considerables retos en la creación de métodos eficaces para evaluar a los alumnos. Entre ellos los actuales enfoques tradicionales en enseñanza y evaluación, la prevalencia de enseñar para el examen en todo el sistema escolar; la excesiva dependencia en exámenes de opción múltiple; el gran número de objetivos de ENLACE, las prácticas de calificación con poca pertinencia pedagógica; la falta de coherencia de las evaluaciones del alumnado en todas las escuelas y clases; la limitada capacidad a nivel estatal y local para apoyar la evaluación dentro del aula; y la necesidad de mejorar los instrumentos para el reporte de calificaciones.*

Este capítulo se centra en los enfoques para la evaluación de los alumnos dentro del marco de evaluación de México. La evaluación de los alumnos se refiere a los procesos mediante los cuales se recogen evidencias del aprendizaje de una manera planificada y sistemática, con el fin de hacer un juicio sobre el aprendizaje de los alumnos (EPP1, 2002). En este capítulo se analiza tanto la evaluación sumativa (evaluación *del* aprendizaje) como la evaluación formativa (evaluación *para* el aprendizaje) de los alumnos.

## Contexto y características

### *Panorama*

La evaluación de los alumnos en México está formada por tres componentes principales: (i) la evaluación dentro del aula con fines formativos y sumativos; (ii) las evaluaciones externas con fines de diagnóstico y selección para el ingreso al primer ciclo de enseñanza secundaria (IDANIS) y a la educación media superior (EXANI I); y (iii) evaluaciones externas con fines de diagnóstico y mejora (ENLACE).

La evaluación dentro del aula, con fines formativos y sumativos, está a cargo de los maestros y es regulada por el Acuerdo oficial 200 (DOF, 1994). Esta norma ha estado en vigor desde 1994 y establece la obligación de que las escuelas públicas y privadas lleven a cabo la evaluación de los alumnos con fines formativos y de certificación, basadas en los conocimientos, habilidades y actitudes incluidas en el plan de estudios nacional. Este acuerdo regula la periodicidad de las evaluaciones sumativas, la escala que debe utilizarse para asignar calificaciones y las condiciones para aprobar o reprobar a un alumno. En los últimos años, se lanzó un nuevo enfoque de evaluación basado en el aula bajo el contexto de la ACE y la RIEB. Las nuevas iniciativas se encuentran en sus etapas iniciales de implementación (incluidas las propuestas para ajustar el Acuerdo 200).

El IDANIS (Instrumento para el Diagnóstico de Alumnos de Nuevo Ingreso a Secundaria) y el EXANI I (Examen Nacional de Ingreso a la Educación Media Superior) son exámenes nacionales externos con fines de diagnóstico y selección. IDANIS fue creado en 1989, es operado por la SEP y evalúa a los alumnos que entran en la educación secundaria. EXANI I fue creado en 1994, es operado por el CENEVAL y evalúa a los alumnos que ingresan a la educación media superior. Ambos instrumentos son administrados con el fin de proporcionar información a las autoridades y a las escuelas para la selección de alumnos (Vidal, 2009).

ENLACE es una evaluación nacional estandarizada que ha sido administrada cada año desde 2006 por la SEP, con una base censal. En la actualidad, todos los alumnos de tercero a sexto grado de educación primaria y de los tres grados de educación secundaria son evaluados en español y matemáticas. Desde 2008, se evalúa una tercera asignatura que varía cada año: ciencias en 2008, educación cívica en 2009, historia en 2010 y geografía en 2011. Los resultados están disponibles en forma individual para cada alumno y las calificaciones promedio por escuela se difunden ampliamente. En un inicio, el propósito explícito de ENLACE era proporcionar información sobre el desempeño académico de los alumnos en los objetivos curriculares, para fomentar la mejora de las prácticas profesionales docentes y aumentar la calidad del aprendizaje. Posteriormente, se añadieron nuevos objetivos y consecuencias a estas evaluaciones (consulte más adelante; (SEP e INEE, de próxima aparición).

### ***Normas para la evaluación dentro del aula***

Las evaluaciones dentro del aula se rigen por el Acuerdo oficial número 200 (DOF, 1994), donde se establece que la evaluación de los alumnos debe ser formativa, sistemática, continua e integral. También especifica que, para fines sumativos, los maestros deben utilizar una escala numérica de 5 a 10. En esta escala, el 6 es la calificación mínima aprobatoria, 5 significa insuficiente y 10 es excelente.

Cada dos meses (en octubre, diciembre, febrero y abril) y en los últimos cinco días hábiles del año escolar, se debe asignar calificaciones e informar a los padres de familia, como se indica en el Acuerdo 499 (DOF, 2009), que introdujo modificaciones menores al Acuerdo 200. Esta norma es cumplida por todas las escuelas y maestros; como resultado, las actividades de enseñanza y aprendizaje están organizadas en cinco periodos o bloques, con una duración de dos meses cada uno. Los directores de escuela deben notificar las calificaciones a los alumnos y padres de familia y fomentar la comunicación entre estos y los maestros. Para ello, las escuelas deben completar una boleta de calificaciones estandarizada para cada alumno con las calificaciones obtenidas en cada asignatura, que debe ser firmada por los padres, en cada periodo. El diseño de la boleta de calificaciones actual corresponde al plan de estudios nacional de 1993. En el contexto de la RIEB, una nueva boleta de calificaciones llamada Cartilla de Educación Básica está siendo diseñada y probada en las escuelas (consulte más adelante).

Al final del año lectivo, se debe calcular una calificación final para cada asignatura, que es el promedio de las calificaciones obtenidas durante el año escolar. Cada alumno recibe un promedio general anual, que es la media de las calificaciones finales de cada asignatura (SEP, 2011a).

En la educación primaria no debe haber reprobación entre el primero y el segundo grado, ya que ambos son considerados parte de un periodo único de aprendizaje. Entre los grados segundo y sexto, los alumnos deben obtener al menos 6.0 como calificación final en español y matemáticas para aprobar, así como un promedio general anual igual o superior a 6.0. Los alumnos que no obtengan estas calificaciones deben repetir el año, a menos que asistan a un “procesos de regularización”, que consiste en cursos de verano o tutorías organizados e impartidos en los estados. Los alumnos pueden pasar al siguiente grado si están certificados por un examen extraordinario antes de empezar el siguiente año escolar (SEP, 2011a; SEP e INEE, de próxima aparición).

En la educación secundaria, los alumnos que no logren obtener una calificación mínima de 6.0 en más de cinco asignaturas deberán repetir todo el ciclo. Al final del año escolar, los alumnos que reprobren cinco o menos asignaturas deberán cursar actividades de nivelación y aprobar un examen para cada una de las materias reprobadas. Los alumnos no podrán inscribirse en el curso siguiente, si no han aprobado más de dos asignaturas después del “periodo de regularización” en septiembre (al inicio del año escolar) (SEP, 2011a; SEP e INEE, de próxima aparición).

### ***Prácticas de evaluación en el aula***

Los procedimientos y las normas descritas antes se centran en los aspectos formales de la evaluación sumativa. La evaluación formativa se considera importante, pero no cuenta con lineamientos concretos para realizarla. En cuanto a la evaluación sumativa, el Acuerdo 200 no incluye lineamientos sobre el significado de las calificaciones en términos del desempeño esperado de los alumnos. Este asunto empieza a abordarse en la RIEB, la reforma curricular actual, que incluye propuestas de revisión al Acuerdo 200 (consulte más adelante).

Dentro de este marco formal, las escuelas y los maestros tienen plena libertad para determinar los criterios y llevar a cabo la evaluación de los alumnos. Los docentes utilizan criterios muy diferentes para evaluar a sus alumnos, en particular para asignar calificaciones al logro académico. Según lo observado por el Equipo de Revisión de la OCDE durante su visita, muchos maestros utilizan un enfoque normativo para calificar, lo que significa que primero asignan la máxima calificación (10) a los mejores alumnos de su clase y después califican a los demás alumnos con base en esta referencia. También hay casos comunicados por diferentes actores escolares al Equipo de Revisión de la OCDE, en los que los maestros ajustan sus criterios de evaluación para aparentar un menor número de alumnos con bajo desempeño.

Para la mayoría de los maestros en México calificar consiste en asignar puntos a los alumnos mediante una serie de elementos: cumplimiento de tareas, asistencia a clase, participación en las actividades del salón de clases, pulcritud en las tareas, disciplina, trabajo en equipo, presentaciones y exámenes (por lo general con un formato de opción múltiple). Estos aspectos reciben “puntos” que al final se promedian para obtener una calificación. Cuando se les preguntó cuál elemento era el más importante para asignar la calificación bimensual (además de los resultados de los exámenes), 71% de los maestros de primaria respondieron “atención y participación” (García *et al.*, 2011).

El peso asignado a cada uno de estos elementos varía entre los maestros y entre los ciclos para el mismo docente, en función de los contenidos impartidos y las actividades realizadas. La ponderación de estos elementos no es explícita, pero los exámenes suelen tener más influencia en la calificación final (García *et al.*, 2009; Loureiro, 2009; Picaroni, 2009; Ravela, 2009a; SEP e INEE, de próxima aparición).

La evaluación formativa en México se realiza de una manera muy limitada: consiste básicamente en dar calificaciones por exámenes, tareas y actitudes, y comunicarles a los alumnos dónde han fallado. Hay poca evidencia de que los maestros estén conscientes de la importancia de dar retroalimentación a los alumnos durante el proceso de la enseñanza, con el fin de ayudarlos a reflexionar sobre su propio aprendizaje o resultados (Picaroni, 2009; Ravela, 2009a).

### ***Capacidades de los maestros para la evaluación de los alumnos***

De acuerdo con el plan de estudios nacional de formación inicial para docentes de educación primaria (SEP, 2002) y secundaria (SEP, 2010), los maestros deben ser capacitados durante su educación para realizar evaluaciones del alumnado en el aula. Se supone que la capacidad para evaluar eficazmente el aprendizaje de los alumnos se desarrolla mediante un curso llamado “Planeación de la enseñanza y evaluación del aprendizaje”, impartido seis horas a la semana en el sexto semestre de la licenciatura en Educación Primaria y cuatro horas a la semana en el cuarto semestre de la licenciatura en Educación Secundaria (SEP e INEE, de próxima aparición).

Este curso tiene como objetivo preparar a los futuros maestros para que organicen actividades de enseñanza, y evalúen los avances y dificultades en el aprendizaje de los alumnos. Los futuros docentes deben aprender a utilizar instrumentos adecuados a los contenidos curriculares y las características de los alumnos. El curso también debe prepararlos para evaluar los procesos de aprendizaje, definir el momento de administrar instrumentos estandarizados y escribir reactivos de examen que demanden reflexión de los alumnos para que utilicen los conocimientos y habilidades intelectuales que han adquirido. Los futuros maestros deben aprender que el objetivo principal de la evaluación es mejorar la enseñanza y el aprendizaje en lugar de asignar calificaciones solamente (SEP, 2002, 2010; SEP e INEE, de próxima aparición).

En cuanto a la formación y al desarrollo profesional de los maestros en servicio, la evaluación de los alumnos dentro del aula es una línea de trabajo importante para la Dirección General de Formación Continua para Maestros en Servicio (DGFCMS) de la SEP. La principal iniciativa nacional en ese ámbito, el Catálogo Nacional de Formación Continua y Superación Profesional de Maestros de Educación Básica en Servicio,<sup>1</sup> se publica anualmente y contiene una amplia oferta de más de 1 000 cursos. En él se incluyen desde cursos de corta duración (40 horas) hasta programas de maestría y doctorado (SEP e INEE, de próxima aparición). Estos cursos, que son impartidos por instituciones de educación superior de todo el país, son evaluados y acreditados por la DGFCMS. Los temas relacionados con la evaluación están recibiendo mayor énfasis en las ofertas disponibles para los docentes. Mientras que hace dos años sólo dos programas abordaban específicamente temas de evaluación, el catálogo actual 2011-2012 incluye más de 30 programas sobre la materia, entre los cerca de 1 100 ofrecidos. La mayoría de ellos están dirigidos a supervisores escolares y se enfocan en la evaluación basada en competencias. Al mismo tiempo, un gran número de cursos sobre temas curriculares incluyen nuevos métodos, técnicas e instrumentos de evaluación basados en el aula (SEP, 2011b; SEP e INEE, de próxima aparición).

La SEP también ha implementado otros mecanismos de apoyo para facilitar el desarrollo de las capacidades docentes para la evaluación de los alumnos. Entre ellos libros para maestros, una sección especial sobre evaluación y autoevaluación en los libros de texto y una variedad de materiales disponibles en el sitio web de la SEP, como artículos de investigación sobre evaluación, indicadores de evaluación y enlaces a otros sitios web pertinentes relacionados con la evaluación (SEP e INEE, de próxima aparición).

### ***Evaluación externa de los alumnos con fines de selección***

IDANIS fue creado en 1989 con el objetivo de aportar información para la selección de alumnos que desean ingresar a la educación secundaria. Esta prueba es administrada y procesada por la DGEP-SEP sólo en aquellos estados que lo soliciten. Tiene alrededor de 60 reactivos de opción múltiple y preguntas para completar en los espacios en blanco; evalúa las habilidades básicas de los alumnos en tres áreas: comunicación, uso de las matemáticas y razonamiento abstracto. Los conceptos de la prueba son muy tradicionales y evalúan capacidades muy sencillas. Los resultados son utilizados para asignar a los nuevos alumnos en las escuelas secundarias, seleccionadas por ellos. Por ejemplo, como se menciona en el sitio web de la SEP-DF, durante el proceso de selección en el Distrito Federal, los alumnos que desean ingresar a secundaria eligen las tres escuelas que prefieren. Los resultados de IDANIS son el principal criterio para asignar a los alumnos a cada escuela.<sup>2</sup> Por lo tanto, estos resultados son muy importantes para los alumnos, ya que determinan su acceso a las escuelas secundarias seleccionadas (SEP e INEE, de próxima publicación; SEP-DF, 2011). Como se observó durante la visita de revisión, en otros estados el proceso de selección es llevado a cabo normalmente por cada escuela secundaria, la cual tiene más candidatos que lugares disponibles.

EXANI I, creado en 1994 y administrado por el CENEVAL, se aplica de forma voluntaria y los alumnos o instituciones tienen que pagar para tomarlo. Debido a que los resultados de EXANI I son muy utilizados por las escuelas de nivel medio superior para la selección de los alumnos, también tienen una fuerte influencia sobre ellos y sus maestros de educación secundaria. En la actualidad, EXANI I está formado por dos instrumentos. El primero es un examen de selección tradicional que mide las habilidades intelectuales generales. El segundo es una prueba diagnóstica de reciente introducción que evalúa el contenido de las principales asignaturas que se debería haber aprendido en la educación secundaria y que son importantes para el siguiente nivel.

Los dos instrumentos son normas de referencia (es decir, los evaluados son comparados entre sí). Al igual que en IDANIS, los resultados de EXANI I son utilizados por las escuelas de nivel medio superior con fines de selección (Vidal, 2009; SEP e INEE, de próxima aparición).

IDANIS y EXANI I ejercen una fuerte influencia sobre alumnos y maestros, los cuales dedican tiempo a su preparación ya que determinan las posibilidades de los alumnos para asistir a la escuela de su preferencia en el siguiente nivel. Por otra parte, una búsqueda de “IDANIS” o “EXANI I” en Internet lleva a varios sitios que ofrecen cursos de preparación para los alumnos (por un pago). En la página web del CENEVAL se ofrecen guías y materiales para la preparación de las pruebas que están disponibles de forma gratuita, aunque los mismos materiales pagados se ofrecen en otros sitios.

### *Evaluación externa de los alumnos con fines formativos*

ENLACE es un importante esfuerzo de evaluación que abarca más de 14 millones de alumnos de tercer grado de educación primaria a tercer grado de secundaria de todas las escuelas públicas y privadas de México. Desde 2008 también incluye a los alumnos que están a punto de terminar la educación media superior, evaluándolos en comprensión de lectura y matemáticas. En este nivel de educación ENLACE es diferente, ya que pone a prueba las competencias en lugar de los contenidos curriculares, e implica una logística y una planeación separada.<sup>3</sup>

La administración de la prueba es controlada externamente por los padres de familia de cada escuela. Los maestros supervisan su administración para un grupo de alumnos diferente al suyo (para evitar que vigilen a sus propios alumnos). Las pruebas incluyen de 50 a 70 preguntas de opción múltiple sobre el contenido curricular estudiado. Mediante el uso de un modelo con tres parámetros de la Teoría de Respuesta al Ítem, los resultados se calculan empleando una escala horizontal que tiene como referencia una media de 500 puntos, que corresponden a los resultados promedio de los alumnos en la primera aplicación de ENLACE (en 2006). Los alumnos son colocados en cuatro niveles de desempeño: “insuficiente”, “básico”, “bueno” y “excelente”. “Insuficiente” significa que los alumnos carecen de los conocimientos y habilidades necesarios para continuar el aprendizaje de la asignatura a un ritmo adecuado. El nivel “Básico” significa que los alumnos son competentes sólo en una pequeña parte de los conocimientos y habilidades evaluadas en una materia y grado escolar, pero que es suficiente para seguir aprendiendo de manera satisfactoria. Los alumnos con un nivel “Bueno” son competentes en la mayoría de los conocimientos y habilidades evaluados en una materia y grado escolar. Los alumnos con un nivel “Excelente” son competentes en todos los conocimientos y habilidades evaluados en una asignatura y grado escolar (SEP e INEE, de próxima aparición).

Los resultados se presentan con puntuaciones y porcentajes de alumnos por nivel de desempeño y están disponibles en el sitio web de la SEP. Los alumnos reciben un código para consultar sus resultados individuales, pero durante su visita, el Equipo de Revisión de la OCDE se dio cuenta de que muchos alumnos no consultan sus resultados a través de Internet. Un estudio realizado por Mendoza Trejo (2010) revela que sólo 27% de los padres de los alumnos de escuelas primarias y 31% de los padres de los alumnos de las escuelas secundarias buscan los resultados de ENLACE a través de Internet (en un contexto donde 52% de los padres de alumnos de escuelas de educación básica desconocen los resultados de ENLACE). Esto sucede a pesar de los considerables esfuerzos de la SEP que envía reportes impresos a los padres con los resultados de sus hijos. Los resultados agregados por la escuela también están disponibles en línea y pueden ser consultados por el público en general. Los directores y maestros de las escuelas tienen acceso a los resultados individuales de los alumnos por componente de la prueba; además, pueden conocer las respuestas correctas

e incorrectas de cada educando, lo que les permite identificar los temas que son más difíciles para los alumnos (SEP e INEE, de próxima aparición). Los maestros también reciben reportes impresos con los resultados de sus grupos. Además, se envían a la comunidad educativa carteles con los resultados de la escuela y su comparación con otras escuelas, y deben exponerse en un lugar visible del plantel.

La evaluación puede ser de “alto impacto” o “bajo impacto” (como lo indica Messick, 1999). ENLACE se presentó inicialmente como una evaluación de “bajo impacto”, pero ha ido evolucionado poco a poco hasta convertirse en un esquema de “alto impacto”. Esto se debe a que se ha dado otros usos a sus resultados, siendo el más importante su uso en el Programa Nacional de Carrera Magisterial (PNCM). Como consecuencia, la asignación de incentivos económicos, que se suman al salario de los docentes, están vinculados a los puntajes de sus alumnos en ENLACE (consulte el análisis en el capítulo 4). Si bien ENLACE se ha utilizado en el contexto del PNCM desde su creación en 2006, su diseño original se basó en el uso formativo de sus resultados para la mejora de la enseñanza y el aprendizaje en el aula (consulte, por ejemplo, Zúñiga Molina y Gaviria, 2010).

Originalmente la SEP sugería los siguientes usos de “bajo impacto” para los resultados de ENLACE:

- Actividades para que los padres de familia participen y apoyen el aprendizaje de sus hijos, como talleres y reforzamiento del aprendizaje en el hogar.
- Creación de materiales para mejorar la enseñanza y fomentar el intercambio de experiencias y buenas prácticas entre los maestros y las entidades federativas.
- El desarrollo de las ofertas de formación continua, basadas en los resultados de ENLACE.
- Iniciativas especiales para fortalecer a las escuelas con bajos resultados en ENLACE.
- Acciones para mejorar la gestión educativa, como el establecimiento de estándares de aprendizaje al final de cada nivel educativo.
- Promoción de programas para mejorar la calidad de la educación (SEP e INEE, de próxima aparición).

Pero ENLACE se convirtió rápidamente en una evaluación de “alto impacto” con la publicación de los resultados en las escuelas, la divulgación de las clasificaciones de los planteles en los medios de comunicación, los incentivos monetarios para los docentes con base en los resultados de sus alumnos en ENLACE y el reconocimiento público de los alumnos con las calificaciones más altas. Estos otros usos de ENLACE son muy distintos de los previstos inicialmente. Pueden provocar que las escuelas, maestros y alumnos dediquen una gran cantidad de tiempo a practicar para las pruebas de ENLACE (Backhoff *et al.*, 2008; Loureiro, 2009), opinión que fue comunicada al Equipo de Revisión de la OCDE por un gran número de interesados, incluyendo las escuelas visitadas.

Durante la visita, el Equipo de Revisión de la OCDE fue testigo de la fuerte influencia de ENLACE en una amplia gama de aspectos de la vida escolar. Los ejemplos de conductas mencionadas constantemente por el personal entrevistado en las siete escuelas visitadas (y que fueron corroborados por un gran número de grupos de actores entrevistados durante la visita), incluyen:

- En todas las escuelas visitadas, los maestros y directores concedían gran importancia a los resultados de ENLACE.
- En todas las escuelas visitadas, los alumnos practicaban mucho para ENLACE utilizando las pruebas aplicadas en años anteriores.
- Algunas escuelas tratan de mejorar el logro de sus alumnos motivándolos a participar en programas promocionales desarrollados por las autoridades educativas, tales como

“Vamos por 600 puntos” (organizado por la Administración Federal de Servicios Educativos en el Distrito Federal, AFSEDF).

- Algunos maestros opinaron que como ENLACE no tiene consecuencias directas para los alumnos, estos no se esfuerzan por contestar la prueba. En consecuencia, como se advirtió al Equipo de Revisión de la OCDE en las escuelas visitadas, algunos maestros pueden tratar de motivar a los alumnos diciéndoles que los resultados de ENLACE contribuyen a su expediente de logros globales, lo cual no es cierto.
- Muchos docentes expresaron su preocupación acerca de que ENLACE no toma en cuenta las diferentes circunstancias y contextos que enfrentan escuelas y alumnos.
- Algunos maestros reconocen un efecto positivo en ENLACE, ya que “ejerce algo de presión para que los maestros hagan un esfuerzo por mejorar el desempeño de los alumnos”.
- ENLACE ocasionó que los maestros se preocuparan de alguna forma por los alumnos con un bajo desempeño. También generó una preocupación especial por la lectura de comprensión.

En cuanto a los alumnos con necesidades educativas especiales, se realizan algunas adaptaciones para administrarles la prueba. ENLACE permite que un familiar esté en el aula y ayude al estudiante sin darle las respuestas. En algunos casos, dependiendo del tipo específico de necesidades del alumno, es posible leerle las preguntas o ayudarlo a escribir en la hoja de respuesta (SEP e INEE, de próxima aparición).

En el caso de los alumnos indígenas, se supone que la pertinencia lingüística y cultural de las pruebas está asegurada gracias a la labor de acompañamiento de la DGEI (SEP e INEE, de próxima aparición). La DGEI forma parte del Consejo Técnico de ENLACE desde 2008 y ha trabajado con la DGEP para mejorar la pertinencia y confiabilidad de la prueba para los alumnos indígenas. Este trabajo ha implicado la revisión de más de 2 500 reactivos. Sin embargo, durante las visitas a las escuelas, los maestros argumentaron que muchas palabras incluidas son palabras o situaciones desconocidas por los alumnos de las zonas rurales e indígenas. En 2008 los maestros de una escuela del estado de Chiapas presentaron una denuncia contra la SEP en el Consejo Nacional para Prevenir la Discriminación (CONAPRED) por discriminación lingüística en la aplicación de ENLACE. La denuncia recibió el apoyo de organizaciones representativas como el Congreso Nacional de Educación Indígena e Intercultural. En 2011, el CONAPRED emitió una resolución a favor de los maestros, afirmando que la SEP debería introducir modificaciones a ENLACE para asegurarse de que no discrimine culturalmente a los alumnos indígenas. La SEP aceptó la resolución y actualmente está trabajando con organizaciones de educación de indígenas con el fin de adaptar ENLACE a las necesidades de los alumnos indígenas (CONAPRED, 2011).

### ***Iniciativas de política relacionadas con la evaluación de los alumnos dentro del aula***

En el marco de la Alianza por la Calidad de la Educación (ACE), se está llevando a cabo un esfuerzo importante de reforma curricular en México, a saber: la Reforma Integral de la Educación Básica (RIEB). Actualmente se está poniendo a prueba un plan de estudios unificado que articula los programas desde preescolar hasta el final de la educación básica, el cual está siendo introducido progresivamente en las escuelas. La reforma tiene como objetivo mejorar la coherencia de todo el sistema, centrándose en proporcionar a los alumnos una formación integral, de manera que puedan adquirir las habilidades necesarias para la vida (OCDE, 2011a; SEP, 2011c) (Capítulos 1 y 2 de esta publicación).

El nuevo plan de estudios está basado en un perfil explícito que debe ser alcanzado por todos los alumnos que egresan de la educación básica. Este perfil se compone de diez aspectos



que se espera alineen todos los esfuerzos curriculares y sirvan como un punto de referencia para evaluar la eficacia del proceso educativo (SEP, 2011c) (consulte el capítulo 2).

El Plan de Estudios 2011 establece que la enseñanza debe estar orientada a las competencias para la vida, por lo que debe comprender más que simplemente explicar los conceptos. Cada curso se organiza en cinco bloques y los “aprendizajes esperados” son indicados de manera explícita en cada bloque. Se ha trabajado mucho para establecer estándares curriculares que puedan ser utilizados como referencia para la evaluación externa y en el aula de los alumnos, al final de los principales ciclos de la educación básica: preescolar, tercer grado, sexto grado y educación secundaria. Los estándares curriculares expresan lo que los alumnos deben saber y ser capaces de hacer al final de cada ciclo. Existen estándares para las áreas formativas importantes: español, inglés, matemáticas, ciencias, capacidad de lectura en español y habilidades digitales. Los estándares se han establecido con la asesoría técnica de la Universidad de Londres y están alineados con el nivel de dominio 3 de PISA en lectura, matemáticas y ciencias, como un objetivo que deben alcanzar los alumnos que egresan de la educación secundaria (SEP, 2011c; SEP, 2011d).

A principios de 2011, se pusieron en marcha los trabajos para diseñar una nueva generación de pruebas ENLACE de acuerdo con los estándares de la RIEB y los resultados de aprendizaje esperados. Estas nuevas evaluaciones se pusieron a prueba en el ciclo 2011-2012 y se generalizarán para todos los alumnos en 2012-2013 cuando la RIEB se amplíe a todos los grados. Este trabajo ha reunido a SEB, DGAIR, INEE, autoridades educativas estatales, maestros y expertos, bajo la coordinación de investigadores de la Universidad Autónoma de Baja California.

En esta reforma se está haciendo un gran esfuerzo para establecer un nuevo enfoque de evaluación basado en el aula. Se pone más énfasis en la evaluación como un proceso continuo y una parte esencial del proceso enseñanza-aprendizaje. Los maestros siguen estando a cargo de evaluar los logros de los alumnos. Los documentos oficiales establecen explícitamente que en la educación básica, el enfoque formativo debe prevalecer en todas las actividades de evaluación, ya que el objetivo principal es mejorar el desempeño de los alumnos. Los maestros deben explicar claramente a los alumnos y padres de familia los resultados de aprendizaje esperados para los alumnos y los criterios de evaluación, así como las medidas que deben tomar los alumnos para superar sus dificultades. Los docentes no deben dar una calificación a los alumnos sin recomendarles cómo mejorar su desempeño (SEP, 2011d; SEP e INEE, de próxima aparición).

En cuanto a los instrumentos que se utilizarán para la evaluación en el aula, el plan de estudios especifica que deben utilizarse rúbricas, listas de control, registros de observaciones, trabajos escritos, proyectos de equipo, mapas conceptuales, portafolios y exámenes escritos y orales (SEP, 2011c; SEP e INEE, de próxima aparición). Los alumnos deben participar continuamente en su autoevaluación y en las actividades de evaluación de sus compañeros.

En relación con este nuevo énfasis en la evaluación dentro del aula, la SEP está trabajando en un nuevo enfoque para reportar las calificaciones a los alumnos y padres de familia. La boleta de calificación actual será reemplazada por una Cartilla de Educación Básica, que deberá combinar las perspectivas cualitativas y cuantitativas sobre el desempeño de los alumnos, y centrarse en su progreso en relación con los resultados de aprendizaje esperados para cada bloque curricular (SEP, 2011c; SEP, 2011d; SEP e INEE, de próxima aparición).

Las calificaciones se seguirán expresando con la escala numérica actual del cinco al diez, pero cada calificación llevará una descripción del nivel de desempeño alcanzado por el alumno en relación con los resultados de aprendizaje esperados. Los niveles de desempeño se clasificarán mediante letras: “A” para “sobresaliente” (correspondiente a un 10); “B” para “satisfactorio” (que corresponde a un 8 o 9); “C” para “suficiente” (correspondiente a un 6 o 7); y “D” para “básico” o “elemental”. Para cada calificación, la cartilla incluirá una declaración sobre el tipo

de apoyo necesario para el alumno, que deberá ser complementado con las observaciones de los maestros.<sup>4</sup> Este nuevo instrumento fue desarrollado por un grupo de trabajo formado por la SEP y el INEE en 2009. En este momento se está probando en 5 000 escuelas primarias y 1 000 escuelas secundarias.

## Fortalezas

### *Los docentes están comprometidos con el aprendizaje de los alumnos*

Durante las visitas a las escuelas, el Equipo de Revisión de la OCDE coincidió en que la mayoría de los maestros mexicanos están realmente involucrados en el aprendizaje y el logro de sus alumnos. En todas las reuniones, los alumnos mencionaron consistentemente el empeño de sus maestros para explicarles las materias y para ayudar a sus compañeros con bajo desempeño académico. El Equipo de Revisión de la OCDE presenció casos en que los docentes dedicaron su tiempo para impartir clases y actividades extras y reunirse con los padres de familia, de forma voluntaria, los sábados o después de horas laborales. Los maestros también son conscientes de la importancia de adaptar la enseñanza y la evaluación a las necesidades individuales y los contextos culturales de los alumnos —aunque la forma en que lo hacen no siempre es la adecuada, como se mencionará más adelante. En algunas escuelas, los docentes pedían a los alumnos avanzados que actuaran como monitores para ayudar a sus compañeros de clase.

En México, los maestros juegan un papel importante en la evaluación de los alumnos, ya que tanto la evaluación formativa continua como la evaluación sumativa son una parte esencial de sus responsabilidades profesionales. La evaluación es parte integral de su trabajo. La evidencia del aprendizaje de los alumnos se reporta de manera regular. Para su evaluación se tiene en cuenta varios aspectos: tareas, esfuerzo, presentaciones, pruebas y proyectos. Sin embargo, se puede mejorar aún más en la manera como los docentes aplican la retroalimentación formativa y en la forma como se asignan las calificaciones.

### *Se está desarrollando un nuevo marco integral para la evaluación de los alumnos dentro del aula, con mayor énfasis en los resultados educativos*

La RIEB está aportando a la educación un enfoque fundamentado para la evaluación dentro del aula. Constituye un esfuerzo importante e inteligente para conformar el plan de estudios en toda la educación obligatoria, desde preescolar hasta el final de la educación secundaria. La RIEB organiza los programas de estudio con un enfoque integral y coherente manteniendo espacio y flexibilidad para las adaptaciones a los diferentes niveles educativos, orígenes y necesidades especiales de los alumnos. El nuevo plan de estudios parece tener un buen equilibrio entre lo que es obligatorio y lo que se puede adaptar a nivel local.

Una de las principales características del enfoque de la RIEB es la alineación curricular en torno al desempeño esperado de los alumnos. Cada “bloque” de cada curso incluye los “resultados de aprendizaje esperados”, que serán la guía principal para el trabajo de los maestros. Los resultados de aprendizaje esperados están alineados con los estándares curriculares establecidos al final de cada uno de los ciclos de tres años (preescolar, primero a tercer grado, cuarto a sexto grado y séptimo a noveno grado). Los estándares curriculares están alineados con un perfil de egreso definido para los alumnos al final de la educación

básica (es decir, la educación secundaria), que incluye el nivel de desempeño de PISA 3 como componente principal (SEP, 2011c).

Al unificar los esfuerzos curriculares alrededor de los resultados de aprendizaje esperados, la RIEB está generando un movimiento positivo que va de un plan de estudios basado en contenidos a uno orientado a las competencias. Esto constituye un importante paso adelante para México. El hecho de tener en cuenta las competencias de PISA como parte del perfil de los alumnos egresados es una directriz inteligente, en el sentido de que las autoridades están enviando un mensaje unificado a los maestros sobre lo que se espera que generen en cuanto a competencias.

La RIEB también incluye un enfoque claro e interesante para la evaluación de los alumnos en el aula, tanto con fines formativos como sumativos, que está establecido expresamente como responsabilidad profesional de los maestros. Como se puntualizó antes, la RIEB amplía el significado de la evaluación, concibiéndola como una parte esencial de la enseñanza y el aprendizaje; propone el uso de una amplia gama de instrumentos de evaluación; enfatiza el propósito formativo de la evaluación dentro del aula; además, como se explica a continuación, introduce un cambio fundamental al dar un nuevo significado a las calificaciones.

Según lo observado por el Equipo de Revisión de la OCDE durante las visitas a las escuelas, los maestros que han conocido el nuevo plan de estudios —a través de la participación directa en las experiencias piloto o por medio de cursos de formación continua sobre la RIEB— están empezando a introducir nuevas prácticas en sus aulas: listas de control orientadas a aspectos como la participación, la colaboración y las actitudes de investigación de los alumnos; rúbricas; nuevos productos auténticos que se exigen a los alumnos (por ejemplo, elaborar un anuncio como tarea escrita), proyectos orientados hacia los resultados de aprendizaje esperados; la autoevaluación de los alumnos y la evaluación por sus compañeros.

### ***Se ha avanzado en la alineación de las calificaciones con los resultados de aprendizaje esperados***

Otro cambio importante que la RIEB está introduciendo en la evaluación dentro del aula se refiere a los cambios en la boleta de calificaciones (ahora llamada “Cartilla de Educación Básica”). Es significativa la intención de dar un nuevo significado a las calificaciones en términos de resultados de aprendizaje esperados.

El enfoque actual es muy inteligente, en el sentido de que no está en el cambio de la escala, sino en dar un nuevo significado a las calificaciones actuales en términos del desempeño de los alumnos. En algunas reformas de las escalas de desempeño, se dedica una gran cantidad de tiempo y energía a cambiar la escala, por ejemplo, de números a letras. Esto es por lo general sólo un cambio cosmético, pues las prácticas de calificación siguen siendo las mismas. El problema principal con la asignación de calificaciones es la ausencia de un significado para cada una y la práctica de calificar como una cuestión de asignar puntos por diversos aspectos, promediándolos a una calificación final sin sentido (Wiggins, 1998; Ravela, 2009a). En México, actualmente se está trabajando para establecer una relación entre las calificaciones de la escala actual y los niveles de desempeño, aunque todavía no se han desarrollado las descripciones específicas. La nueva Cartilla de Educación Básica está en fase piloto y aún queda trabajo por hacer para definir el significado específico de cada nivel de desempeño para cada una de las asignaturas, de manera que las calificaciones adquieran un significado claro para los maestros, alumnos y familias.

Sin embargo, también hay que decir que el nuevo sistema de reporte de calificaciones incluye algunos aspectos discutibles, como los indicadores de la velocidad de lectura y comprensión

(una prueba específica requerida que consiste en medir el número de palabras que un alumno es capaz de leer en un minuto) y la práctica de promediar las calificaciones para diferentes asignaturas en la educación secundaria.

### ***Se está atendiendo adecuadamente la reducción de la reprobación en la educación primaria***

La reprobación es un problema importante en América Latina. Los análisis de la educación primaria indican que Guatemala (26%) y Colombia (21%) presentan los mayores porcentajes de alumnos de 7 a 11 años de edad que van retrasados (es decir, no están en el grado que corresponde a su edad). Los porcentajes para Brasil, Uruguay, Chile y Argentina son 11%, 10%, 9% y 6%, respectivamente (OEI, 2010). El porcentaje para México (3%) fue el más bajo de los 17 países.

Es bien sabido que: (i) los alumnos que repiten un grado continúan teniendo un bajo desempeño durante sus estudios (es decir, reprobar un grado no tiene el supuesto efecto de nivelar a los repetidores a la altura de sus compañeros de clase); y (ii) los alumnos que repiten un grado aumentan su probabilidad de abandonar la escuela. Muchos países han tratado de eliminar la reprobación, pero en algunos casos el efecto ha sido contraproducente porque los alumnos avanzan en sus estudios sin haber adquirido el aprendizaje esperado (Torres, 1995; Schiefelbein y Wolff, 1992).

En México, los alumnos no pueden repetir entre el primero y el segundo grado de educación primaria (SEP, 2011a). En el contexto de la RIEB, esta regla se extenderá al tercer grado. Se hará hincapié en oportunas acciones compensatorias para los alumnos con un bajo desempeño, en lugar de la reprobación de grado. Un hecho notable es que, al observar los datos internacionales de desempeño de los alumnos (Laboratorio Latinoamericano para la Evaluación de la Calidad de la Educación de la UNESCO, LLECE), México está por encima del promedio regional en lectura y matemáticas, tanto en tercero como en sexto grado: el 25% de alumnos de tercer grado están en o por debajo del nivel uno de lectura (por debajo del promedio regional de 32%). Sólo Cuba, Costa Rica y Uruguay tienen un mejor desempeño que el de México (UNESCO/OREALC, 2008). Así que la evidencia sugiere que el enfoque de México en cuanto a la reprobación está funcionando bien, ya que los alumnos no se están retrasando y, al mismo tiempo, están logrando un buen nivel de desempeño en el contexto regional de América Latina.

Las tasas de reprobación de México son más altas en la educación secundaria, donde 22% de los alumnos de 15 años de edad, han repetido al menos un grado durante su escolaridad (OCDE, 2010). Sin embargo, con ese 22%, México, tiene en realidad el porcentaje más bajo dentro de los países latinoamericanos que participan en PISA. En países como Argentina, Brasil, Colombia y Uruguay, alrededor de 35% de los alumnos de 15 años han repetido al menos un grado (OCDE, 2010). La cifra de México es también inferior a las de España y Portugal (ambas alrededor de 35%) (Consulte el anexo D).

### ***Se está fomentando la participación de los padres en el aprendizaje de sus hijos***

Durante sus reuniones y visitas a las escuelas, el Equipo de Revisión de la OCDE observó que, en México, los directores de escuela y maestros están muy conscientes de la importancia de la comunicación frecuente con los padres de familia. Aunque es posible que en algunos casos los padres sean contactados sólo por cuestiones administrativas, en todas las escuelas

visitadas había una preocupación por comunicar las calificaciones a los padres cada dos meses y por involucrarlos para que apoyaran a los alumnos con bajo desempeño. También observaron notables iniciativas para que los padres participen en los esfuerzos actuales para fomentar la lectura y promover su participación activa en los Consejos Escolares de Participación Social en la Educación.

La implementación actual de la nueva Cartilla de Educación Básica también es importante, ya que tiene la intención de dar un significado más claro a las calificaciones con respecto a los resultados de aprendizaje esperados de los alumnos. Además, es encomiable que la nueva Cartilla aliente el apoyo de los padres en el aprendizaje de sus hijos.

### ***Algunas iniciativas fomentan el desarrollo de las competencias de los docentes para la evaluación de sus alumnos***

Como ya se puntualizó, existen algunas iniciativas incipientes pero importantes para fomentar el desarrollo de competencias docentes para la evaluación dentro del aula. Este aspecto está cada vez más presente entre las ofertas de formación continua, aunque los maestros todavía no están satisfechos con la calidad de los cursos ofrecidos. Recientemente, el INEE y la SEP comenzaron a trabajar en el desarrollo de instrumentos para la evaluación en el aula, que estarán disponibles en Internet. El INEE está desarrollando herramientas para los docentes y distribuirá un millón de copias entre los maestros. Estas iniciativas deben apoyar la implementación de la RIEB y el nuevo enfoque de la evaluación. Sin embargo, es importante señalar que no basta con distribuir materiales. El reto es asegurar que los maestros tengan la oportunidad de interactuar colectivamente con los nuevos materiales en sus escuelas y probar las nuevas prácticas en las aulas con sus colegas.

### ***Es notable la capacidad de diseñar e implementar evaluaciones a gran escala***

La evaluación externa de los alumnos es una práctica común en las escuelas mexicanas. Por más de 20 años, IDANIS se ha administrado a los alumnos que egresan de la escuela primaria. Lo mismo ocurre con EXANI I y los alumnos que terminan la educación secundaria, desde 1994. Durante la década de 1990, se llevaron a cabo grandes evaluaciones estandarizadas nacionales en el marco del Programa Nacional de Carrera Magisterial (PNCM). Además, EXCALE y ENLACE se pusieron en marcha en 2005 y 2006, respectivamente. Millones de evaluaciones estandarizadas son administradas, procesadas y reportadas en México, cada año.

Como se mencionará más adelante, a pesar de que la diversidad de las evaluaciones externas de los alumnos es algo confusa para los maestros, contribuyen a enfocar su atención en el logro de los alumnos. Los maestros analizan los resultados de cada alumno y con frecuencia se concentran en aquellos que tienen resultados más bajos en ENLACE. Otro efecto ha sido que los maestros se han concientizado de la importancia de la lectura de comprensión en la vida de los estudiantes y la habilidad para responder exámenes. Varios docentes mencionaron la importancia de que los alumnos comprendan el texto de las preguntas de modo que, por ejemplo, entiendan lo que se pide en los problemas de matemáticas. Muchos maestros sienten que los resultados de ENLACE desafían su competencia profesional, porque los presiona para mejorar el desempeño de sus alumnos (medido por ENLACE).

IDANIS y EXANI I también juegan un papel en la motivación de los alumnos y maestros para el logro. Sin embargo, hay algunas cuestiones que deben abordarse en torno a la alineación entre las diferentes evaluaciones externas y el nuevo plan de estudios.

Cada una de las evaluaciones externas implica mucho trabajo técnico de varios comités encargados de la elaboración de instrumentos y el procesamiento de datos. Se cuenta con una gran capacidad logística instalada. Esta capacidad no sólo existe dentro de la SEP y el INEE sino también dentro de CENEVAL, el cual cuenta con una amplia experiencia en la organización de evaluaciones estandarizadas. La experiencia con la evaluación externa se ha basado casi exclusivamente en pruebas de opción múltiple, que dejan espacio para la introducción de tipos más complejos de tareas que podrían enviar nuevas e importantes señales pedagógicas a los maestros, alumnos, padres de familia y escuelas.

## Retos

### *Es necesario cambiar la cultura de la enseñanza*

Mejorar el método de evaluación de los maestros implica un cambio fundamental en las prácticas de enseñanza establecidas. Según lo declarado por Stigler y Hiebert (2009), las prácticas docentes son difíciles de cambiar porque son parte de una cultura. Los maestros aprenden a enseñar principalmente a través de su participación informal en la escuela y sus prácticas en el aula, durante largos períodos — y en menor medida a través de los programas de formación inicial o continua. La enseñanza es una práctica que se aprende al vivir en una cultura más que a través de estudios formales. La práctica docente en las escuelas está determinada más bien por guiones culturales que pueden equipararse al ADN de la enseñanza. Por estas razones, el cambio en las prácticas educativas es un proceso complejo y de largo plazo (Stigler y Hiebert, 2009). Teniendo esto en cuenta, es crucial para los esfuerzos de México cambiar las prácticas de evaluación y aplicar las modificaciones curriculares. No es sólo una cuestión de escribir un nuevo plan de estudios, mejorar la boleta de calificaciones, ofrecer nuevos cursos de formación continua o entregar nuevos materiales a las escuelas, sino de promover un cambio cultural significativo.

El sistema educativo mexicano está muy centralizado. Las escuelas y los docentes son objeto de control por parte de los supervisores, las autoridades estatales y federales y el sindicato de maestros. La simulación es una parte importante de la cultura del sistema ya que, a mayor control, mayor simulación se necesita para cumplir con las expectativas del trabajo. Durante la visita de revisión, la palabra “simulación” fue mencionada con frecuencia como parte de la cultura por varios actores: autoridades, supervisores, directores de escuela, docentes e investigadores. Este hecho debe tenerse muy en cuenta por los líderes del cambio a nivel central. En ese tipo de cultura es aún más difícil introducir cambios reales por decreto en las aulas.

### *Es común enseñar para la prueba en todo el sistema escolar*

Enseñar para la prueba se ha convertido en una práctica pedagógica generalizada en México. Los directores de escuela, maestros y alumnos consideran que practicar para las pruebas estandarizadas es la mejor estrategia para elevar el logro de los alumnos. Este mensaje fue comunicado constantemente por el personal entrevistado en las siete escuelas visitadas por el Equipo de Revisión y confirmado por muchos de los grupos de interesados que fueron entrevistados durante la visita de revisión. Las escuelas motivan a los alumnos para mejorar

sus resultados de ENLACE, incluso mediante la participación en los programas oficiales, como “Vamos por 600 puntos” (organizada por la Administración Federal de Servicios Educativos en el Distrito Federal, AFSEDF). Varias semanas antes de la aplicación de ENLACE, los maestros dedican un tiempo de clase considerable a lo que ellos llaman el “Pre-ENLACE”, que son sesiones de prueba similares al examen real con ejemplos de ejercicios anteriores. Algo similar ocurre con IDANIS y EXANI I. En algunas escuelas, los maestros y alumnos dedican mucho tiempo a practicar para estas pruebas. Si bien los resultados de ENLACE tienen consecuencias para los docentes y directores —parte de su salario depende de ellos (consulte el capítulo 4)— IDANIS y EXANI I son importantes para los alumnos y padres de familia ya que la posibilidad de asistir a una escuela seleccionada depende de sus resultados. Del mismo modo, la SEP también exhorta a las escuelas para que preparen a los alumnos para las pruebas de PISA. La publicación “Hacia PISA 2012” es ampliamente distribuida en una versión para docentes (250 000 ejemplares) y otra para alumnos (cuatro millones de ejemplares), para animar a los alumnos y docentes de secundaria a que practiquen cada semana con los reactivos de PISA, durante todo el curso escolar. Obtener buenos resultados en PISA se presenta a los alumnos como una cuestión de orgullo nacional y lealtad a México (SEP, 2011e y 2011f).<sup>5</sup>

El problema fundamental que debe ser considerado con atención es si el objetivo es simplemente elevar los puntajes promedio (en ENLACE o PISA) o, por el contrario, mejorar la amplia gama de competencias de los niños mexicanos. Es muy importante diferenciar el objetivo del indicador. Como lo expresan Linn y Gronlund (2000):

*Casi siempre estamos interesados en hacer inferencias que van más allá de la prueba específica. Nos gustaría, por ejemplo, poder decir algo sobre el grado de comprensión de los conceptos matemáticos, basados en la puntuación que se obtiene en un examen de este tipo. Debido a que los reactivos de una prueba sólo muestran el dominio en cuestión, la calificación de la prueba y la inferencia sobre el grado de comprensión no son los mismos. Se necesita una generalización y es la generalización, no la calificación de la prueba en sí, la que es importante. Cuando los reactivos específicos de la prueba son enseñados, la validez de la inferencia sobre el nivel de logro del estudiante se ve amenazada.*

La estrategia para fomentar la práctica de los reactivos de ENLACE y PISA entre los alumnos puede ocasionar un aumento de las puntuaciones en el corto plazo, pero no está claro si esto conducirá a un mejor aprendizaje sostenido de la población estudiantil, en el largo plazo. Desafortunadamente, el riesgo es que las prácticas de enseñar para la prueba refuercen la creencia cultural de que practicar las pruebas estandarizadas es el camino para mejorar el aprendizaje. Esto se agrava por los incentivos que tienen los docentes y directores de escuela por promover la enseñanza de la prueba, ya que los resultados de ENLACE tienen un impacto directo en sus salarios (como se menciona con mayor detalle en el capítulo 4). Estos incentivos no están necesariamente alineados con las mejores estrategias para sostener el proceso de mejora de los conocimientos, habilidades y actitudes de los alumnos. Se descuidan importantes objetivos educativos que no se evalúan en las pruebas. Como las pruebas estandarizadas abarcan un rango limitado de competencias y habilidades del plan de estudios, enseñar para la prueba limita las experiencias educativas de los alumnos (consulte Morris, 2011, y Rosenkvist de 2010, para una discusión). Sería lamentable que estas prácticas socavaran los muchos efectos positivos que ENLACE ha tenido en el sistema educativo mexicano, tales como el énfasis en la mejora de resultados de los alumnos, la mayor atención a los alumnos con dificultades de aprendizaje y la transparencia de los resultados para los interesados en la educación (como se mencionó más arriba en este capítulo y en el capítulo 2).

### ***Hay una excesiva dependencia de los exámenes de opción múltiple***

Un problema importante en la evaluación externa de los alumnos en México es el uso casi exclusivo de exámenes de opción múltiple, que podría ocasionar distorsiones en la educación de los niños. La prevalencia de este tipo de exámenes en la evaluación externa debería preocupar a las autoridades educativas, ya que esto puede reducir la visión de los alumnos sobre lo que es el desempeño. El éxito de la RIEB requiere la introducción de tareas más complejas en las evaluaciones externas, así como una mayor variedad de evaluaciones.

Del mismo modo, IDANIS y EXANI I deberían introducir una mayor variedad de preguntas, de manera que puedan evaluar las competencias promovidas por la RIEB. El argumento de que estas pruebas evalúan “habilidades” en lugar de contenidos curriculares es bastante anticuado. Si las autoridades mexicanas desean mantener estos instrumentos con fines de selección (lo cual, por sí mismo, es discutible) o simplemente como una herramienta administrativa eficaz para asignar a los alumnos a las escuela secundarias (como parece ser el caso del Distrito Federal), debe alinear estas evaluaciones con los resultados de aprendizaje esperados y los estándares de la RIEB, para reforzar el mensaje unificado para los maestros y alumnos sobre lo que es importante en el aprendizaje.

### ***ENLACE tiene demasiados objetivos***

Como se explicó antes, si bien ENLACE era originalmente un instrumento de evaluación diagnóstica y formativa, nuevos objetivos y consecuencias fueron añadidos después, siendo el más importante el uso de sus resultados para proporcionar incentivos monetarios a los maestros y directores de escuela (consulte el capítulo 4 para más detalles). Según lo explica Linn (2000), los sistemas de evaluación que son útiles para propósitos formativos suelen perder gran parte de su credibilidad cuando se les asocia con consecuencias de alto impacto, porque los efectos no deseados a menudo acaban por prevalecer sobre los efectos positivos esperados.

Durante la visita de revisión, maestros, directivos, alumnos y expertos en educación plantearon algunos casos de los efectos indeseables de ENLACE. Los ejemplos incluyen el tiempo invertido en la preparación especial para la prueba de ENLACE que se reduce del plan de estudios regular; practicar reactivos de la prueba sin analizarlos a profundidad; y las dificultades para asegurar la integridad de la administración de la prueba. Como se documenta en Zúñiga Molina y Gaviria (2010) y OCDE (2011b), los experimentos realizados en México sugieren que las posibilidades de copia en la prueba pueden ser altas — entre 4%-10% de los porcentajes globales de casos probables de copia para los grados 3° a 6° en el período de 2006 a 2009. Como señalan Zúñiga Molina y Gaviria (2010) “Cada año, la prueba ENLACE se vuelve más importante en términos de impacto social y mediático. Esto podría reflejarse en el aumento de conductas de engaño”. En su análisis sobre las oportunidades para un mayor desarrollo de ENLACE para incentivos y evaluación docente en México, Zúñiga Molina y Gaviria (2010) concluyen:

*... desde su implementación inicial, los resultados del programa [ENLACE] han sido utilizados para distintos fines, a pesar de las repetidas advertencias que aparecen en los documentos oficiales del programa en cuanto a la necesidad de evitar algunos de esos usos. En ese sentido, se puede argumentar que algunos de los usos más visibles del programa están relacionados con los propósitos que han sido considerados inadecuados por los encargados del diseño y operación del programa. Tal es el caso, por ejemplo, del uso frecuente de ENLACE,*



*en los medios de comunicación, para establecer clasificaciones nacionales o estatales de las escuelas, basadas exclusivamente en los puntajes promedio obtenidos por los alumnos en cada escuela en particular; o el uso ocasional que los funcionarios de educación y otras partes interesadas han hecho de los resultados de las pruebas como si fueran un indicador inequívoco de la calidad del trabajo realizado por los docentes.*

Además, aunque EXCALE es la evaluación de los alumnos específicamente diseñada para monitorear los objetivos de aprendizaje del sistema (a través del tiempo y en todas las entidades federativas) (consulte el capítulo 6 para más detalles), en la práctica, esta función también es llevada a cabo por ENLACE como consecuencia del uso que dan a los resultados las autoridades educativas y los medios de comunicación. Esto no es deseable ya que ENLACE no sólo evalúa a los alumnos en un contenido curricular mucho más limitado que EXCALE, sino que además utiliza un número menor de reactivos. Como lo expresó uno de los actores entrevistados por el Equipo de Revisión de la OCDE, “El censo mata a la muestra”. ENLACE se ha vuelto mucho más visible que EXCALE y se está utilizando en todos los niveles, con una gran diversidad de objetivos (función formativa y de diagnóstico con los alumnos, evaluación del sistema, rendición de cuentas de las escuelas, rendición de cuentas del Estado, incentivos monetarios para los maestros y directores de escuela).

Como se indicó antes, ENLACE ha aportado considerables beneficios, como la atención del maestro en el desempeño de los alumnos, en particular el de los alumnos con bajo desempeño, o una mayor conciencia de la importancia de la lectura de comprensión. Sin embargo, los efectos no intencionales de ENLACE parecen ser importantes. A pesar de las grandes cantidades de datos recabados, su uso con fines formativos no está claro. Lo anterior exige una reflexión seria sobre los usos que deben tener los resultados de ENLACE para que sus múltiples beneficios no sean minados.

### ***Los maestros tienen un enfoque limitado de la docencia y la evaluación formativa***

Si bien los maestros son conscientes de la importancia de la dimensión formativa de la evaluación dentro del aula, el Equipo de Revisión de la OCDE encontró poca evidencia de que se aplique de manera adecuada. Parecería que los docentes no entienden el significado pleno de la evaluación formativa. Dar retroalimentación a los alumnos se concibe como proporcionarles calificaciones o puntos para una tarea, informarles si su trabajo es o no aceptable, o pedirles que revisen su trabajo o hagan un esfuerzo extra (Ravela, 2009a). Todos estos conceptos están profundamente arraigados en la cultura de los maestros. La evaluación formativa consiste básicamente en dar a los alumnos una idea general acerca de lo que estuvo mal en la prueba o la tarea evaluada. Según lo declarado por los alumnos en las entrevistas con el Equipo de Revisión de la OCDE, “nos explican lo que hicimos mal” y “nos dicen lo que tenemos que mejorar”.

La evaluación formativa es en esencia una cuestión de retroalimentación adecuada. Debe impregnar el proceso de enseñanza-aprendizaje en lugar de ser algo que sucede después del aprendizaje. En el mismo sentido, la retroalimentación y su uso no debe ser algo que se produzca entre evaluaciones. Debe interpretarse en el núcleo de lo que se está evaluando. La retroalimentación debe ser continua e inmediata (Wiggins, 1998). Durante las visitas a las escuelas, el Equipo de Revisión de la OCDE percibió que en la cultura de los maestros mexicanos falta reflexión sobre estos temas. Hay poca conciencia de la importancia de dar retroalimentación durante el proceso de aprendizaje y poco conocimiento de los procesos cognitivos del aprendizaje. Por ejemplo, al analizar los resultados de las pruebas, los docentes

se centran en los reactivos en lugar de los procesos cognitivos implicados. También se encontró poca evidencia de que los alumnos reflexionaran sobre sus propios procesos de aprendizaje y tareas auténticas (Wiggins, 1998). Como se afirma en un estudio publicado recientemente, en México existe la necesidad de que los maestros proporcionen una retroalimentación más descriptiva a los alumnos (García *et al.*, 2011, pp. 33-34).

Las prácticas de autoevaluación y evaluación entre pares también están en una primera etapa de desarrollo (García *et al.*, 2011). Los maestros rara vez muestran ejemplos de buen trabajo a los alumnos de manera que puedan entender el nivel de desempeño que se espera de ellos y evaluar su propio trabajo. Además, los docentes no suelen mostrar ejemplos de trabajos deficientes que poco a poco se transformaron en trabajos de buena calidad (García *et al.*, 2011, pp. 75-76).

Los supervisores, directores y asesores técnicos pedagógicos (ATP) también tienen un enfoque limitado de la observación en el aula. Su preocupación principal es controlar los aspectos administrativos de la enseñanza. Entre ellos la puntualidad de los maestros, el control de asistencia de los alumnos, recordar el tema de la clase anterior, escribir en el pizarrón el tema de la clase actual, controlar la disciplina en el aula, facilitar la participación de los alumnos en las actividades del aula, revisar las tareas de los alumnos y el manejo de contenidos en los 50 minutos de clase. Los aspectos más específicos de los procesos de enseñanza- aprendizaje reciben menos atención.

### ***Las prácticas de calificación carecen de significado pedagógico***

Como se describió antes, las prácticas actuales de calificación están profundamente arraigadas en la cultura docente. En México, la asignación de calificaciones consiste en otorgar puntos a los alumnos por una serie de tareas y comportamientos, para después promediarlos y obtener una calificación total. Los puntos por el trabajo de los alumnos son asignados de manera normativa, mediante la comparación de los alumnos dentro de las clases y el otorgamiento del mayor número de puntos a los “mejores” alumnos, independientemente de si los estándares son realmente excelentes.

Las calificaciones obtenidas al promediar los puntos asignados por una serie de tareas y conductas realizadas, por lo general, conducen a un “número importante” sin algún significado claro (Ravela, 2009a). Los alumnos “ganan” puntos por asistir a clases, hacer tareas, participar en trabajos en equipo, dar presentaciones y presentar exámenes. El maestro asigna los puntos de una manera muy subjetiva, ya que no existen rúbricas que especifiquen el sentido de los puntos otorgados a cada una de estas actividades. Como resultado, un “7” puede significar cosas muy distintas en diferentes escuelas, en cada aula de una escuela y, peor aún, para diferentes alumnos del mismo salón y para el propio alumno en distintos momentos del año escolar. Cada maestro determina los criterios de calificación, pero esto no significa que docente será consistente en la aplicación de sus criterios entre los alumnos y con el tiempo. Además, como no existe una definición clara que describa el tipo de desempeño que debe alcanzarse, los alumnos no pueden entender lo que se espera de ellos. Por lo que obtener una calificación alta se convierte en su objetivo principal, lo que posiblemente distorsiona los esfuerzos educativos (Shepard, 2006). Los padres de familia también están más preocupados por las calificaciones que por el aprendizaje real de sus hijos.

Un tema que merece especial atención es la asignación de calificaciones mediante la comparación de los alumnos en el aula, en ausencia de estándares explícitos o puntos de referencia, lo que lleva a los docentes a adaptar sus expectativas, sus necesidades cognitivas y su enseñanza, al “nivel” actual de los alumnos en su salón de clases. El resultado es enseñar

menos a los más desfavorecidos (el problema de las adaptaciones a diferentes poblaciones y contextos culturales se detallará más adelante).

La práctica de combinar la evaluación del esfuerzo y la motivación con la evaluación del desempeño real es también una cuestión que debe abordarse, porque atenta contra las calificaciones académicas como indicadores de desempeño (Ravela, 2009a; García *et al.*, 2011). Esta práctica también ocasiona que los alumnos simulen sus esfuerzos, ya que aprenden rápidamente cómo comportarse con cada maestro para dar una buena impresión. Por último, las calificaciones se convierten en un instrumento disciplinario para los docentes, en lugar de una herramienta que informe acerca del aprendizaje. Controlar el comportamiento de los alumnos con las calificaciones no es lo mismo que crear un entorno de aprendizaje motivador (Shepard, 2006).

Todas estas características de la asignación de calificaciones están arraigadas de manera sólida y profunda en la cultura de los docentes. Al igual que sucede con todas las prácticas culturales, cambiarlas será todo un reto. Es probable que no funcione una resolución central estandarizada para todo el país. Las escuelas y los maestros deben tener espacio para hacer su trabajo y probar nuevos enfoques para asignar calificaciones dentro del marco nacional. Es necesario discutir, a nivel local y escolar, los diferentes procesos involucrados: disciplina, asistencia, tareas, pruebas, el tipo de tareas, etc. Como se dijo anteriormente, también es crucial desarrollar una visión más sofisticada de los procesos cognitivos. Deben participar todos los actores de todos los niveles: alumnos, padres de familia, maestros, directores, supervisores, ATP y jefes de sector; de lo contrario, persistirán las normas culturales actuales para asignar calificaciones.

### ***Los instrumentos para reportar las calificaciones necesitan mejorar más***

Los instrumentos utilizados para registrar y notificar los logros de los alumnos son de importancia fundamental en todo sistema educativo y en todos los niveles. Desempeñan varias funciones: comunicar el aprendizaje esperado a los alumnos, maestros y padres de familia; motivar y orientar los esfuerzos de los alumnos.

Como ya se mencionó, algunos cambios importantes en las boletas de calificaciones están siendo promovidos por la SEP en el marco de la RIEB. Sin embargo, todavía hay varios aspectos que deben atenderse. En primer lugar, se necesitan descripciones más específicas y detalladas sobre el desempeño de los alumnos en los diferentes niveles. Es fundamental asegurar que los padres de familia y alumnos puedan entender la información proporcionada por la boleta de calificaciones.

En segundo lugar, el enfoque de la nueva Cartilla de Educación Básica para exponer observaciones y orientación para los alumnos y padres de familia parece estar demasiado estandarizado. Es necesario concebir un instrumento más flexible, con más espacio para que los docentes comuniquen lo que esperan y cómo puede mejorar el estudiante.

En tercer lugar, la nueva Cartilla de Educación Básica incluye algunos aspectos preocupantes, en particular los estándares para la velocidad de lectura. La justificación para que este indicador tenga tanta visibilidad no está clara. Durante las reuniones con el Equipo de Revisión de la OCDE, varios alumnos expresaron su ansiedad y señalaron la falta de sentido de tal medida: “Leer rápido es más difícil, ya que no puedo entender lo que estoy leyendo”. La forma en que la lectura de comprensión se presenta en la Cartilla de Educación Básica también es motivo de preocupación. El desempeño se considera adecuado si el alumno responde correctamente tres de cuatro preguntas.<sup>6</sup> Sin embargo, no se dice nada sobre el tipo de texto, su extensión y complejidad, ni de la dificultad de las preguntas a ser contestadas. Es evidente que se requiere de un enfoque más sofisticado para la evaluación de la lectura.

Por último, cabe señalar que para la educación secundaria, la práctica de promediar las calificaciones de diferentes asignaturas y unidades, la cual no tiene un significado real, se mantiene como un elemento central en el esquema de la asignación de calificaciones. Además, promediar las calificaciones de cada bimestre para obtener una calificación final única no reconoce el avance de los alumnos durante el año (García *et al.*, 2011).

### ***La evaluación de los alumnos se da con poca interacción entre los maestros***

Los maestros rara vez se reúnen en torno a temas de evaluación de los alumnos. No suelen interactuar en la elaboración de instrumentos de evaluación ni en el desarrollo de criterios de calificación. De hecho, el intercambio de prácticas en el aula es muy poco común entre los maestros mexicanos. En las escuelas, los maestros trabajan en relativo aislamiento, aunque esta práctica no es tan frecuente en las escuelas primarias. Visitarse en las aulas es poco común. Según la encuesta TALIS, en México sólo el 27.5% de los maestros participa en una red de desarrollo profesional, siendo el tipo de actividad profesional menos mencionada (consulte el anexo D)<sup>7</sup>. En apariencia, sólo se presenta en pequeñas escuelas multigrado, donde los maestros comparten materiales, experiencias e instrumentos de evaluación, de manera sistemática.

La mayoría de los profesores no intercambian sus instrumentos de evaluación con los demás, no discuten resultados de aprendizaje esperados y no desarrollan un enfoque común para las calificaciones. Esta parece ser otra de las características culturales de la enseñanza en México, que podría convertirse en un obstáculo para la RIEB y la aplicación del nuevo método de evaluación en el aula. Por otra parte, los maestros parecen no comprender por completo la necesidad de aprender más acerca de la evaluación de los alumnos. Cuando se les preguntó sobre las áreas en las que tienen un “alto nivel de necesidad” de desarrollo profesional, sólo el 15% mencionó “las prácticas de evaluación de los alumnos” y el 13.7% mencionó “estándares de desempeño y contenidos” (consulte el anexo D).

La evaluación de los alumnos no se está utilizando como una actividad de desarrollo profesional docente ni como un camino para que mejoren su criterio profesional. La ponderación, que implica el trabajo auténtico del alumno, no está del todo desarrollada, siendo que debería ser una estrategia clave para la formación docente. Lo mismo sucede con la construcción de una nueva cultura de evaluación en el aula. La ponderación de la evaluación y la calificación también pueden vincularse con las prácticas en el aula.

### ***Falta consistencia para evaluar a los alumnos en todas las escuelas y clases***

Las escuelas no tienen criterios explícitos para asignar calificaciones y, por lo general, no documentan sus métodos para evaluar a los alumnos. Este hecho, aunado a la ausencia de procedimientos de ponderación para alinear el significado de las calificaciones asignadas por los maestros, ocasiona que el significado de las calificaciones difiera de una región a otra, de una escuela a otra e incluso de un aula a otra dentro de la misma escuela y de un alumno a otro dentro de una misma aula. En palabras de uno de los alumnos entrevistados por el Equipo de Revisión de la OCDE: “El significado de un 10 varía de un maestro a otro y depende de que demuestres tu empeño”.

Las relaciones entre las calificaciones asignadas por los maestros y los resultados de ENLACE no han sido analizadas. Los estudios de este tipo ofrecerían una imagen de la correspondencia entre las calificaciones en el aula y los resultados de las evaluaciones externas de los alumnos, así como una idea de la variabilidad en el significado de las calificaciones para los docentes.

### ***Existe una capacidad limitada a nivel estatal y local para apoyar la evaluación dentro del aula***

A nivel central se lleva a cabo un impresionante esfuerzo para promover la RIEB, incluyendo la formación continua de los maestros, un nuevo enfoque para asignar las calificaciones, herramientas para la evaluación en el aula, y los ajustes a la evaluación externa de los alumnos y al contenido de los libros de texto. Sin embargo, debido a que México es un país grande y diverso y las competencias docentes para la evaluación formativa y la asignación de calificaciones son limitadas, es necesario que las autoridades estatales y locales apoyen las estrategias. No va a ser suficiente con entregar los materiales a las escuelas de todo el país; las autoridades educativas estatales, jefes de sector, supervisores y jefes de enseñanza también deben responsabilizarse.

Sin embargo, los estados parecen tener una capacidad limitada para ayudar a los maestros en cuestiones pedagógicas y de evaluación. El trabajo de los supervisores se centra principalmente en verificar el cumplimiento de los reglamentos escolares. Hay una clara necesidad de cambiar el perfil de los supervisores y ATP hacia una función más especializada en temas pedagógicos y de evaluación (consulte también el capítulo 5). A nivel estatal, en lugar de duplicar las funciones del INEE, se necesita desarrollar organismos que apoyen directamente a las escuelas, con una interacción en torno a cuestiones pedagógicas, didácticas y de evaluación. Esto también podría incluir estrategias para que las escuelas y los maestros colaboren e intercambien experiencias (consulte el capítulo 5). Al mismo tiempo, la SEP debe considerar estrategias para mejorar las capacidades a nivel estatal para proporcionar un apoyo mucho más significativo a las escuelas en los aspectos más prometedores para mejorar el aprendizaje del estudiante.

### ***Lograr que la evaluación de los alumnos sea inclusiva sigue siendo un desafío***

Las escuelas de educación básica en México tienen una proporción importante de alumnos con necesidades especiales en las clases regulares, a los cuales se les proporciona poco apoyo extra. Tanto ENLACE como EXCALE no cuentan con una estrategia para incluir a esta población. La implementación de las pruebas y la aplicación de las variaciones basadas en servicios educativos o en adaptaciones para los alumnos con necesidades educativas especiales no se han puesto en marcha, excepto en algunos casos particulares de ENLACE (SEP e INEE, de próxima aparición) (consulte también el capítulo 6 sobre EXCALE).

Además, existe una preocupación importante entre los maestros y otras partes interesadas sobre el sesgo cultural de la evaluación externa de los alumnos. El Equipo de Revisión de la OCDE escuchó varios recuentos sobre el uso de preguntas que incluyen situaciones o condiciones que son totalmente desconocidas para los alumnos que no viven en entornos urbanos o que tienen antecedentes indígenas. Como se explicó antes, este problema fue reconocido formalmente a través de la resolución del Consejo Nacional para Prevenir la Discriminación (CONAPRED) al señalar que ENLACE está culturalmente sesgada en contra de los alumnos indígenas y que la SEP necesita trabajar más para lograr que sea una evaluación más justa para esos alumnos.

El enfoque de la RIEB en cuanto a la diversidad es adecuado, ya que establece los estándares comunes y los resultados esperados de aprendizaje, pero, al mismo tiempo, da lugar a adaptaciones locales de acuerdo con los antecedentes, necesidades y contexto de los alumnos. Sin embargo, existen riesgos en la forma en que los docentes pueden llegar a interpretar y aplicar las adaptaciones. Según lo observado por el Equipo de Revisión de la OCDE durante las visitas a las escuelas, muchos maestros que trabajan con los alumnos en desventaja no entienden la necesidad de adaptar el plan de estudios en un sentido pedagógico. En vez de esto, sus adaptaciones consisten en bajar las expectativas de logro y reducir las dificultades

cognitivas para esos alumnos (por ejemplo, pidiéndoles que lean textos más cortos o utilizando números menos complicados en sus cálculos matemáticos). Otros maestros interpretan la necesidad de hacer adaptaciones como ajustar sus criterios de calificación, dando a los alumnos desfavorecidos una calificación mejor que la que se obtendría con el conjunto de criterios establecidos para toda la clase. Por tanto, existe un riesgo importante en torno a la interpretación y aplicación de las adaptaciones, lo que significa menores expectativas para los alumnos desfavorecidos.

## Recomendaciones de política educativa

### *Asegurar una estrategia integral y coherente para la implementación de la RIEB*

En el presente contexto, la mejora de la evaluación dentro del aula necesita desarrollarse junto con la implementación de la RIEB, debido a la inclusión de las iniciativas pertinentes relacionadas con la evaluación formativa y sumativa. Además, la implementación y el impacto de la RIEB dependen fundamentalmente de la introducción exitosa de los cambios en las prácticas de evaluación de los alumnos y en la alineación de éstas con los resultados de aprendizaje esperados y los estándares definidos en el nuevo plan de estudios.

La implementación de un nuevo plan de estudios es una tarea a largo plazo que debe ser diseñada de una manera cuidadosa e inteligente. Durante las últimas dos décadas, América Latina ha tenido varias experiencias con reformas curriculares ambiciosas que tienen poco impacto en las aulas. Si se desea garantizar que la reforma llegue al aula no puede dejarse sólo en manos de las autoridades centrales, especialmente en un país tan grande y diverso como México.

Se debe prestar atención especial en garantizar que el alcance del plan de estudios y las metas de aprendizaje establecidas se conserven en la evaluación de los alumnos, asegurándose que todas las asignaturas y objetivos sean atendidos. Esto implica no sólo la evaluación dentro del aula, sino también las evaluaciones externas (consulte más adelante). En cuanto a la evaluación dentro del aula, los docentes deben integrar en sus prácticas una gama mucho más amplia de actividades e instrumentos para promover y asimilar procesos cognitivos más complejos.

### *Consolidar el dominio del aprendizaje y la evaluación formativa de los maestros*

La implementación exitosa de la RIEB requiere la introducción de una visión más precisa de los procesos de aprendizaje y las cuestiones didácticas en la cultura docente de México. Esta es una condición necesaria para poder implantar las nuevas prácticas de evaluación formativa. Los maestros deben recibir apoyo y capacitación para pasar de una visión tradicional de la enseñanza, concebida como la exposición de temas y conceptos, a un concepto más amplio basado en la facilitación del aprendizaje y el desarrollo de competencias. En este contexto, el repertorio de enfoques sobre el aprendizaje y la evaluación necesita ser ampliado, y alejarse de la asignación de ejercicios y de la práctica de pruebas.

La evaluación formativa es intrínseca a las buenas prácticas docentes. Cualquier esfuerzo para mejorar la enseñanza y el aprendizaje en México debe implicar la mejora de las competencias de los maestros en evaluación formativa de una manera reflexiva y coherente. En la actualidad la mayoría de los docentes utilizan tres tipos de procesos en lo que ellos llaman la evaluación

formativa: (i) indicar a los alumnos sus errores en una prueba o tarea; (ii) pedirles que se esfuercen más; y (iii) alabarlos para motivarlos (Picaroni, 2009; Ravela, 2009a).

Pero la auténtica retroalimentación formativa implica otras cosas. Como lo expresa Wiggins (1998):

*Salvaguardar la premisa fundamental de que la evaluación debe mejorar el desempeño, no sólo auditarlo, requiere que la evaluación incorpore y exija autoajuste basado en una buena retroalimentación... el momento en que el alumno entiende por qué una parte de su trabajo está incorrecta es completamente distinto al momento cuando percibe que al maestro no le gusta esa parte de su trabajo. La mejor retroalimentación es la que especifica, revela o describe ampliamente el resultado real, es clara para el autor y está disponible o se ofrece en términos de objetivos y estándares específicos.*

La evaluación formativa también incluye el desarrollo de instrumentos, como las rúbricas, para que los alumnos reflexionen por sí mismos acerca de la brecha entre lo que se espera que logren y su desempeño real (Wiggins, 1998; Ravela, 2009a).

Es importante reconocer que el repertorio completo de prácticas e instrumentos para evaluación formativa no es sólo una cuestión de enviar nuevos materiales a las escuelas o subirlos a un sitio web (y esperar que los maestros los utilicen). Para asegurar una implementación exitosa, es fundamental crear redes de docentes y desarrollar una interacción profesional sostenida en torno a los nuevos procesos de evaluación (Ravela, 2010). El diseño y la introducción de instrumentos y prácticas adecuadas para la evaluación formativa requiere de un trabajo colaborativo, continuo e interactivo.

### ***Desarrollar un nuevo enfoque para asignar calificaciones***

Asumir la responsabilidad de la certificación del logro de los alumnos es inherente al profesionalismo de los maestros. Los padres de familia y la sociedad confían en el criterio exhaustivo y preciso sobre el desempeño de los alumnos. Las evaluaciones externas pueden contribuir a esta función, pero no pueden reemplazar el juicio profesional de los maestros.

En México, los criterios de calificación son definidos por cada maestro y no están documentados. Por consiguiente, el significado de las calificaciones es específico de cada escuela, salón de clases y alumno. Si se desea que la asignación de calificaciones se alinee con los resultados de aprendizaje esperados y los estándares de la RIEB de manera consistente en todo el país, entonces una de las prioridades debe ser establecer los mecanismos de ponderación para asignar calificaciones, tanto dentro como entre las escuelas. El objetivo es reducir las variaciones en la evaluación docente de los alumnos y la asignación de calificaciones de manera que se mejore la equidad de la evaluación. Las estrategias de ponderación deben incluir una interacción frecuente entre los docentes en torno al significado de las calificaciones (dentro y entre las escuelas), centrada en la relación entre las calificaciones y los niveles de desempeño para cada bloque del contenido curricular, así como en el tipo de evidencia adecuada para cada uno de los niveles de desempeño. La ponderación de las calificaciones también puede incluir el análisis estadístico de las correlaciones entre las calificaciones basadas en el maestro y los resultados de los alumnos en las pruebas externas como ENLACE, IDANIS o EXANI I. Sin embargo, la ponderación no debe entenderse como una manera de estandarizar las calificaciones. En la asignación de calificaciones debe haber espacio para la flexibilidad y las decisiones locales. Lo importante es que el significado de las relaciones entre los resultados de aprendizaje esperados y las calificaciones sea claro y compartido.

La ponderación tiene el potencial para ser una oportunidad profesional de aprendizaje muy eficaz para los maestros, ya que pueden relacionarla estrechamente con sus prácticas en el aula. La ponderación también contribuye a mejorar los juicios profesionales sobre el trabajo de los alumnos y a compartir estándares o criterios de calificación comunes. En los programas de desarrollo profesional en lectoescritura, aritmética y evaluación para el aprendizaje de Nueva Zelanda, se ha demostrado que las discusiones profesionales relacionadas con el trabajo de los alumnos tienen un impacto muy positivo en los resultados de su aprendizaje (Timperley *et al.*, 2007).

El desarrollo de los procesos de ponderación debe ir acompañado de la elaboración de lineamientos nacionales de evaluación teniendo como referencia los objetivos de aprendizaje. Los maestros necesitan utilizar ejemplos de trabajos de los alumnos para ilustrar los logros en los diferentes niveles, calificaciones, indicadores del desempeño deseado de los alumnos, tareas opcionales de evaluación y exámenes.

Otros aspectos relevantes que deben abordarse son:

- Los procedimientos para llegar a una decisión sobre la calificación de un alumno en particular (que no sea un promedio de puntos);
- Alejarse de los enfoques normativos para calificar (es decir, la comparación de los alumnos dentro de una clase) y comprender el significado y la importancia de la evaluación basada en criterios (es decir, dar una calificación según los estándares establecidos para los diferentes niveles de desempeño); y
- Reconocer la importancia de reportar por separado el logro de los alumnos y sus actitudes y compromiso.

Los distintos aspectos del desempeño de los alumnos deben prever informes de evaluación independientes, para que quede en claro lo que es necesario mejorar. Por ejemplo, es importante distinguir entre la capacidad de entender lo que se pide en un problema de matemáticas y la capacidad de hacer cálculos o las actitudes en el aula.

El trabajar con los problemas y retos reales que enfrentan los maestros al calificar a los alumnos puede ser una estrategia de capacitación de gran alcance para desarrollar las capacidades de los docentes. Talleres en los que los maestros analicen y califiquen, de manera independiente, muestras del trabajo de los alumnos, para después comparar y discutir las calificaciones asignadas por cada uno, es un buen ejemplo del tipo de capacitación necesaria. Como se mencionará más adelante, la implementación de estas estrategias requiere de un gran número de especialistas en evaluación a nivel local, visitas continuas a las escuelas y el trabajo con directores y maestros.

La capacidad de los maestros para calificar con base en los objetivos nacionales de aprendizaje de los alumnos también debe evaluarse en el contexto de la evaluación de las escuelas. Esto podría incluir la comparación de las calificaciones del maestro con los resultados de ENLACE, revisión de los instrumentos y los criterios utilizados para calificar, evaluar hasta qué punto las calificaciones se relacionan con los niveles de desempeño en los resultados de aprendizaje esperados y examinar si las calificaciones son comunicadas claramente a los alumnos y padres de familia.

### ***Desarrollar una buena estrategia para fortalecer las capacidades de los maestros para la evaluación de los alumnos***

El desarrollo de habilidades y competencias docentes para la evaluación de los alumnos requiere de una gran inversión y una planeación racional. Como lo expuso Crozier (1989), “la inversión en recursos humanos es la más difícil de lograr. Pero si tiene éxito, es la más eficaz de todas”.



El desarrollo profesional continuo debe ser concebido como algo más que cursos para los maestros en servicio. También debe implicar una interacción entre los docentes en sus propias escuelas y con todas las escuelas locales, y un enfoque puntual en las prácticas de enseñanza. Según TALIS, participar en “cursos y talleres” es el tipo más común de actividades de desarrollo profesional llevada a cabo por los maestros en México (consulte el anexo D), mientras que la participación en redes es la actividad menos común (consulte el anexo D).

Del mismo modo, mejorar las habilidades de evaluación de los alumnos durante la formación inicial docente requiere algo más que impartir cursos sobre evaluación. La estrategia fundamental para mejorar las habilidades de evaluación de los futuros maestros debería ser la implementación de un nuevo enfoque de la evaluación en las instituciones de formación docente. Durante sus estudios, los futuros maestros deberían ser evaluados, con fines formativos y sumativos, en la misma forma en que se espera evalúen a sus alumnos con: uso de rúbricas, tareas complejas, criterios de calificación basados en los niveles de desempeño, explícitos resultados de aprendizaje esperados y así sucesivamente. Esto requiere de una inversión considerable en programas de formación de profesores sobre la forma de dar retroalimentación y asignar calificaciones con base en criterios.

Lo mismo es válido para los esfuerzos actuales de la SEP en apoyar a los maestros a través de la publicación de documentos y artículos acerca de las prácticas de evaluación dentro del aula y las prácticas sugeridas a través de la página web de la SEP. Si bien estas iniciativas son importantes, no son suficientes para promover un cambio real. Cuando este tipo de estrategia se aplica de manera aislada equivale a “disparar al cielo y esperar que caiga un pato” (Ravela, 2010).

Además, con el fin de mejorar las prácticas docentes, es necesario pasar del concepto de la enseñanza como actividad aislada hacia una visión como actividad profesional, lo cual implica interactuar con colegas y estar abiertos a su revisión (Ravela, 2011). Se debe pedir a los maestros que reflexionen sobre su propia práctica y aprendan de la experiencia. En el núcleo de una estrategia de desarrollo docente debe haber un espacio para que los docentes experimenten, compartan y reflexionen sobre sus prácticas en el aula (Shepard, 2006; Ravela, 2009b).

Se necesita hacer un esfuerzo importante para capacitar a los maestros en el desarrollo de las rúbricas de evaluación y otros tipos de instrumentos cualitativos para evaluar el trabajo diario de los alumnos de una manera más significativa, así como en los enfoques para calificar las unidades de aprendizaje. Un aspecto importante es que no es suficiente producir los instrumentos a nivel central y enviarlos a las escuelas. Los maestros mismos deben participar en la elaboración de sus propios instrumentos, en el nuevo enfoque de la evaluación.

### ***Rediseñar y fortalecer el papel de los supervisores***

El papel de las entidades federativas y su sistema de supervisión es crucial para lograr el cambio en el aula, dada su proximidad a las escuelas. Los estados son responsables de los servicios de educación dentro de sus fronteras y los supervisores son el vínculo principal entre las escuelas, las autoridades y las políticas educativas. El éxito de cualquier reforma nacional depende primordialmente de la capacidad de los estados y de la supervisión para garantizar los vínculos necesarios en las aulas. En este contexto, es necesario invertir de manera substancial en la capacidad de los supervisores, ATP, jefes de enseñanza y jefes de sector para que puedan contribuir en la implementación de las reformas. Un área de atención especial debe ser el liderazgo educativo y pedagógico, el cual incluye estrategias fundamentadas para observaciones en el aula.

Las visitas a las escuelas por el Equipo de Revisión de la OCDE pusieron de manifiesto dos características principales del trabajo de los supervisores, a través de las entrevistas

con los actores. Primera, los supervisores son esenciales para implementar el cambio, dada su autoridad y cercanía con las escuelas. Segunda, los supervisores no son necesariamente confiables para efectuar el cambio pedagógico y educativo dado su enfoque en el control político y administrativo de las escuelas. El Equipo de Revisión de la OCDE tuvo la impresión de que las principales preocupaciones de los supervisores parecen estar relacionadas con la recopilación de información y el control de las actividades en las escuelas. Sus intervenciones pedagógicas son rudimentarias (consulte también el capítulo 5).

Al mismo tiempo, en un sistema que carece de una tradición de trabajo docente colectivo en torno a cuestiones pedagógicas y prácticas de enseñanza, existe una enorme necesidad de que algunos agentes específicos faciliten y orienten la interacción de los docentes (Ravela, 2009b). Estos agentes deben reunir a los maestros a nivel escolar o local, promover y dirigir el intercambio entre docentes, construir un marco común de evaluación, crear espacios de colaboración, observar con frecuencia las prácticas en el aula y dar retroalimentación externa a los maestros sobre sus prácticas. Teniendo en cuenta la realidad actual, los actores de la estructura de supervisión (inspectores, ATP, jefes de enseñanza, jefes de sector) parecen estar en una posición ideal para convertirse en agentes de este tipo. Esto implicaría la redefinición de su papel de manera que se enfoque mucho más en el liderazgo educativo y pedagógico (consulte también el capítulo 5). Esto también podría incluir la creación de nuevos puestos y la contratación de nuevo personal con la formación adecuada.

Algunas entidades, como Aguascalientes, ya están desarrollando experiencias en este ámbito y los supervisores se están responsabilizando de la creación de redes de escuelas y la construcción de alineación en todos los diferentes niveles. Otros diez estados están adoptando prácticas similares (OCDE, 2011a). Dichas prácticas responden a los actuales esfuerzos por fortalecer el apoyo a las escuelas: “la reorganización del sistema educativo en México se basa en gran parte en la creación de Regiones para la Gestión de la Educación Básica (RGEB), que son unidades geográficas definidas en torno a la escuela para apoyar los diversos aspectos relacionados con los servicios educativos, tales como la planeación, la ejecución de programas, la distribución de recursos, la recopilación de datos, la distribución de materiales, la evaluación y la rendición de cuentas. Cada región (RGEB) tendrá un Centro de Desarrollo Educativo (CEDE), encargado de asegurarse de que las condiciones locales administrativas y académicas sean apropiadas para apoyar la mejora en el desempeño escolar y los resultados de aprendizaje de los alumnos” (OCDE, 2011a). Es esencial que los CEDE tengan un claro énfasis en la promoción de redes de escuelas y docentes en torno a los enfoques pedagógicos y las prácticas e instrumentos de evaluación. Es importante tener en cuenta que existe un riesgo significativo de que la cultura actual absorba el potencial de innovación de los CEDE y éstos se conviertan en un nuevo centro administrativo para el control de las escuelas.

### ***Promover el uso formativo de las evaluaciones estandarizadas de los alumnos***

Una prioridad de la política educativa debería ser promover el uso formativo adecuado de las evaluaciones estandarizadas como ENLACE y PISA, así como alejarse de los incentivos otorgados a las escuelas para practicar las pruebas. Por ejemplo, se ha distribuido ampliamente un documento titulado “Sugerencias para el uso pedagógico de los resultados de ENLACE”, cuyo objetivo principal es utilizar los resultados de ENLACE como una herramienta pedagógica y técnica para que los maestros mejoren sus prácticas de enseñanza y para mejorar la calidad del aprendizaje en el aula (SEP 2011g; SEP e INEE, de próxima aparición).<sup>8</sup> Sin embargo, el alcance del uso de este documento por parte de los docentes es poco claro para el Equipo de Revisión

de la OCDE. Cuando se les preguntó sobre la forma en que trataban de mejorar los resultados de ENLACE, la respuesta más frecuente fue “mediante la práctica de las pruebas”. Además, el hecho de que los resultados de ENLACE se pongan a disposición de los alumnos, padres de familia y maestros hasta el año escolar posterior a la aplicación de ENLACE no facilita el uso formativo de los resultados (Mendoza Trejo, 2010). En estas circunstancias, los resultados de ENLACE son menos pertinentes para fundamentar las estrategias de mejora del aprendizaje. Por tanto, se debe reflexionar sobre cómo mejorar el tiempo de entrega de los resultados de manera que puedan apoyar las estrategias de aprendizaje en el mismo año escolar en que se aplica ENLACE.

En el caso de PISA, las autoridades deberían enfocar la atención de los docentes en la comprensión de su marco — lo que evalúa PISA — y en la reflexión y discusión sobre cómo desarrollar las competencias evaluadas, en el aula. Los usos de la evaluación externa para mejorar la enseñanza no deben centrarse en los reactivos de la prueba, sino en el marco de evaluación conceptual. Lo importante es que los maestros entiendan lo que se está evaluando y el por qué, de manera que reflexionen sobre los procedimientos para mejorar sus prácticas de enseñanza (Ravela, 2010, 2011).

### ***Desarrollar un marco de trabajo más articulado y coherente para la evaluación externa***

Mientras que la RIEB incluye un marco sólido para la evaluación dentro del aula, las evaluaciones externas son muy diversas. IDANIS, EXANI I, ENLACE y EXCALE no están claramente articuladas dentro de una estrategia de evaluación externa. Cada una de ellas emerge en un período diferente, en respuesta a las distintas necesidades históricas, pero no ha habido un esfuerzo por redefinir y articular claramente su papel en el marco de evaluación de los alumnos.

Se recomienda que se aborden los siguientes puntos:

- Los efectos de ENLACE deben ser revisados y comunicarse con claridad. ENLACE se utiliza actualmente para una gran variedad de propósitos, incluyendo el monitoreo del sistema nacional y estatal, una tarea que logra más apropiadamente EXCALE.
- Todas las evaluaciones externas deben ser rediseñadas y alineadas con los estándares de la RIEB y los resultados de aprendizaje esperados, y estar orientadas a las competencias. En el caso de ENLACE, hay algunos avances en este sentido y en 2012-2013 se introducirá una nueva generación de pruebas ENLACE, después de su piloto en 2011-2012. En el caso de IDANIS y EXANI I, su enfoque en “habilidades” que “pronostican” el desempeño futuro parece ser un poco anticuado. Si se desea mantener estos exámenes para fines de selección —que ya, de por sí, es un problema que debería considerarse— su contenido también debe alinearse con la RIEB.
- EXCALE tendrá que desarrollar un diseño más complejo para mantener tanto las tendencias de desempeño iniciadas en 2005 con la prueba alineada al plan de estudios antiguo como la evaluación de los logros de los alumnos en las nuevas competencias fomentadas por la RIEB (consulte también el capítulo 6). Cabe señalar que este trabajo comenzó en 2012 con el uso de submuestras en el tercer grado de educación secundaria.
- Se debe emprender un esfuerzo importante para introducir tareas más diversas en las evaluaciones externas, no sólo preguntas de opción múltiple. Lo anterior se aplica no sólo a las evaluaciones externas de los alumnos, antes mencionadas, sino también a las evaluaciones administradas para la evaluación de los maestros (consulte el capítulo 4 para más detalles). Si se desea que los docentes adopten un nuevo enfoque global de evaluación, se debe reducir la constante presencia de preguntas de opción múltiple. En las pruebas de evaluación docente, los maestros también deben tener la experiencia de ser evaluados

con instrumentos y tareas que rebasen las preguntas de opción múltiple, de modo que ellos las repitan en la evaluación de sus alumnos.

Si bien la introducción de ítems de construcción de respuestas u otro tipo de tareas complejas en las evaluaciones a gran escala es agotadora, en la actualidad la tecnología hace que las cosas sean más fáciles. En Chile, por ejemplo, se ha desarrollado una importante capacidad en torno a la codificación electrónica de las preguntas abiertas y las respuestas escritas de los alumnos en la prueba nacional estandarizada (Sistema de Evaluación de la Calidad de la Educación, SIMCE), así como en la codificación de los portafolios en el programa de evaluación docente (Santiago *et al.*, de próxima aparición; Manzi *et al.*, 2011). Aunque México ya ha desarrollado una capacidad significativa en el diseño, administración y procesamiento de las pruebas de opción múltiple, el siguiente reto debe ser desarrollar la capacidad de introducir ítems de construcción de respuestas y tareas más complejas en las evaluaciones a gran escala. Hacer este importante cambio a ítems de construcción de respuesta y tareas complejas en las evaluaciones externas, junto con el esfuerzo para alinearlos con el enfoque de competencias de la RIEB, deben ser prioridades de corto plazo.

### ***Desarrollar estrategias para hacer frente a los efectos no deseados de ENLACE***

ENLACE es un elemento dominante del sistema educativo mexicano. Ha traído beneficios considerables para el aprendizaje de los alumnos, pero también ha generado importantes efectos no deseados. Como resultado, una de las principales prioridades de la política educativa debe ser el desarrollo de estrategias para eliminar, o al menos reducir, los efectos negativos de la prueba actual. Este esfuerzo debe basarse en un estudio a profundidad del impacto de ENLACE en las prácticas de las escuelas y las aulas.

Una estrategia podría ser reducir el alto impacto de ENLACE. Para lograr esto se puede recurrir a varias opciones que dependen de hasta qué punto pueda reducirse el impacto para los actores. Una posibilidad es volver a plantear los objetivos de ENLACE, incluyendo regresar a su propósito original de ser una herramienta meramente diagnóstica y formativa para la evaluación de los alumnos. Otra posibilidad es agregar a este objetivo original algún papel en el sistema de evaluación para determinar si, a nivel nacional, se logran o no los objetivos de aprendizaje en las asignaturas cubiertas por ENLACE. La mayoría de los países de la OCDE limitan el uso de las evaluaciones estandarizadas a estas dos funciones. Si se mantiene el objetivo de utilizar ENLACE para la rendición de cuentas de las escuelas (publicación de los resultados de la prueba en la escuela) y para la evaluación de los maestros, entonces es imperativo desarrollar técnicas de valor agregado para capturar el impacto real de cada escuela (para la publicación de los resultados y su uso en RNAME) (consulte también el capítulo 5) y reducir considerablemente el peso de los resultados de ENLACE en la evaluación de los maestros por las razones que se explican en este capítulo y en el capítulo 4. Al mismo tiempo, es importante vigilar los posibles efectos no deseados de los usos de alto impacto de ENLACE mediante estudios de investigación apropiados.

Otra estrategia es transformar ENLACE en una herramienta para la evaluación sumativa externa de los alumnos, es decir, un sistema de exámenes externos. Esto implicaría ampliar los objetivos de aprendizaje evaluados mediante la inclusión de más asignaturas y la ampliación de la gama de tareas asignadas. También tendría consecuencias para los alumnos ya sea en la contribución de sus calificaciones finales o como un mecanismo de certificación al final de las etapas clave de la educación (al final de los ciclos educativos). Esto supondría una fuerte

motivación para los alumnos (Shepard, 2006; Messick, 1999) y para los maestros, porque la mayoría de los docentes están realmente preocupados por el éxito de sus alumnos. Como se señala en un informe reciente de la OCDE sobre las lecciones de los países con un alto desempeño en PISA, las evaluaciones de alto impacto para los alumnos al final de ciertos niveles del sistema educativo constituyen fuertes incentivos para los alumnos y maestros. Esto refuerza lo que se conoce como rendición de cuentas horizontal —cuando los docentes tienen que responder a sus colegas y a los padres de familia— en lugar de reforzar la rendición de cuentas vertical —cuando los maestros tienen que responder a las autoridades e instancias administrativas (OCDE, 2011c). La rendición de cuentas horizontal es más difícil de simular que la vertical, debido a las relaciones interpersonales cotidianas.

Un camino para reducir la carga de las escuelas, maestros y alumnos, así como los gastos de administración, sería aplicar la evaluación sólo en las etapas clave de la educación, por ejemplo, al final de cada ciclo educativo (es decir, tercero y sexto de primaria, y tercero de secundaria), en lugar de hacerlo en cada grado escolar como sucede actualmente. Otra posibilidad es dejar al menos una parte de la calificación a los maestros junto con procedimientos de ponderación sofisticados. Esto podría implicar el intercambio de docentes entre escuelas, rúbricas y manuales de calificación diseñados centralmente, el control central de la administración de la evaluación y de la calificación de una muestra de escuelas. La experiencia de Suecia (Nusche *et al.*, 2011) con sus evaluaciones externas y de Nueva Zelanda (Nusche *et al.*, 2012) con su sistema de certificación en la educación media superior, pueden ser particularmente útiles.

Este enfoque podría tener varias ventajas, por ejemplo:

- Confiar más en el profesionalismo de los docentes y promoverlo;
- Contar con una valiosa experiencia de aprendizaje profesional para los maestros;
- Involucrar a los docentes en la evaluación nacional y fomentar su conocimiento de la RIEB;
- Promover el esfuerzo de los alumnos;
- Hacer posible el uso de una gama más amplia de tareas y preguntas de evaluación; y
- Permitir la evaluación de un conjunto más amplio de materias y de competencias dentro de las asignaturas.

### ***Asegurar que la evaluación de los alumnos sea inclusiva***

Los sistemas de evaluación deben subrayar la importancia de responder a las necesidades individuales de los alumnos y al contexto de la comunidad escolar y diseñar estrategias de evaluación que se adapten a las necesidades de los diferentes grupos de alumnos. El objetivo es desarrollar un sistema de evaluación incluyente basado en el principio de que todos los alumnos deben tener la oportunidad de participar en las actividades educativas —entre ellas las actividades de evaluación— y demostrar sus conocimientos, habilidades y competencias de una manera justa. Por lo tanto, las prácticas de evaluación de los maestros y el formato y contenido de las pruebas externas estandarizadas (tales como ENLACE y EXCALE) deben ser sensibles a determinados grupos de alumnos, a saber, los alumnos indígenas, los alumnos con necesidades especiales y los alumnos que viven en contextos sociales desfavorecidos. En el contexto de México, esta es sin duda una tarea formidable dada la existencia de 68 lenguas indígenas, por nombrar un caso.

El contexto cultural debe tenerse muy en cuenta en el diseño de pruebas a fin de evitar el uso de palabras, expresiones y situaciones que sean desconocidas en ciertos contextos culturales. Se sugiere preparar lineamientos de aseguramiento de calidad y adoptar prácticas que garanticen que las evaluaciones externas sean revisadas para evitar un posible sesgo

en estos aspectos. Lo anterior podría incluir la consideración de varios formatos de evaluación (basados en pruebas, tareas de desempeño, orales, escritos) para que cada alumno/grupo de alumnos no esté sistemáticamente en desventaja, así como la revisión del contenido de las preguntas del examen. Además, es necesario prever adaptaciones específicas para que los alumnos con necesidades especiales puedan tomar las pruebas.

Por último, con respecto a la evaluación basada en el maestro, existe un dilema en torno a los criterios de calificación, las adaptaciones locales de los planes de estudio y los resultados de aprendizaje esperados. Por un lado, todos los niños deben alcanzar los resultados de aprendizaje esperados que se prevén para cada grado y materia. Esto significa que los maestros tienen las mismas expectativas académicas para sus alumnos, independientemente de su situación socioeconómica. Por otro lado, las grandes diferencias culturales y socioeconómicas son parte de la compleja y desigual sociedad mexicana. En consecuencia es necesario trabajar estrechamente con los docentes para alcanzar el equilibrio correcto entre la no exclusión de los alumnos (como resultado de expectativas demasiado exigentes) y el no bajar las expectativas de su aprendizaje.

### ***Mejorar la información para los alumnos y padres de familia***

La introducción de la nueva y encomiable Cartilla de Educación Básica debe ir acompañada de algunos ajustes a fin de que se convierta en un auténtico instrumento de aprendizaje. Un primer ajuste sería plantear definiciones asociadas con los niveles de desempeño más específicos. En la actualidad, dichas definiciones son muy generales. Por ejemplo, para “C” o 6-7 (en la antigua escala de calificación), la definición es “muestra un desempeño suficiente en los resultados de aprendizaje esperados para este bloque”, sin importar el grado escolar, la materia o el bloque de aprendizaje. Sería conveniente que esta definición de carácter general fuera específica para cada grado, materia y bloque, y se asociara a los resultados concretos de aprendizaje que los alumnos deben alcanzar.

Los resultados de aprendizaje esperados para cada grado, materia y bloque deben explicarse claramente a los alumnos y padres de familia, e incluir ejemplos de lo que es un desempeño aceptable, satisfactorio y excepcional. Se debe desarrollar una versión de los resultados de aprendizaje esperados en cada bloque de aprendizaje para los padres de familia, de manera que puedan entender el significado de “muestra un desempeño suficiente en los resultados de aprendizaje esperados para este bloque”, por citar un ejemplo. Los mapas curriculares de avance de Chile pueden ser una fuente de inspiración interesante para este esfuerzo (Santiago *et al.*, de próxima aparición).

Otra cuestión clave es la necesidad de información más detallada del desempeño de cada alumno (García *et al.*, 2011, p. 90). El espacio dedicado a ello en la nueva Cartilla de Educación Básica podría no ser útil si los maestros no están preparados para dar indicaciones más precisas sobre cómo mejorar el desempeño a los alumnos y padres de familia. El riesgo es que este instrumento puede quedar limitado a la exhortación de los docentes para que los alumnos realicen un mayor esfuerzo. Las experiencias de Dinamarca y Suecia con los planes individuales, que contienen recomendaciones detalladas y específicas para cada alumno son ejemplos valiosos de enfoques con mayor potencial para generar un incremento en el logro educativo (Shewbridge *et al.*, 2011; Nusche *et al.*, 2011).

## Notas

- 1 El Catálogo Nacional de Formación Continua y Superación Profesional de Maestros de Educación Básica en Servicio 2011-2012 se encuentra disponible en <http://formacioncontinua.sep.gob.mx>.
- 2 La información sobre el proceso de asignación de alumnos a las escuelas en el Distrito Federal se encuentra disponible en [www.sepdf.gob.mx/principal/archivos/nota\\_preg\\_frecuentes\\_ingreso\\_secundaria\\_2011.pdf](http://www.sepdf.gob.mx/principal/archivos/nota_preg_frecuentes_ingreso_secundaria_2011.pdf).
- 3 Sin embargo, debe tenerse en cuenta que el análisis en este informe se refiere sólo a ENLACE en los niveles de primaria y secundaria.
- 4 Las Cartillas de Educación Básica están disponibles en [www.boleta.sep.gob.mx](http://www.boleta.sep.gob.mx).
- 5 Las dos publicaciones están disponibles en [www.pisa.sep.gob.mx/descargas.html](http://www.pisa.sep.gob.mx/descargas.html).
- 6 Las Cartillas de Educación Básica están disponibles en [www.boleta.sep.gob.mx/](http://www.boleta.sep.gob.mx/).
- 7 El Estudio Internacional sobre Enseñanza y Aprendizaje de la OCDE se llevó a cabo en 2007-2008, abarcó la educación secundaria y contó con la participación de 23 países (OCDE, 2009). Los resultados derivados de TALIS se basan en los informes de los propios maestros y directores, por lo que representan sus opiniones, percepciones, creencias y recuentos de actividades. Para más información consulte [www.OECD.org/edu/talis](http://www.OECD.org/edu/talis). Los resultados de México en TALIS se presentan en el anexo D.
- 8 Las publicaciones de ENLACE para cada grado están disponibles en [www.enlace.sep.gob.mx/ba/apoyos\\_para\\_el\\_uso\\_pedagogico](http://www.enlace.sep.gob.mx/ba/apoyos_para_el_uso_pedagogico).

## Bibliografía

- Backhoff, E., E. Andrade, A. Sánchez y M. Peon (2008), *Estudio comparativo del aprendizaje en sexto de primaria 2005-2007*, INEE, México, [www.inee.edu.mx/index.php?option=com\\_content&view=article&id=3469&Itemid=1051](http://www.inee.edu.mx/index.php?option=com_content&view=article&id=3469&Itemid=1051).
- CONAPRED (Consejo Nacional para Prevenir la Discriminación) (2011), Resolución por Disposición 1/2011, EXP. CONAPRED/DGAQR/419/08/DR/II/CHIS/R289, [www.cneii.org/images/stories/resolucion-por-disposicion.pdf](http://www.cneii.org/images/stories/resolucion-por-disposicion.pdf).
- Crozier, M. (1989), *Estado modesto, Estado moderno. Estrategia para el cambio*, Fondo de Cultura Económica, México.
- DOF (*Diario Oficial de la Federación*) (1994), Acuerdo número 200 por el que se establecen normas de evaluación del aprendizaje en educación primaria, secundaria y normal. 19 de septiembre 1994, SEP, México.  
[www.sep.gob.mx/work/models/sep1/Resource/3f848841-8476-4596-9835-d71e306221a3/a200.pdf](http://www.sep.gob.mx/work/models/sep1/Resource/3f848841-8476-4596-9835-d71e306221a3/a200.pdf).
- DOF (2009), Acuerdo número 499 por el que se modifica el diverso número 200 por el que se establecen normas de evaluación del aprendizaje en educación primaria, secundaria y normal. Noviembre 2009, SEP, México, [http://dof.gob.mx/nota\\_detalle.php?codigo=5117051&fecha=04/11/2009](http://dof.gob.mx/nota_detalle.php?codigo=5117051&fecha=04/11/2009).
- EPPI (Evidence for Policy and Practice Information and Co-ordinating Centre) (2002), *A Systematic Review of the Impact of Summative Assessment and Tests on Students' Motivation for Learning*, Review conducted by the Assessment and Learning Synthesis Group, Evidence for Policy and Practice Information and Co-ordinating Centre, Londres.
- García, B., J. Mejía y A. Meza (2009), *Propuesta para evaluar y reportar el proceso de desarrollo de competencias de los alumnos de educación básica, mediante una nueva boleta de calificaciones*, SEP, México.
- García, A.M., M.A. Aguilera, M.A. Pérez y G. Muñoz (2011), *Evaluación de los aprendizajes en el aula – Opiniones y prácticas de docentes de primaria en México*, INEE, México.
- Linn, R. (2000), “Assessments and Accountability”, *Educational Researcher*, Vol. 29, No. 2, pp. 4-16.
- Linn, R. y N. Gronlund (2000), *Measurement and Assessment in Teaching*, Prentice Hall, New Jersey.
- Loureiro, G. (2009), *Evaluación en el aula, currículo y evaluaciones externas*, Partnership for Educational Revitalization in the Americas (PREAL), Santiago de Chile, [www.ucu.edu.uy/Default.aspx?tabid=1267](http://www.ucu.edu.uy/Default.aspx?tabid=1267).
- Manzi, J. R. Gonzalez e Y. Sun (eds.) (2011), *La evaluación docente en Chile*, MIDE-UC, Santiago de Chile.
- Mendoza Trejo, F. (2010), *Estudio sobre el impacto de la difusión de los resultados de la prueba ENLACE 2009*, Instituto de Fomento e Investigación Educativa, A.C.
- Messick, S. (1999), “Key Issues”, en F. Ottobre (ed.), *The Role of Measurement and Evaluation in Education Policy*, UNESCO, Educational Studies and Documents No. 69, París.
- Morris, A. (2011), “Student Standardised Testing: Current Practices in OECD Countries and a Literature Review”, *OECD Education Working Papers*, No. 65, OECD, París, [www.oecd.org/edu/workingpapers](http://www.oecd.org/edu/workingpapers).
- Nusche, D., G. Halász, J. Looney, P. Santiago y C. Shewbridge (2011), *OECD Reviews of Evaluation and Assessment in Education: Sweden*, OECD, París, [www.oecd.org/dataoecd/38/42/47169533.pdf](http://www.oecd.org/dataoecd/38/42/47169533.pdf).
- Nusche, D., D. Laveault, J. MacBeath y P. Santiago (2012), *OECD Reviews of Evaluation and Assessment in Education: New Zealand*, OECD, París, [www.oecd.org/edu/evaluationpolicy](http://www.oecd.org/edu/evaluationpolicy).



- OCDE (2009), *Creating Effective Teaching and Learning Environments: First Results from TALIS*, OECD, París, [www.oecd.org/edu/talis](http://www.oecd.org/edu/talis).
- OCDE (2010), PISA 2009 Results: What Makes a School Successful? Resources, Policies and Practices, Volume IV, OECD, París, <http://dx.doi.org/10.1787/9789264091559-en>.
- OCDE (2011a), *Progress with Educational Reform in Basic Education in Mexico: An OECD Perspective*, OECD, París.
- OCDE (2011b), *Establishing a Framework for Evaluation and Teacher Incentives: Considerations for Mexico*, OECD, París.
- OCDE (2011c), *Strong Performers and Successful Reformers in Education. Lessons from PISA for the United States*, OECD, París.
- OEI (Organización de Estados Iberoamericanos) (2010), *Metas educativas 2021 -La educación que queremos para la generación de los bicentenarios*, OEI, Madrid.
- Picaroni, B. (2009), *La Evaluación en las aulas de primaria: Usos formativos, calificaciones y comunicación con los padres*, Partnership for Educational Revitalization in the Americas (PREAL), Santiago de Chile, [www.ucu.edu.uy/Default.aspx?tabid=1267](http://www.ucu.edu.uy/Default.aspx?tabid=1267).
- Ravela, P. (2009a), “Consignas, Devoluciones y Calificaciones: los Problemas de la Evaluación en las Aulas de Educación Primaria en América Latina”, *Páginas de Educación*, Vol. 2, Universidad Católica del Uruguay, Montevideo, pp. 49-89, [www.ucu.edu.uy/Default.aspx?tabid=1267](http://www.ucu.edu.uy/Default.aspx?tabid=1267).
- Ravela, P. (2009b), “La Evaluación del Desempeño Docente para el Desarrollo de las Competencias Profesionales”, en E. Martín y F. Martínez Rizo, *Avances y desafíos en la evaluación educativa*, OEI-Fundación Santillana, Madrid, pp. 113-126.
- Ravela, P. (2010), *¿Qué pueden aportar las evaluaciones estandarizadas a la evaluación en el aula?*, Documento No. 47, Partnership for Educational Revitalization in the Americas (PREAL), Santiago de Chile, [www.ucu.edu.uy/Default.aspx?tabid=1267](http://www.ucu.edu.uy/Default.aspx?tabid=1267).
- Ravela, P. (2011), *¿Qué hacer con los resultados de PISA en América Latina?*, Documento No. 58, Partnership for Educational Revitalization in the Americas (PREAL), Santiago de Chile, [www.ucu.edu.uy/Default.aspx?tabid=1267](http://www.ucu.edu.uy/Default.aspx?tabid=1267).
- Rosenkvist, M. (2010), “Using Student Test Results for Accountability and Improvement: A Literature Review”, *OECD Education Working Papers*, No. 54, OECD, París, [www.oecd.org/edu/workingpapers](http://www.oecd.org/edu/workingpapers).
- Santiago, P., F. Benavides, C. Danielson, L. Goe y D. Nusche (próximamente), *OECD Reviews of Evaluation and Assessment in Education: Teacher Evaluation in Chile*, OECD, París, [www.oecd.org/edu/evaluationpolicy](http://www.oecd.org/edu/evaluationpolicy).
- Schiefelbein, E. y L. Wolff (1992), *Repetition and Inadequate Achievement in Latin America's Primary Schools: A Review of Magnitudes, Causes, Relationships and Strategies*, UNESCO/OREALC, Santiago de Chile.
- SEP (Secretaría de Educación Pública) (2002), *Licenciatura en Educación primaria. Plan de estudios 1997*. Documentos Básicos, SEP, Subsecretaría de Educación Básica y Educación Normal, Dirección general de Estándares. 4ª Edición. México, México, [www.dgespe.sep.gob.mx/sites/default/files/planes/lepri/plan.pdf](http://www.dgespe.sep.gob.mx/sites/default/files/planes/lepri/plan.pdf).
- SEP (2010), *Licenciatura en Educación secundaria. Plan de estudios 1999*. Documentos básicos, [www.dgespe.sep.gob.mx/sites/default/files/planes/les/plan.pdf](http://www.dgespe.sep.gob.mx/sites/default/files/planes/les/plan.pdf).
- SEP (2011a), *Normas de control escolar relativas a la inscripción, reinscripción, acreditación, regularización y certificación en la educación básica*, SEP, Dirección general de Acreditación, Incorporación y Revalidación. México, [http://dgece.sev.gob.mx/docs/NORMAS\\_BASICA\\_2011-2012.pdf](http://dgece.sev.gob.mx/docs/NORMAS_BASICA_2011-2012.pdf).

- SEP (2011b), Catálogo nacional. *Formación continua y superación profesional para maestros de Educación básica en servicio 2011-2012*, SEP Dirección General de Formación Continua de Maestros en Servicio
- DGFCMS [http://formacioncontinua.sep.gob.mx/index.php?option=com\\_content&view=article&id=234:catalogonaciona20112012&catid=35:novedades&Itemid=110](http://formacioncontinua.sep.gob.mx/index.php?option=com_content&view=article&id=234:catalogonaciona20112012&catid=35:novedades&Itemid=110).
- SEP (2011c), Plan de Estudios 2011. Educación básica, SEP, Mexico.
- SEP (2011d), Acuerdo 592 por el que se Establece la Articulación de la Educación básica, SEP, México.
- SEP (2011e), *Competencias para el México que queremos: Hacia PISA 2012. Manual de alumnos*, SEP, México.
- SEP (2011f), *Competencias para el México que queremos: Hacia PISA 2012. Manual de maestros*, SEP, México.
- SEP (2011g), *Características generales e información de los reactivos aplicados para su uso pedagógico. Sexto Grado de Primaria 2011*, SEP, Mexico.
- SEP (Secretaría de Educación Pública) e INEE (Instituto Nacional para la Evaluación de la Educación) (próximamente), *Country Background Report for Mexico*, prepared for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, to be available from [www.oecd.org/edu/evaluationpolicy](http://www.oecd.org/edu/evaluationpolicy).
- SEP-DF (Secretariat of Public Education in Federal District) (2011), *Preguntas frecuentes ingreso a secundarias 2011-2012*, SEP-DF, México, [www.sepdf.gob.mx/principal/archivos/nota\\_preg\\_frecuentes\\_ingreso\\_secundaria\\_2011.pdf](http://www.sepdf.gob.mx/principal/archivos/nota_preg_frecuentes_ingreso_secundaria_2011.pdf).
- Shepard, L. (2006), “Classroom Assessment”, en R. Brennan (ed.), *Educational Measurement*, ACE/Praeger, Westport, pp. 623-646.
- Shewbridge, C., E. Jang, P. Matthews y P. Santiago (2011), *OECD Reviews of Evaluation and Assessment in Education: Denmark*, OECD, París.
- Stigler, J. y J. Hiebert (2009), *The Teaching Gap*, Free Press, Nueva York.
- Timperley, H., A. Wilson, H. Barrar e I. Fung (2007), *The Teacher Professional Learning and Development Best Evidence Synthesis*, Ministry of Education, New Zealand, [www.educationcounts.govt.nz/publications/series/2515](http://www.educationcounts.govt.nz/publications/series/2515).
- Torres, R. (1995), “Repetition: A Major Obstacle to Education for All”, en *Education News*, No. 15, UNICEF, Nueva York.
- UNESCO/OREALC, 2008, *Los aprendizajes de los estudiantes de América Latina y el Caribe. Primer reporte de los resultados del Segundo Estudio Regional Comparativo y Explicativo*, UNESCO/OREALC, Santiago de Chile.
- Vidal, R. (2009), *¿ENLACE, EXANI, EXCALE o PISA? CENEVAL*, México.
- Wiggins, G. (1998), *Educative Assessment – Designing Assessments to Inform and Improve Student Performance*, Joey-Bass, San Francisco.
- Zúñiga Molina, L. y J. L. Gaviria (2010), *Challenges and Opportunities for the Further Development of the ENLACE Assessment for Evaluation and Teacher Incentives in Mexico*, expert paper commissioned by the OECD for the Co-operation Agreement between the OECD and the government of Mexico.