

Uso de los resultados de evaluaciones educativas a gran escala en América Latina

Patricia Arregui

Grupo de Trabajo sobre Estándares y Evaluación del PREAL

Octubre 2008

Así como durante la etapa de expansión del acceso a la educación en América Latina se dio prioridad al recojo y utilización de indicadores de cobertura y niveles de escolaridad¹, cuando hacia fines de los ochenta lograr la efectividad de las escuelas y el aprendizaje real se convirtió en el principal desafío y se tuvo que diseñar y justificar nuevas estrategias para la asignación de recursos, la medición de la calidad de la educación y de los aprendizajes se hizo impostergable.

Así, a mediados y fines de los años noventa, la mayoría de países de América Latina empezó a dar sus primeros pasos hacia el desarrollo de capacidades técnicas básicas para diseñar e implementar evaluaciones a gran escala y para emprender análisis estadísticos de los datos recogidos².

Hoy en día, tras varios años de experiencias más o menos continuas en este quehacer, en buen número de países se están dando discusiones y acciones orientadas al rediseño de los sistemas de evaluación. Una parte del debate se refiere a necesarias mejoras técnicas en el desarrollo de instrumentos y procedimientos para la recolección y el análisis de los datos, así como al alineamiento de los mismos con los objetivos curriculares vigentes. Otra parte tiene que ver principalmente con la constatación de que los esfuerzos para comunicar y difundir los resultados de las pruebas en América Latina no han alcanzado a muchos agentes clave, incluyendo profesores/as y padres³. Tanto o mayor desasosiego genera el hecho de que, aparentemente, la abundante información que se está produciendo está siendo usada de modo insuficiente para sustentar y apoyar procesos de mejoras tanto en las políticas como en la enseñanza o como instrumento para la responsabilización o "accountability". Las evaluaciones parecerían, en opinión de algunos, no estar teniendo mucho impacto "más allá de informar, sensibilizar y crear conciencia"⁴. Por supuesto, no faltan quienes están más preocupados aún por los usos no válidos o hasta perversos de las evaluaciones para el logro de objetivos para los cuales no fueron diseñadas.

En nuestra región, y hasta el momento, los resultados de las evaluaciones a gran escala se han utilizado principalmente para:

1. elevar la conciencia ciudadana sobre los retos de la educación
2. mejorar el diseño de las políticas educativas y la asignación de recursos
3. diseñar programas de apoyo y seleccionar escuelas o grupos de estudiantes para participar en ellos
4. realizar investigación y evaluar el impacto de programas
5. establecer o revisar estándares educativos
6. promover la responsabilización
7. mejorar la gestión pedagógica e institucional en las escuelas o áreas escolares

En este documento, se pasa revista a estos diversos tipos de usos de los resultados de evaluaciones nacionales, sub-nacionales, regionales y/o internacionales en nuestra región, se

¹ Para ese entonces, las oficinas de estadística recogían información sobre varios insumos clave tales como calificaciones formales de los/as profesores/as, número de estudiantes por profesor/a, disponibilidad del espacio y equipamiento en el aula, provisión de libros de texto y de bibliotecas escolares, etc. También producían información como tasas promedio de escolaridad y tasas de promoción y repetición.

² Para una visión general de la instalación y desarrollo de los sistemas nacionales de evaluación de logros de aprendizaje, ver Ferrer 2006.

³ Ravela 2002.

⁴ Moncada et al. 2005: 279-281.

presenta algunas generalizaciones al respecto y se enuncian algunas advertencias sobre los desafíos que supone una utilización más productiva de los mismos. Los siguientes acápite describe a grandes líneas cada uno de estos tipos de usos, dedicando más atención a los dos últimos.

--oOo--

1. Elevar la conciencia ciudadana sobre los retos de la educación

Si bien en América Latina las evaluaciones del rendimiento educativo se iniciaron generalmente a partir de la necesidad de informar a los tomadores de decisiones y teniendo a ello como su principal propósito, el contexto de resurgimiento democrático de la época y la idea de “rendición de cuentas” que empezaba a instalarse en el discurso político también hacía necesario satisfacer los derechos de los ciudadanos y de algunas audiencias clave de conocer la situación y las tendencias de los resultados educativos. Informar a la opinión pública ha sido quizás por ello el uso más común y efectivo de las pruebas de logros de aprendizaje. Los resultados han provisto de evidencia “dura” sobre los logros insuficientes y desiguales, han estimulado una preocupación pública sobre el problema, y reorientado el foco de la demanda social hacia una distribución más equitativa de oportunidades de aprendizaje.

Aun así, en la mayoría de casos, la prensa recogió sólo los resultados más llamativos sobre el bajo nivel de logro general, y en el caso de las pruebas internacionales, sobre los bajos puntajes obtenidos por los participantes latinoamericanos. A menudo se puso más énfasis en las grandes diferencias entre las escuelas privadas y las públicas, y mucho menos en aquéllas entre las poblaciones urbana y rural o entre niñas y niños⁵. Rara vez se mencionaron las diferentes características socioeconómicas y culturales de las poblaciones atendidas por escuelas privadas y públicas, pues los informes oficiales no incluían esa información. Por otro lado, al menos al inicio, los funcionarios públicos rara vez estaban disponibles para analizar los resultados y promover un diálogo capaz de orientar a la opinión pública a partir de las evaluaciones.

Curiosa, aunque no sorprendentemente, entre las audiencias o grupos de *stakeholders*, son quizás los padres de familia el grupo menos atendido en términos de la provisión de información útil para monitorear el desempeño de las escuelas a las que asisten sus hijos/as o del sistema educacional en su conjunto y para convertirlos en agentes de cambio mejor informados. Aún hoy en día, en muchos de nuestros países los reportes preparados por las agencias de evaluación están más allá de las posibilidades de comprensión de la mayoría de los padres, particularmente si se considera el bajo nivel educativo promedio hasta ahora alcanzado por los adultos latinoamericanos. El manejo de la información por parte de la prensa probablemente no ha ayudado a que la información publicada se interprete de forma precisa, y las capacidades de “traducción” de los sofisticados procedimientos técnicos de evaluación y análisis estadístico de las mismas agencias evaluadoras son aún limitadas.

2. Mejorar el diseño de las políticas educativas y la asignación de recursos

Aun así, gracias a las evaluaciones nacionales e internacionales, se cuenta con evidencia nueva y más sólida sobre el bajo rendimiento de los estudiantes, así como sobre las grandes diferencias en los resultados de varios grupos poblacionales y tipos de escuelas, se ha hecho posible el establecimiento y la generación de consensos sobre nuevas prioridades generales de política.

Por ejemplo, la necesidad de que los ministerios de educación, los administradores y las escuelas se centren en el aprendizaje de los estudiantes, es ahora un mandato público generalizado en la región. Y, si bien la “focalización” y la discriminación positiva o acción

⁵ Se debe resaltar que muchas de las primeras pruebas fueron construidas con un enfoque referido a normas, lo que podría quizás explicar estos énfasis.

afirmativa están hasta ahora lejos de ser el eje principal de las políticas públicas, hay una innegable y creciente conciencia de la necesidad de realizar una distribución más equitativa de las oportunidades de aprendizaje.

En algunos países muchas políticas y planes inspirados, según propias aseveraciones, por los resultados de las evaluaciones no han logrado trascender el papel donde han sido formulados. En otros, en cambio, sí se se ha pasado a la acción con, por ejemplo, reformas curriculares y desarrollo de estándares, producción y distribución masiva de libros de texto y cuadernos de trabajo (que los primeros análisis sugirieron gravitaban fuertemente en el aprendizaje), el fomento de la inclusión de estudiantes pobres en escuelas particulares subvencionadas (por aparentes mejores resultados y el efecto que tienen los pares en los logros estudiantiles), programas de capacitación o desarrollo profesional docente focalizados en el fortalecimiento de capacidades de comprensión lectora y de razonamiento matemático entre los maestros (que algunos análisis de factores asociados a los rendimientos estudiantiles sugieren tienen fuerte impacto sobre los mismos) o, más recientemente, la propuesta de subvenciones diferenciadas que favorezcan a las escuelas y los estudiantes más pobres.

Varios países cuentan con políticas y programas que procuran concentrar más recursos e intervenciones diferenciadas y a gran escala en áreas geográficas y grupos poblacionales con bajo rendimiento y condiciones generales de pobreza, aunque generalmente de modo experimental o demostrativo, más que como política generalizada. Esto incluye desde grandes programas de apoyo integral a zonas y escuelas más vulnerables y de bajos resultados hasta, en algunos casos, una distribución no siempre del todo racional de nuevos insumos “empezando” por los más pobres – como podría ser el caso de reciente distribución de computadoras portátiles en el Perú. Sobre estos programas se describen aspectos adicionales en el siguiente acápite.

Puede argumentarse que esta tendencia al “alineamiento” entre resultados y políticas se ha dado de manera más clara en países donde los resultados de las primeras evaluaciones nacionales han reforzado la validez de algunas decisiones que ya habían sido tomadas, ya sea por recomendación de las agencias internacionales de financiamiento o a partir de un genuino proceso interno de deliberación. En cualquier caso, los resultados de las pruebas proporcionaron fuertes argumentos para validar socialmente esas propuestas de política.

3. Diseñar programas de apoyo y seleccionar escuelas o grupos de estudiantes para participar en ellos

Las evaluaciones también han permitido identificar áreas más específicas donde se experimenta problemas, analizar las posibles causas de éstos y diseñar programas experimentales o generales que pudieran servir para superarlos - así como identificar y seleccionar poblaciones objetivo o instituciones educativas que puedan ser beneficiarias de esos programas.

Típico ejemplo de estos programas se dio en Chile, cuando los resultados de las pruebas SIMCE mostraron grandes brechas de rendimiento entre escuelas privadas, subvencionadas y municipales y se dio inicio como una estrategia de emergencia al "P-900" -el "Programa para 900 Escuelas Ayudando a los Pobres". El 10% de escuelas con el rendimiento más bajo en cada una de las regiones del país y con grandes proporciones de estudiantes pobres urbanos y rurales de nivel primario fue invitado a participar en actividades del programa - inicialmente concentradas en remodelar aulas y abastecer de equipo básico, libros de texto y bibliotecas a las escuelas, pero luego incluyendo talleres para profesores que buscaban promover la autonomía profesional y la responsabilidad, y espacios de tutoría en grupos pequeños para estudiantes con bajo rendimiento que eran dirigidos por miembros de la comunidad. El P-900 se convirtió más adelante en un programa permanente del Ministerio de Educación, centrado en proveer servicios de capacitación de profesores/as *in situ*, y desarrollo profesional o capacitación en servicio de directores/as a lo largo de tres años. Empezó a incluir también escuelas de nivel secundario y abrió un subprograma especial para escuelas rurales.

El programa se fue modificando gradualmente, conforme la investigación y la práctica iban mostrando más elementos que necesitaban ser atendidos para promover cambios de manera efectiva y permanente. Sin embargo, el desarrollo de capacidades fundamentales en lectura, escritura y matemáticas – que son las áreas curriculares que el SIMCE evalúa -- se mantuvo como una prioridad. Las escuelas P-900 – que llegaron a sobrepasar las 1500 hace algunos años – fueron el foco de muchos programas específicos de mejora escolar y eran supervisadas entre 8 y 16 veces al año por funcionarios provinciales y centrales.

Similares en enfoque y objetivos han sido la iniciativa de Nivelación para la Excelencia de Bogotá, cuyas diversas actividades son guiadas por el análisis de los resultados de las escuelas en las pruebas, las escuelas de jornada completa de Uruguay establecidas preferentemente en áreas urbanas empobrecidas, surgidas de la constatación que ellas habían obtenido mejores logros en la primera evaluación nacional, la jornada completa introducida también en Chile con posterioridad.

En términos generales, las políticas de "focalización", en un contexto caracterizado por la presencia de programas económicos de ajuste estructural de corte más bien ortodoxos durante los noventa y la hasta ahora insuficiente recuperación del crecimiento económico, se ha convertido en una característica común de las políticas sociales en la mayoría de los países de América Latina. Esto requiere indicadores técnicos sólidos y socialmente aceptables con los cuales identificar las poblaciones objetivo de programas especiales de apoyo, y las evaluaciones se han constituido en un insumo importante en este sentido.

4. Realizar investigación y evaluar el impacto de programas

Si consideramos que el aprendizaje - académico o de otra índole - es la misión central de la educación, es posible afirmar que la "variable dependiente" más importante para la investigación educativa orientada a la formulación de políticas no estaba disponible en gran parte de Latinoamérica antes de que los sistemas de evaluación comenzaran a generar resultados. Y dado que la investigación debería ser un insumo importante para la toma de decisiones, la región mostraba un serio retraso.

Antes de los 90, cuando se quería estimar los efectos sobre el aprendizaje de, por ejemplo, la relación estudiante/profesor, las calificaciones formales de profesores/as, la disponibilidad de libros de texto y otros materiales de aprendizaje, así como del nivel de gasto público en educación, la información sobre el rendimiento debía extraerse de otras medidas como las tasas de matrícula, repetición o deserción escolar. Los resultados de las pruebas nacionales han proporcionado un indicador crucial para determinar los efectos de la escolaridad en el aprendizaje. Hasta 1999, en toda América Latina no se podía encontrar más de una veintena de estudios empíricos sobre posibles asociaciones entre las diferencias en rendimiento, diferencias en las características intra-escolares o de aula y diferencias en factores externos tales como las características educativas y la condición socioeconómica de las familias.

Apenas seis años más tarde, una consulta bibliográfica rápida de 48 horas reveló la existencia de no menos de 150 estudios para 16 países, que exploran los efectos sobre el rendimiento de factores tales como (a) la condición socioeconómica de los estudiantes, el trabajo durante el año escolar, la raza y la condición nutricional, (b) el salario, calificaciones y prácticas de enseñanza en aula de los docentes, (c) características de la escuela tales como la cultura organizacional, la autonomía, la composición socio-económica de los distintos grupos, las prácticas de evaluación en aula, la cobertura del plan de estudios, las políticas sobre la realización de tareas en casa, el liderazgo de los directores, la infraestructura, la disponibilidad de libros de textos y (d) el financiamiento y/o administración pública o privada de la escuela, la presencia de programas de bonos escolares (*vouchers*), de programas de alimentación escolar, las prácticas de retención/repetición, el número de alumnos por aula, y la duración de la jornada escolar.

Algunos de estos estudios exploran un rango completo de posibles influencias sobre el desempeño y procuran medir su importancia relativa, mientras que algunos examinan el peso de y los mecanismos por los cuales una variable particular influye sobre (o cambia con) el desempeño. Otros estudian los cambios en las relaciones entre las variables predictivas y los resultados a lo largo del tiempo.

Algunos investigadores utilizan los puntajes promedio de las escuelas para seleccionar escuelas "exitosas" y escuelas con un alto índice de "fracaso", en las cuales después conducen estudios cualitativos en profundidad sobre los factores que pueden explicar las diferencias en el desempeño o recogen evidencia para determinar la validez de los puntajes de las pruebas como medidas de la calidad de las escuelas.

Los resultados de las pruebas se están utilizando cada vez más para estimar el impacto de políticas y programas educativos y sociales en general. Así, por ejemplo, los ministerios de educación incluyen muestras mayores de algunas poblaciones escolares de manera de poder hacer algunas generalizaciones válidas sobre el impacto de intervenciones especiales, tales como escuelas bilingües indígenas (en Perú y Guatemala), programas de expansión escolar basados en la comunidad (Nueva Escuela Unitaria de Guatemala y escuelas del PRONADE), perspectivas nuevas o experimentales para planes de estudios (Escuelas de Excelencia y el SIMAC en Guatemala), y mecanismos diversos de gestión y financiamiento (escuelas privadas subvencionadas en Chile y Nicaragua). Con frecuencia se ha permitido e incluso se ha estimulado a los investigadores a usar los resultados de las pruebas nacionales y/o re-aplicaciones o adaptaciones de los instrumentos de las pruebas para evaluar programas específicos, tales como la Escuela Nueva en Colombia, Educatodos en Honduras, Desayunos Escolares, Escuelas Activas y escuelas Fe y Alegría en Perú, etc.

5. Revisar expectativas curriculares o establecer estándares de aprendizaje

Hoy en día, se entiende que incrementar y aclarar las expectativas de aprendizaje constituye un elemento crucial para el mejoramiento de la enseñanza y el aprendizaje.

La mayoría de especialistas educativos, incluyendo algunos de los que estuvieron cercanamente involucrados en las reformas curriculares que los países de América Latina emprendieron a finales de los 80 y comienzos de los 90, estaría de acuerdo en que los objetivos curriculares no son los suficientemente claros, a menudo son contradictorios y por lo general están lejos de fomentar niveles de competencia que puedan asegurar un desarrollo adecuado de la ciudadanía y de la productividad que requiere el actual mundo globalizado.

Por otra parte, los resultados de las evaluaciones sugieren que incluso los objetivos existentes, que son difusos y no demasiado exigentes, no se están alcanzando. Esto es así en parte porque han sido mal definidos y no pueden ser utilizados para articular o alinear en forma efectiva la producción de libros de texto, los planes de estudio y la capacitación docente. Las unidades de evaluación demandan que los equipos de desarrollo curricular revisen y aclaren las expectativas de aprendizaje, pero con frecuencia se abocan ellas mismas a esta tarea, manteniendo distintos niveles de consulta con los equipos curriculares. En este proceso, los equipos toman en consideración tanto los resultados de aprendizaje deseables y el rendimiento "real" tal y como es medido por las evaluaciones nacionales previas.

Cuando la mayoría de sistemas de evaluación nacional en América Latina empezó a desarrollar pruebas referidas a criterios, y no ya a normas, experimentaron grandes dificultades para definir los tipos de tareas académicas que se debería esperar que los estudiantes puedan realizar si es que han logrado dominar un área particular de los contenidos curriculares. También confrontaron – y muchos siguen confrontando – grandes desafíos para definir los varios niveles de desempeño posibles (generalmente básico, aceptable o suficiente y competente) en cada uno

de los criterios seleccionados. Allí donde seató alguna consecuencia a los resultados de esas evaluaciones, o donde se ha realizado esfuerzos importantes por devolver resultados a las escuelas, los contenidos de las pruebas han empezado a cumplir la función de "aclaradoras" del currículo – si bien no todos los curriculistas aprobarían el que ello sea así. La amplitud de las consultas a docentes, expertos y productores de textos que pueda realizar la agencia evaluadora también ha contribuido en algunos casos a generar un proceso de redefinición de metas curriculares nacionales.

Sin embargo, la influencia de los resultados de las pruebas nacionales en la redefinición de las expectativas de aprendizaje y de las metas curriculares ha sido relativamente menor si se la compara con el impacto que viene teniendo en ese sentido la participación en pruebas internacionales.

Los resultados de los pocos países que han participado en TIMSS, PISA y algunas otras evaluaciones internacionales no siempre fueron congruentes con la distribución de estudiantes entre los distintos niveles de rendimiento fijados en sus propias pruebas nacionales (por ejemplo, aprobado/reprobado; aceptable/no aceptable; o niveles competente, satisfactorio o básico) o con los índices aprobatorios nacionales⁶. Esto impulsó a que varios países echaran una nueva mirada a sus currículos, libros de texto, prácticas pedagógicas y políticas de capacitación de profesores en busca de posibles explicaciones y a profundizar sus esfuerzos de priorizar y aclarar metas de aprendizaje mediante el establecimiento de estándares académicos.

Los primeros esfuerzos fueron los desplegados por los países centroamericanos en 1999, bajo la conducción de la Secretaría General del organismo de Coordinación Educacional y Cultural de la región (SGCECC). Equipos de profesionales de los ministerios de educación de **Panamá, Costa Rica, Nicaragua, Guatemala, Honduras y El Salvador** se abocaron a construir indicadores de contenido y desempeño sobre la base de sus currículos de educación primaria -, cuando éstos existían y/o habían sido recientemente revisados. En unos pocos casos, donde no había un currículo establecido, estos indicadores se construyeron sobre la base de otras pautas curriculares, tales como el contenido de los libros de texto.

Pocos de estos países habían desarrollado antes algún tipo de prueba a escala nacional, aunque sí, en algunos casos, tenido alguna experiencia con pruebas estandarizadas aplicadas en algunos proyectos de cooperación internacional y participado en el desarrollo de la primera evaluación regional de la UNESCO, cuyos resultados pueden haber sido un referente implícito para el desarrollo de dichos estándares. Se esperaba que tan pronto como los estándares fueran "aprobados" y puestos en ejecución formalmente, la capacitación de profesores, la elaboración de libros de textos y las pruebas estandarizadas tendrían que ser alineadas con éstos. Pocas, si alguna, de estas metas se materializaron, en parte porque el proceso de diseño de estándares ocurrió como parte de esfuerzos "locales", poco publicitados y realizados con pocos recursos.

Desde entonces, **Honduras y Guatemala** han avanzado sendas propuestas de estándares para todos los grados de la educación básica, para lo cual sí han tomado en consideración los resultados logrados en evaluaciones nacionales anteriores y con los cuales empiezan ya a alinear un conjunto de instrumentos de evaluación de aprendizajes. En **Nicaragua**, el conjunto de estándares de la SGCECC (que había comenzado a ser desarrollado incluso antes de que el proyecto regional empezara) fue utilizado para revisar el currículo y sirvió de insumo para la construcción de su primera evaluación nacional, aplicada en el 2003.

Colombia desarrolló y estableció estándares en el 2003, tomando como uno de sus referentes la evidencia disponible sobre las capacidades actuales de los estudiantes según lo evaluado por el ICFES, institución colombiana que ha sido directa o indirectamente responsable de la

⁶ La repetición de grado para estudiantes que no aprobaron las evaluaciones escolares es una práctica común en los países de América Latina.

construcción y análisis de la mayoría de las evaluaciones nacionales llevadas a cabo en el curso de varios años. Los estándares se centran en competencias para resolver problemas de tipo real en lenguaje, matemáticas, ciencias naturales y sociales, y, más recientemente, en ciudadanía. Se espera que sirvan, en primer lugar, como guías para el desarrollo de programas de mejoramiento escolar, y que se conviertan en el conjunto básico de expectativas de aprendizaje que el ICFES utilizará en el corto plazo. También es significativo el esfuerzo para incluir elementos éticos y valores que deben acompañar el desarrollo de cada una de las habilidades académicas incluidas bajo cada estándar. Esto es una respuesta a una muy fuerte demanda social para que la escuela continúe teniendo un papel primario en forjar "personas completas", y que no sólo funja como centro de instrucción académica. Esta demanda contrasta con la naturaleza estrictamente académica de la mayoría de estándares en los Estados Unidos y constituirá un gran desafío para el ICFES cuando se precise evaluar el logro de tales elementos.

En **Chile**, el análisis del SIMCE y de los resultados de las pruebas internacionales ha mostrado poco incremento significativo en los puntajes promedio, a pesar de los grandes aumentos en el gasto público en educación, de las reformas curriculares, y de la inversión en varios proyectos a gran escala. Se introdujo sucesivamente varios nuevos instrumentos y aclaraciones curriculares, antes de optar por la construcción de estándares de contenido y desempeño, conocidos ya como Mapas de Progreso y Niveles de Logro, respectivamente. Los Mapas y Niveles se nutrieron, entre otros elementos, de los resultados obtenidos por los alumnos chilenos en sus pruebas nacionales e internacionales. Se convertirán gradualmente en referentes básicos para la elaboración de los instrumentos y se usan ya en el reporte de resultados del SIMCE. Se pretende que sirvan también para guiar la reflexión, la planificación y la práctica pedagógica de los docentes, en subáreas específicas de matemáticas, lenguaje, ciencias sociales, ciencias naturales e inglés como segunda lengua. También se está elaborando estándares para educación física, educación preescolar, educación técnica y arte.

El gobierno federal **argentino** aprobó en el 2004 los denominados Núcleos de Aprendizajes Prioritarios. Se trata de un intento por establecer prioridades y aclarar expectativas de aprendizaje de modo de contrarrestar, con un mandato nacional, las brechas de rendimiento que, al parecer, han crecido desde la descentralización. En general, los bajos niveles de desempeño de los estudiantes y las diferencias en los puntajes promedio de las provincias, sirvieron como argumento en las negociaciones entre el gobierno nacional y los gobiernos provinciales y autoridades sectoriales, que condujeron a un acuerdo preliminar sobre los contenidos del NAP.

6. Promover la responsabilización, la rendición de cuentas y la "libre elección" de escuelas

En general, las pruebas estudiantiles con altas consecuencias no son comunes en América Latina, a excepción de pruebas de tránsito a niveles superiores del sistema educativo como en Costa Rica, República Dominicana y, más recientemente, en El Salvador. Pero pueden incluirse también entre ellas las pruebas nacionales cuyos resultados sirven para seleccionar a escuelas para recibir premios o sanciones, ya sean simbólicos o efectivos, tales como la publicación de rankings –de todas las escuelas o de las mejores --, premios a la "excelencia" de instituciones, incentivos para equipos docentes o ascensos individuales en la carrera magisterial.

La práctica más simple y extendida a este respecto es la identificación de las escuelas más destacadas en rankings nacionales. Algunos países como **Chile**, que realizan pruebas censales, publican la lista completa de escuelas con sus resultados promedio en cada una de las áreas evaluadas. A veces lo hacen en forma separada para grupos de escuelas según su condición socioeconómica. Otros países, como **El Salvador**, no publican la lista completa, sino que incluyen sólo a las escuelas de mejores desempeños, resaltando aquellas que recién han logrado posicionarse en lo más alto del ranking o que han "avanzado" más desde la ronda de evaluación previa.

Un trabajo delicado y efectivo con la prensa a llevado a veces a la publicación de historias de interés general sobre los esfuerzos que hayan conducido a mejoras significativas y que aparentemente motivan a otras escuelas a esforzarse para lograr mejores calificaciones.

Esta práctica suele convertirse en un incentivo mayor a los "ganadores" para lograr progresos adicionales. Hay que señalar que los educadores en América Latina, como los de muchos otros países del mundo, suelen estar opuestos a la publicación de rankings, y consideran que los resultados de las pruebas sólo deberían compartirse al interior de la comunidad escolar⁷.

Desde 1997, la ciudad de **Bogotá** concede El Galardón a la Excelencia a las escuelas que constituyen buenos modelos de gerencia y cuyos resultados en las pruebas presentan mayores mejoras. Hay una participación cada vez mayor de escuelas privadas en este esquema. De otro lado, las escuelas públicas más destacadas han sido invitadas a participar de una Red de Acción para la Excelencia a través de la cual, y mediante el uso de tecnologías de comunicación virtual, recogen y comparten información sobre prácticas exitosas - información que es luego difundida a otras escuelas.

En **Colombia**, los datos de las pruebas son uno de los criterios usados para seleccionar y supervisar escuelas privadas que atienden a estudiantes de pocos recursos económicos provistos con bonos.

Dadas las dificultades para identificar exactamente qué insumos educativos "garantizan" mejores resultados de rendimiento en distintas condiciones, los diseñadores de políticas analizan varias combinaciones de incentivos y mecanismos de responsabilización con los cuales movilizar actores e instituciones clave en su propia búsqueda de las "mejores" direcciones a seguir. Los incentivos al desempeño docente y escolar son un ejemplo de tales mecanismos, y pueden ser concedidos sobre la base de varios criterios.

Aunque en muchos países del mundo la investigación ha demostrado que las compensaciones económicas no son el mejor incentivo al desempeño docente, se ha mostrado también que éstas pueden servir para tal propósito cuando se les incluye dentro de una estrategia mayor para asegurar que los profesores concentren sus esfuerzos en el rendimiento estudiantil y, particularmente, cuando requieren de un esfuerzo colectivo para lograr los premios⁸.

En América Latina, la escala salarial docente suele ser más bien plana, y está determinada principalmente por los años de experiencia, la permanencia en la carrera pública y por grados académicos o capacitaciones adicionales. Los retornos a los esfuerzos o logros en aula son pocos o nulos. Dado que diversos estudios sobre factores que afectan el rendimiento estudiantil han demostrado que el efecto de los profesores sobre el aprendizaje depende no de las calificaciones formales o años de experiencia de los docentes, sino de su esfuerzo y dedicación y de sus prácticas pedagógicas, éste parecería un buen terreno para establecer incentivos monetarios.

Un criterio que se ha utilizado para asignar este tipo de incentivos es el éxito de un profesor/a (o de un grupo de profesores) en ayudar a sus estudiantes a aprender. Los resultados en las evaluaciones nacionales - o cambios en dichos resultados- cuando están disponibles para todas las escuelas, junto con otros indicadores de los esfuerzos exitosos de las escuelas en proveer a sus estudiantes con oportunidades eficaces y equitativas de aprendizaje, pueden ser utilizados como medidas de tales resultados.

⁷ Ver, por ejemplo, el documento de discusión y contribuciones al foro electrónico sobre este tema en [http://www.preal.org/Biblioteca.asp?Id_Carpeta=216&Camino=315|Grupos%20de%20Trabajo%20\(GDT\)/38|Evaluaci%F3n%20y%20Est%20Indares/213|Foro%20Mitosis/216|Segundo%20Ciclo%20Los%20Rankings](http://www.preal.org/Biblioteca.asp?Id_Carpeta=216&Camino=315|Grupos%20de%20Trabajo%20(GDT)/38|Evaluaci%F3n%20y%20Est%20Indares/213|Foro%20Mitosis/216|Segundo%20Ciclo%20Los%20Rankings)

⁸ Odden y Kelly, 1997

Desde 1996, el SNED (Sistema Nacional de Evaluación del Desempeño de Docentes en escuelas con financiamiento público) **chileno** utiliza los resultados de las pruebas y las mejoras en éstos para asignar un "premio a la excelencia en el desempeño" a las escuelas más destacadas – premio que cubre cerca del 25% del total de la matrícula en cada región.

Además de los puntajes promedio y de las mejoras en éstos, cada escuela debe demostrar también que no está seleccionando estudiantes (lo que significa que no le niegan la matrícula a niños/as que les plantean desafíos especiales y que podrían poner en riesgo el promedio en los resultados SIMCE), que dan cabida a la participación de los padres, que proporcionan buenas condiciones de trabajo y operan sin problemas, y que asumen iniciativas para desarrollar actividades y proyectos especiales. Se utiliza múltiples fuentes de información para construir un índice agregado del desempeño, que asigna pesos distintos a cada uno de los factores mencionados. Para evitar prejuicios que favorezcan a ciertos tipos de escuela, la selección se realiza dentro de grupos homogéneos de escuelas dentro de cada región, y tomando en cuenta si son urbanas o rurales, si son escuelas de nivel primario/secundario, y la condición socio-económica de las familias. Este incentivo pretende compensar la falta de mecanismos de competencia y la información insuficiente con que cuentan los padres de familia, que no permite que el sistema de bonos chileno funcione de la manera prevista en algunas áreas del país (es decir, para facilitar la libre elección de escuelas y como una forma de presión para la mejora de las escuelas)⁹.

Desde 1993, los profesores **mexicanos** pueden participar en la Carrera Magisterial, un sistema de incentivos salariales que también utiliza los resultados de los estudiantes en las evaluaciones como un indicador del desempeño profesional. A diferencia del SNED chileno, que funciona como un incentivo colectivo para toda la escuela¹⁰, la Carrera Magisterial es un premio individual. Este incentivo puede representar entre el 27% y el 224% del salario básico docente¹¹, y es permanente, lo que constituye otra diferencia importante con el modelo chileno, que, según vimos, ofrece un monto mucho menor y se reasigna cada dos años. Otra diferencia entre ambos modelos es que la Carrera Magisterial da menor importancia a los resultados de los estudiantes en las pruebas.

Todavía se sabe poco sobre la efectividad de estos programas desde el punto de vista de la mejora en los aprendizajes. Hasta hace poco tiempo, los resultados de las pruebas nacionales **mexicanas** no estaban disponibles para investigadores o evaluadores, ni habían sido diseñados para permitir comparaciones en el tiempo. Algunos estudios¹² sugieren que el sistema mexicano es defectuoso porque los esfuerzos de los/as profesores/as para mejorar su enseñanza no pesan lo suficiente en el conjunto de criterios que permiten el acceso y la movilidad dentro del programa de Carrera Magisterial. De otro lado, tampoco premia características tales como el dominio de los contenidos del área de enseñanza, cuyo efecto positivo sobre el aprendizaje es reconocido. Un estudio que comparó a un grupo de profesores que había acumulado una buena cantidad de "puntos" antes de que sus estudiantes fueran evaluados, con otro grupo cuyas probabilidades de lograr el puntaje mínimo para calificar era mucho menor, demostró que no existían diferencias significativas en los resultados de los estudiantes de ambos grupos en las pruebas - aunque ambos grupos podrían haber "enseñado para la prueba" de modo de calificar

⁹ Mizala y Romaguera, 2004.

¹⁰ Cada cuatrimestre el 90% del monto total asignado a una escuela es distribuido en partes iguales entre todos los profesores, mientras que el 10% se asigna a los "mejores profesores", quienes son seleccionados de acuerdo a reglas establecidas por cada escuela (Chile n/d).

¹¹ Vegas y Umansky (2005) estiman que los incentivos pueden constituir entre 20% y 197% de los salarios básicos.

¹² Mizala y Romaguera, 2005; Vegas y Umansky, 2005

para ascensos en la "Carrera". Se ha encontrado también algunas consecuencias "involuntarias", tales como la tendencia de los docentes a reubicarse en áreas urbanas donde hay mayor acceso a oportunidades de desarrollo profesional.

Las limitaciones de estos y otros esquemas de incentivos similares - que han sido aplicados en Bolivia y El Salvador, y como parte de programas especiales a gran escala, en varios países de América Central - pueden deberse a ciertas fallas en su diseño o implementación, más que al hecho de que, como muchos arguyen, los incentivos externos no puedan, por sí mismos, producir cambios en el comportamiento docente. Sin embargo, un "impacto" importante de los esquemas existentes es su contribución a una aceptación implícita del uso de las evaluaciones de estudiantes y profesores, como mecanismos válidos no solamente para recompensar el trabajo de los docentes, sino también para determinar en qué consiste una enseñanza de buena calidad. Incluso a pesar de que hasta ahora las pruebas y los métodos de evaluación del desempeño docente existentes sólo pueden evaluar un espectro limitado de lo que se considera como formas buenas y efectivas de enseñanza y aprendizaje, la confianza parcial pero fuerte en el uso de instrumentos verificables y objetivos para medir el aprendizaje y los cambios introducidos es claramente un paso adelante en el esfuerzo por mejorar la calidad.

Una forma adicional de utilizar los resultados de las evaluaciones es para facilitar que los padres puedan elegir libremente a qué escuela enviar a sus hijos/as, lo que, de forma indirecta, sirve para presionar a las escuelas a mejorar la calidad de sus servicios y sus resultados de manera de "permanecer en el negocio". Esto es así particularmente en aquellos lugares donde el financiamiento público se asigna directamente a los niños (y no a las escuelas) – como en el caso de sistemas de bonos escolares.

El ejemplo mejor conocido de esta clase de utilización de resultados es el del SIMCE, el sistema de evaluación **chileno**, que se aplica en casi todas las escuelas y a estudiantes de algunos grados seleccionados que van cambiando de un año a otro¹³.

Los directores/as de cada escuela chilena reciben un informe detallado del desempeño de su escuela en cada una de las áreas curriculares que han sido evaluadas y que presenta también diversas comparaciones con escuelas similares y promedios nacionales, así como información sobre las tendencias observadas en el tiempo a este respecto. Según se mencionó antes, se incentiva a los directores para que compartan esta información con padres y profesores a fin de planificar mejoras. Además, se pone a disposición del público general (a través de la Internet y la prensa) los resultados obtenidos en las pruebas por cada una de las escuelas de cada región del país, junto con los promedios para cada área, información sobre los cambios observados desde las rondas anteriores y comparaciones con escuelas con características socioeconómicas similares.

La prensa recibe un portafolio con datos adicionales e información contextual, así como pautas orientadas a mejorar la interpretación de los resultados de las pruebas (y evitar algunos errores típicos) y algunas explicaciones posibles sobre las diferencias entre escuelas, según su situación y en el contexto de distintas tendencias. Se proporciona también los promedios nacionales en cada uno de los temas evaluados, y se les compara con rondas de evaluación previas, y se incluye algunas explicaciones posibles para los cambios observados o la inexistencia de éstos. La información provista resalta también los resultados regionales promedio, así como los cambios observados en el tiempo y desviaciones importantes de los promedios nacionales. Asimismo, se informa sobre los promedios y cambios en éstos para varios grupos de escuelas con características socioeconómicas similares, los cuales son desagregados según el tipo de institución (escuelas privadas donde se paga matrícula, privadas subvencionadas y escuelas municipales) y por el nivel de selectividad de las instituciones. Finalmente, se incluye algunos

¹³ 4 y 8 grados del nivel básico de educación y 2do año de media.

ítems seleccionados junto con la distribución nacional de respuestas para cada una de las múltiples opciones de respuesta (o, en el caso de sub-pruebas escritas, muestras de productos satisfactorios, aceptables y de nivel pobre), y explicaciones sobre su relevancia para el currículum y sobre las posibles fuentes de error entre los estudiantes.

Toda esta información procura revertir la tendencia de la prensa a construir clasificaciones simplistas sobre cuáles son "las mejores" o "las peores" escuelas o a publicar titulares de historias escandalosas sobre la carencia de mejoras en el rendimiento general o la calidad "en descenso" de las escuelas municipales. De este modo se busca brindar mejor información para que el público pueda evaluar las políticas y programas, y para que los padres puedan elegir mejor a qué escuela enviar a sus hijos.

Sólo recientemente, desde el 2003, se ha empezado a preparar informes especiales para los padres, que incluyen los promedios de las escuelas de sus hijos/as y comparaciones con los resultados obtenidos en años anteriores, así como con los promedios nacionales y con aquellas escuelas de características socioeconómicas similares. Los informes incluyen también un breve llamado para que los padres se involucren más en el aprendizaje de sus hijos/as, para que exijan información de las escuelas sobre lo que están haciendo para mejorar el rendimiento y para que discutan con otros padres las formas en que pueden colaborar con la escuela.

Una revisión reciente de investigaciones disponibles que buscaba evaluar el supuesto de que el acceso a la información sobre el rendimiento serviría para ejercer una presión efectiva sobre las escuelas para mejorar la oferta, mostró que los resultados al respecto no son concluyentes¹⁴. Las distintas propuestas de investigación y el distinto énfasis en diversas variables dependientes y de control han conducido a hallazgos bastante contradictorios sobre la efectividad relativa de varios tipos de escuelas y sobre si el esquema competitivo ha servido o no para estimular mejoras en el sistema o en las escuelas municipales subvencionadas que abastecen a los pobres. Las escuelas de administración privada parecen haber respondido al mecanismo de libre elección seleccionando estudiantes con mayores probabilidades de éxito y no con mejoras en la calidad.

El estudio sugiere también que los padres no "votan con los pies", pues cuando tienen que elegir una escuela para sus hijos/as no confían sólo en la información "objetiva" acerca de las diferencias en los puntajes de rendimiento de las escuelas. Cerca del 90% de los padres en todos los niveles socio-económicos, considera sólo una o dos escuelas como posibles opciones, situación que no refleja simplemente el acceso diferenciado a distintos proveedores de servicio o a fuentes formales de información sobre la calidad. Además, los padres que sí consideran más de una opción para matricular a sus hijos/as en el primer grado, no parecen mantener un criterio uniforme en la selección de sus alternativas. Solamente el 38% de los padres tiene entre sus opciones a escuelas con resultados SIMCE similares, mientras que casi la mitad restringe su selección a escuelas con condiciones socio-económicas similares, y el 60% elige sólo entre escuelas privadas, o entre las privadas subsidiadas o entre escuelas municipales. Casi el 70% elige sólo entre escuelas religiosas o entre escuelas laicas. De los padres que incluyeron más de una escuela entre sus opciones, sólo el 26% matriculó a sus hijos/as en la que tenía una mejor posición en el ránking, y menos del 1% de todos los padres de niños de primer grado que fueron encuestados dijeron haber elegido la escuela a raíz de SIMCE. La proximidad y el costo jugaron un papel importante entre las razones de aquellos con menos recursos, mientras que la promesa de mejores contactos sociales fue un criterio particularmente fuerte entre los más ricos.

En conclusión, 23 años de mecanismos de elección supuestamente basada en la información sobre los resultados en las pruebas, no han producido una masa crítica de consumidores capaz de presionar a las escuelas para mejorar su calidad. La disponibilidad de información "objetiva", tal como la que proviene de las evaluaciones, tampoco parece haber afectado los criterios de selección de escuelas. Además, se ha documentado la existencia de efectos indeseables o

¹⁴ Elacqua y Fabrega, 2004

"perversos" de la estrategia de promoción de la libre elección de escuelas, tales como el incremento en la brecha entre los pobres y los más ricos, la desmoralización de los "perdedores", la "enseñanza para las pruebas" y la reducción del currículum, así como la evidencia de algunas prácticas corruptas. Como resultado de esto, el marco y las metas del SIMCE están siendo revisados. Se está poniendo un mayor énfasis en la provisión de insumos para la capacitación docente y para el mejoramiento de las escuelas, lo que requiere de cambios sustantivos en los contenidos y el diseño de las pruebas, así como la renovación de las estrategias de difusión. Hay que resaltar que el impacto de todos estos cambios será evaluado con las pruebas SIMCE, que aunque serán renovadas, no perderán las características que permiten las comparaciones en el tiempo.

Colombia y algunos otros países también han experimentado con otros modelos de políticas de elección de escuelas. Las libretas escolares - incluyendo los puntajes de desempeño de los estudiantes en las pruebas de nivel estatal - han sido utilizadas en varios estados **brasileros** para promover una mayor participación de los padres en las escuelas de sus hijos/as y para promover mayores grados de responsabilización de parte de las escuelas, tanto frente a los padres como frente a las oficinas de nivel superior. Aunque no han sido diseñadas para orientar la elección de escuelas, estas políticas eventualmente cumplen ese papel.

7. Mejorar la gestión pedagógica e institucional en las escuelas o áreas escolares

a. Uso de los resultados para diseñar planes de mejoramiento escolar

Es parte de un discurso común asumir que las evaluaciones a gran escala - como tantos otros componentes de la reforma educativa - tendrán un impacto en los logros de aprendizaje solamente si sirven para generar cambios positivos la escuela y el aula. Esto, se piensa, puede lograrse a través de presiones o incentivos a los docentes y estudiantes mediante el establecimiento de altas implicancias para los resultados de las pruebas. Otros plantean que esto puede lograrse mediante una asignación apropiada de recursos, lo que ha demostrado tener un impacto sobre el rendimiento estudiantil. Tal sería el caso de los programas o intervenciones a gran escala discutidas previamente. Pero los resultados diferenciados alcanzados por varios grupos hacia los que esas intervenciones apuntan, de un lado, o la adhesión a perspectivas organizacionales o institucionales particulares sobre el cambio social, de otro, han llevado a muchos actores - incluso antes de que hubiera evidencia disponible - a sostener que las rutas más productivas a través de las cuales se puede afectar el aprendizaje estudiantil es proveyendo información para el diseño e implementación de estrategias de mejoramiento escolar.

Además, el advenimiento de procesos de descentralización en muchos países de la región ha traído consigo un aumento de normas e incentivos para el diseño de planes de mejoramiento educativo a nivel subnacional (regiones, departamentos, provincias) y local (distritos). Estos deben basarse, al menos parcialmente, en una lectura de los resultados de las evaluaciones realizadas a nivel escolar, local o nacional, dondequiera que estas se encuentren disponibles.

También caracterizan a algunos de los procesos en curso de reforma de la gestión educativa la promoción de una mayor autonomía escolar, que conllevan el establecimiento de normas e incentivos para estimular el diseño participativo y la implementación de planes de mejoramiento a nivel de la escuela. Los Proyectos de Desarrollo Institucional (**Perú**), los Proyectos Educativos Institucionales y los Planes de Mejoramiento (**Colombia**), los Proyectos de Mejoramiento Educativo (**Uruguay**), los Proyectos de Innovación Educativa (**Paraguay**) y los Proyectos de Mejoramiento Educativo (**Chile**), se han convertido en elementos comunes en los procesos de reforma.

Estos proyectos usualmente requieren el desarrollo de actividades no rutinarias así como la movilización de habilidades de enseñanza no tradicionales, tales como el diseño de proyectos, el trabajo en equipo y la preparación de informes. Su alcance va más allá de la "simple" mejora de

los logros de aprendizaje, e involucra la definición de rasgos distintivos y metas de las escuelas con respecto a muchos otros objetivos educacionales. Sin embargo, se espera cada vez más que los planes y proyectos estén basados en una autoevaluación de los resultados de aprendizaje obtenidos por las instituciones educativas, y que se derivan de las estadísticas disponibles y de los resultados en evaluaciones a gran escala, ya sean muestrales o censales.

En **Colombia**, se empezó a desarrollar planes de mejoramiento escolar (Proyectos Educativos Institucionales - PEIs) a principios de los 90. Un informe de 1999¹⁵ indica que para ese entonces, aproximadamente el 90% de escuelas había diseñado un plan. Sin embargo, aparentemente sólo desde la elaboración del plan sectorial de educación 2002-2006, se ha establecido orientaciones explícitas que señalan que las escuelas deben desarrollar dichos planes sobre la base de los resultados de evaluaciones disponibles (la prueba Saber para educación básica y las pruebas estatales de salida).

De acuerdo a informes no oficiales, las metas y los objetivos específicos de los PEIs - aunque han sido importantes y útiles para otros propósitos - eran demasiado difusos para servir en forma efectiva como instrumentos de movilización para mejorar los logros de aprendizaje, y han tenido que complementarse con los nuevos *Planes de Mejoramiento*. Éstos tienen que incluir una serie de acciones y tareas orientadas a la consecución de una clara meta de incrementar el aprendizaje de los estudiantes, así como una definición clara de las contribuciones, métodos, procedimientos, recursos y responsabilidades que necesitan ser desplegados para obtener los resultados deseados. Estos planes tendrán que ser diseñados por los directores y coordinadores de área y discutidos con el cuerpo docente, y deberán usar información basada en el desempeño en las pruebas que las Secretarías de Educación departamentales deberán proporcionar a cada escuela con los resultados de las evaluaciones censales de 5º y 9º grado. Los exámenes de salida del onceavo grado, que ahora también están basados en los marcos y estándares curriculares nacionales, también se pueden utilizar como insumo para el diagnóstico¹⁶, mientras que los estándares recientemente establecidos sirven para guiar el establecimiento de metas¹⁷.

Hasta el momento de realización de este estudio, el Ministerio de Educación todavía estaba diseñando guías y estrategias de capacitación para permitir que las escuelas lleven a cabo esta nueva actividad de planificación. Esto probablemente se benefició del hecho que el ICFES¹⁸ había desarrollado una guía para el diseño de planes que contiene una serie de cuadros y procedimientos que utilizan los resultados de las escuelas en las pruebas además de algunas otras variables. Se esperaba que la identificación de las áreas más débiles en las escuelas condujera a la formulación de estrategias específicas de mejoramiento. Las guías fomentarían la participación de todo el personal, así como también alguna participación de los estudiantes.

También en Colombia, la ciudad de **Bogotá** organiza talleres de capacitación que se centran en desarrollar capacidades para analizar datos de rendimiento, así como para diseñar estrategias para mejorar la planificación de aula y las prácticas de enseñanza en las áreas curriculares que más mejoras requieren. Durante las capacitaciones también se hace recomendaciones sobre cómo mejorar las escuelas, las que sirven como posibles opciones a partir de las cuales modelar los planes escolares existentes.

Con frecuencia, los planes de mejoramiento escolar son el objetivo de pequeños fondos a los que las escuelas acceden a partir de procesos competitivos. Tal fue, por ejemplo, el caso de **Paraguay** donde el planeamiento escolar comenzó como un programa competitivo de

¹⁵ "Escuelas que protagonizan el mejoramiento institucional" PREAL Serie Mejores Prácticas 1,1,1999.

¹⁶ Altablero No. 24, 2003.

¹⁷ <http://www.mineducacion.gov.co/estandares/queson.asp>

¹⁸ Se debe notar que si bien el nombre ICFES, Instituto Colombiano de Fomento de la Educación Superior, sugiere que su campo de acción está limitado a los niveles educativos superiores y se le vincula tradicionalmente con el diseño de los procesos de admisión universitaria, el ICFES ha provisto por muchos años la asistencia técnica para los procesos de evaluación de la educación básica y se ha propuesto la misión de proveer servicios para los procesos escolares de planificación institucional.

"innovación educativa". El objetivo principal del programa era promover la participación local o escolar para encontrar soluciones a la baja efectividad de la enseñanza, es decir, para mejorar el desempeño en áreas curriculares clave. Los maestros/as, directores/as y padres de familia recibieron orientaciones sobre cómo diseñar y presentar sus planes ante un comité regional de selección. Los premios podían llegar hasta los 4000 dólares americanos - cantidad bastante significativa en países en los cuales las escuelas rara vez cuentan con un presupuesto propio. Muchas veces la formulación de estos planes se convirtió en una actividad ritual, con escasa relación entre los datos presentados y las actividades con que se pretendía solucionar los problemas de eficiencia interna (repetición, deserciones, etc.) y de eficacia (tasas de promoción, calificaciones, resultados de evaluación externa) de las escuelas. Esto se debió en, en parte, a que las guías y manuales para las escuelas (y para los evaluadores de las propuestas) con frecuencia ponen demasiado énfasis sobre los procedimientos formales como un elemento importante de la "modernización".

Otras experiencias con planes de mejoramiento escolar también han tenido lugar en **Chile** desde 1992 y, más adelante, en **Uruguay**.

Los programas **chilenos** de mejoramiento escolar – PME – fueron diseñados por personal de las escuelas, quienes tuvieron que analizar la situación de las mismas e identificar, entre otras cosas, las áreas de aprendizaje más débiles entre los estudiantes y desarrollar una propuesta especificando objetivos, beneficiarios, estrategias, actividades, responsabilidades de varios actores y un plan de evaluación. La propuesta era presentada a la administración central del proyecto para su evaluación y posible selección para recibir un fondo especial de desarrollo, que desde 1999, concedió subvenciones a más de 5600 planes - a un promedio de 6000 dólares americanos por plan seleccionado. Además de esto, las escuelas obtuvieron algún equipamiento educativo. Informes recientes sugieren un costo promedio de 10.000 dólares para proyectos de 30 meses de duración. Los acuerdos actuales de financiamiento proveen alrededor de 200 subvenciones anuales. El programa se convirtió luego en un componente del "Sistema de Calidad de la Administración de las Escuelas", un proceso de acreditación de la escuela orientado a promover mecanismos de mejoramiento sostenible entre las unidades del sistema.

Las PME **uruguayas** también se asignan sobre una base competitiva, aunque para montos promedio menores (3500 dólares) y periodos más cortos de implementación (24 meses).

Tanto en **Chile** como en **Uruguay**, las mejoras en el gobierno escolar y la participación de padres y profesores/as en la toma de decisiones, aumentaron la colaboración de los profesores/as en el desarrollo de actividades curriculares y planes de estudio relevantes a la escuela. Se ha reportado también que estas mejoras han servido para establecer un mejor clima escolar, así como un incremento en la demanda por capacitaciones en servicio más específicas. Según lo medido por las pruebas, hay también cierta evidencia¹⁹ de que la puesta en práctica de esos planes de mejoramiento ha conducido a mejoras en los resultados.

En el caso **chileno**, un estudio de 1998 encontró un modesto pero promisorio efecto en los resultados SIMCE de escuelas que habían desarrollado planes de mejoramiento²⁰. Otros indicadores de logro, tales como las tasas de repetición, también sugieren un efecto positivo de los PME. Otras evaluaciones, sin embargo, han identificado varias consecuencias involuntarias e indeseables del diseño del programa, y sugieren que su naturaleza competitiva ha conducido a serias desviaciones e incluso a prácticas poco éticas de parte de algunos proveedores de servicios educativos²¹.

¹⁹ http://www.lobos-net.net/ilo/195_base/es/topic_n/t_71chi.htm

²⁰ Hay que tomar en cuenta que los ciclos de evaluación requieren varios años para poder evaluar un mismo grado, y el programa recién se inició en 1992.

²¹ Zibas, 2002

Un estudio realizado por la administración educativa **uruguay**, donde se compararon los resultados obtenidos por escuelas que atienden a estudiantes con características socioeconómicas similares, también mostró un modesto pero positivo efecto de las PMEs en los resultados académicos²².

Dado que la confianza exclusiva en los esquemas competitivos para seleccionar los "mejores" proyectos a financiar no garantiza una implementación satisfactoria, y como las capacidades para el diseño e implementación de buenos proyectos no son sólo escasas, sino que están distribuidas de manera muy desigual, la asistencia técnica resulta muy necesaria – aunque puede ser difícil de proporcionar, particularmente en países con geografías complejas y muchas escuelas.

b. Uso de los resultados para proveer oportunidades de desarrollo profesional para profesores/as y supervisores/as

Según muchos especialistas educativos, el uso de los resultados de las pruebas para la formación y capacitación docente es tan importante como su uso para el planeamiento escolar. No es de sorprender, entonces, que una de las primeras y más sencillas formas de utilización de los resultados de las pruebas en América Latina, haya sido a la publicación y distribución más o menos extensa de folletos de "análisis de ítems".

Con frecuencia, la publicación de los resultados generales de las pruebas ha estado acompañada por la publicación de ítems seleccionados de las pruebas para cada una de las áreas curriculares, y que usualmente reflejan distintos niveles de dificultad. Tales publicaciones suelen presentar el ítem completo con las múltiples opciones de respuesta, seguidas normalmente por una descripción de las habilidades curriculares específicas que estaban siendo evaluadas con esa pregunta y una explicación del por qué se les consideraba lo suficientemente importantes como para ser evaluadas en la prueba. Las publicaciones suelen analizar también las razones probables por las que los estudiantes podrían elegir la respuesta correcta o alguna de las respuestas incorrectas y, a veces, se ofrecen algunas sugerencias – tanto conceptuales como didácticas – sobre cómo los profesores podrían ayudar a los estudiantes a desarrollar la habilidad prevista. Por lo general se incluyó también información sobre la distribución porcentual de respuestas entre varias opciones, de modo de ayudar a los/as profesores/as a reflexionar sobre el desempeño real o posible de sus propios estudiantes y a pensar en cómo desarrollar mejor sus planes para asegurar mejores resultados²³.

En países donde las pruebas sólo se aplicaron a muestras de estudiantes de algunos grados, estas publicaciones con frecuencia sirvieron como una primera oportunidad para que los profesores y supervisores se informaran sobre los propósitos y métodos de las evaluaciones nacionales²⁴. Es por esto que tales informes "pedagógicos" generalmente incluyen algunas páginas introduciendo a los educadores al significado y uso potencial de las evaluaciones para su propio desarrollo profesional y al análisis de su posible desempeño en este tipo de pruebas. Asimismo, y dado que los informes generales rara vez logran llegar hasta las escuelas, las publicaciones incluyen también un resumen de los resultados nacionales o sub-nacionales. En la mayoría de países se esperaba que no sólo todas las escuelas recibieran copias de estos materiales, sino que éstas y los distritos escolares tomarían la iniciativa para analizarlos. Sin embargo, en muchos casos la distribución del material estuvo seriamente limitada por falta de presupuesto y, en la mayoría de países, no se ha hecho ningún esfuerzo para supervisar de qué modo se está utilizando estas publicaciones en la práctica, quiénes las están usando, ni mucho menos con qué resultados.

²² Informado en Espinola s/f.

²³ Un ejemplo de este tipo de informes puede encontrarse en Apéndice 4.

²⁴ Algunas veces, las escuelas y profesores reciben folletos informativos con un análisis similar de ítems, sin información sobre puntajes, los que son distribuidos antes de la aplicación de las pruebas. Un ejemplo puede encontrarse en http://www.simce.cl/doc/folletos_2003_2medio.pdf

En América Latina, una vez que las evaluaciones comenzaron a estar – parcial o totalmente – guiadas por marcos referidos a criterios, los informes dirigidos a los docentes se han centrado en descripciones de los distintos niveles de desempeño en cada una de las áreas o criterios de aprendizaje, cada una con ejemplos de uno o varios ítems o tareas que los estudiantes que alcanzan ese nivel de desempeño podrían responder o realizar correctamente.

Los primeros ejemplos se pueden encontrar en un informe **argentino** de 1999, así como en los informes de 1997 del SAEB **brasileño**. Tal como en el caso del análisis de ítems individuales descrito anteriormente, se presentó la distribución porcentual de estudiantes entre los varios niveles de desempeño y se hicieron algunos comentarios generales sobre qué habilidades generales o particulares parecen plantear las mayores dificultades para los estudiantes. También se discutió las implicancias de dichas limitaciones para el aprendizaje futuro y se hizo algunas sugerencias sobre cómo los docentes pueden promover mejoras.

En **Chile**, se entrega los resultados de las pruebas a cada una de las escuelas junto con una guía para ayudarlas a analizarlos. Esto busca promover una reflexión institucional sobre los resultados, sus posibles determinantes y sobre las prácticas usuales de evaluación y enseñanza en las escuelas. Varios estudios han mostrado que los/as directores/as y profesores/as valoran mucho los informes de evaluación como herramienta para el mejoramiento y especialmente para mantenerse enfocados en la calidad. Sin embargo, también hay evidencia señalando que una gran cantidad de profesores/as - cerca del 40%, según un informe oficial - consideran que la información del SIMCE no es útil para orientar las prácticas de aula²⁵. Además, parece que son las mejores escuelas las que reportan un mayor uso, mientras que aquellas que más mejoras necesitan son incapaces o no están dispuestas a utilizar los resultados de las pruebas para este fin. Actualmente, se está impulsando a las escuelas a organizar eventos especiales con el personal con el objetivo específico de discutir los informes.

En **Uruguay**, cuando salieron los primeros resultados de las pruebas del UMRE, se pudo observar diferencias importantes entre las pruebas nacionales y las calificaciones regulares de las escuelas. Mientras que el 96% de los estudiantes de sexto grado estaba siendo promovido, supuestamente por haber aprendido en forma adecuada lo que se esperaba que aprendan, los resultados del UMRE mostraron que solamente alrededor de dos tercios de los estudiantes estaban logrando aprendizajes "suficientes". Estas diferencias resultaron particularmente provocativas, dando lugar a muchas reflexiones, y fueron analizadas por directores y profesores en talleres especialmente convocados por la administración educativa.

Por otra parte, los instrumentos de evaluación fueron puestos a disposición de todas las escuelas poco después de la aplicación muestral controlada de la prueba junto con una invitación, al parecer extensamente aceptada, para que aplicaran la prueba a sus propios estudiantes y compararan los resultados con aquéllos de sus pares en escuelas similares - ya sea dentro de sus propios distritos o con las escuelas de todos los distritos, así como con los promedios nacionales. Tan sólo cuatro años después de la primera evaluación nacional, un observador externo²⁶ destacó que los profesores/as y directores/as de diversas escuelas públicas y privadas decían estar utilizando tareas e ítems similares a los de las pruebas nacionales para modelar sus propias evaluaciones de aula, y, posiblemente - dado que los ítems de las pruebas simulan buenas prácticas de enseñanza – también para modelar sus prácticas de aula.

Incluso cuando los profesores saben de la importancia de las evaluaciones formativas y sumativas, suele ser muy difícil para ellos desarrollar maneras innovadoras de supervisar y evaluar permanentemente el aprendizaje de sus alumnos. El desafío es incluso mayor cuando, como es el caso en América Latina, el aprendizaje memorístico de algunas habilidades y la acumulación de información han sido las principales características de las estrategias de

²⁵ Chile, 2003.

²⁶ Benveniste, 2000.

enseñanza, y cuando las actuales reformas curriculares y pedagógicas pretenden generar un cambio dramático hacia un enfoque que enfatiza la importancia de "aprender a aprender" y el desarrollo de habilidades complejas. Dada la escasez de oportunidades de capacitación y de fuentes de información pedagógica disponibles para muchos profesores/as en la región, esta clase de utilización de las pruebas puede ser una poderosa herramienta de la reforma - incluso más que los propios resultados en las pruebas. Hace poco, una comisión que revisó la experiencia **chilena** con el SIMCE, recomendó varias maneras a través de las cuales los profesores podrían familiarizarse mejor con los instrumentos de evaluación. Entre ellas se incluye la aplicación y análisis en el aula de pruebas similares a las aplicadas por la administración central, en otros grados que aquellos que son evaluados oficialmente.

Cabe mencionar que estudios recientes realizados en **Uruguay** sugieren que sólo una fracción de los profesores/as - principal público objetivo del sistema de evaluación uruguayo - ha leído o usado los informes emitidos por la agencia central de evaluación, y que muchos directores/as de escuela²⁷ no organizan sesiones especiales para discutir los resultados nacionales o institucionales que supuestamente deberían llegar a las escuelas durante el periodo de planificación que se da al final de cada año escolar. Tales porcentajes podrían bien considerarse un logro en comparación con lo que sucede en otros países. De otro lado, de aquellos que dicen haber leído los informes nacionales o institucionales, más del 60% dice haber introducido algunos cambios en sus planes y acciones tras esta lectura. Como se mencionó antes, los estudios sugieren que el mayor impacto se ha logrado a partir de la provisión de los instrumentos de evaluación para que las mismas escuelas los apliquen después de su aplicación en la muestra central.

Un caso tal vez más típico de otros países de América Latina es **Perú**, donde aún están pendientes mejores esfuerzos para comunicar los resultados a las escuelas y profesores. Algunos análisis interesantes de ítems han sido realizados y publicados tras los ciclos de evaluación de 1998 y 2001, pero probablemente han llegado sólo a un número reducido de escuelas y profesores, ya sea en forma impresa o por Internet. Se espera, sin embargo, que las nuevas evaluaciones - considerablemente mejoradas como consecuencia de los resultados en la prueba PISA - y el desarrollo de mayores y mejores esfuerzos de difusión, brindarán a los/as profesores/as señales más claras que otros instrumentos curriculares sobre el tipo de aprendizajes que sus prácticas de enseñanza y sus propias evaluaciones en aula deberían promover. La más reciente introducción de pruebas censales, con devolución de resultados a escuelas y padres de familia y con guías para su análisis dirigidas a los docentes son también canales que podrían ayudar a transmitir expectativas de aprendizaje con mayor claridad - aunque solo cubren una parte menor del currículo que lo que se puede hacer con pruebas muestrales y formas rotadas.

En **Brasil**, un número importante de profesores se ha capacitado en instituciones de nivel secundario y no cuenta con ninguna capacitación especializada. La prueba ENEM, diseñada principalmente para estudiantes que egresan de la secundaria, está siendo utilizada por los profesores para acceder a educación terciaria y, en algunos casos, a un fondo de becas vinculado a la prueba. Al parecer, el proceso mismo de preparación para la prueba - que es fuertemente interdisciplinaria y está "orientada a la realización de tareas del mundo real" - sirve como una oportunidad para que profesores no certificados reflexionen no sólo sobre su conocimiento de los contenidos y sobre sus competencias generales, sino también sobre sus formas de enseñar y evaluar a sus estudiantes. Los gobiernos locales de zonas aisladas donde el acceso a oportunidades de capacitación es limitado, están proporcionando cierta ayuda a los profesores que están pensando rendir la prueba, pues consideran que sus escuelas se beneficiarán de estas actividades preparatorias.

²⁷ 54.5, 62.9 y 75.6% de los directores y profesores entrevistados en tres diferentes áreas del país reportaron que no realizan ninguna actividad, formal o informal, ligada a la discusión de los resultados.

A pesar de haber realizado consultas con varias fuentes, no se ha encontrado ninguna información sistemática sobre maneras específicas en las cuales los resultados de las pruebas se estén utilizando para diseñar capacitaciones o programas de desarrollo profesional. De hecho, es muy posible que esté habiendo mucha actividad en este sentido, dado que los funcionarios ministeriales suelen contar la "capacitación de profesores" entre los usos dados a los resultados de las pruebas²⁸. Sin embargo, hay razones para creer que el contenido de muchas de tales actividades bien puede consistir sólo en un análisis superficial de los resultados de las pruebas nacionales y en una introducción general para los profesores sobre el sistema de evaluación.

La identificación y descripción de formas más promisorias en que los resultados de las pruebas puede ser utilizada como material para la capacitación docente es aún una tarea pendiente.

--o0o--

Conclusiones

Una primera conclusión de esta rápida revisión es que los resultados de las evaluaciones de logros de aprendizaje escolar parecen estar siendo utilizados más de lo que usualmente se sospecha. Dada la existencia relativamente corta de los sistemas de evaluación de la región y la inestabilidad del compromiso político con su desarrollo, no sorprende que algunos países utilicen los resultados en menor medida de lo que se esperaría. Sin embargo, si se toma en cuenta la seria escasez de recursos para el mejoramiento educativo y la existencia de muchos programas valiosos que valdría la pena financiar, parecería necesario redoblar los esfuerzos para dar un uso adecuado a los resultados de las evaluaciones.

Aunque la mayoría de países latinoamericanos no iniciaron el desarrollo de sus sistemas de evaluación con ideas claras sobre los usos que se podría dar a los resultados, una vez que empezaron a recoger información fueron encontrando diversas maneras de usarla. Así, por ejemplo, se la información ha sido utilizada para: el diseño e implementación de políticas; la asignación de recursos; para estimular cambios pedagógicos y promover la responsabilización entre varios actores e instituciones educativas. Ha habido además una considerable evolución en la forma de presentar los resultados a diversas audiencias, y cada vez se está dando mayor atención a la necesidad de incluir información sobre factores de contexto que deben ser tomados en cuenta al momento de interpretar los resultados. Ciertamente, también se han dado casos de usos no legítimos o válidos, especialmente para "responsabilizar" principalmente a los docentes y a sus gremios por los bajos logros de aprendizaje generalmente registrados.

Dado que en muchos lugares la prensa se ha convertido en el principal comunicador de los resultados a profesores/as, padres y tomadores de decisiones, es claro que no se ha realizado esfuerzos suficientes para educar a los medios de comunicación de modo que éstos se conviertan en aliados de los esfuerzos para movilizar al público en apoyo de la educación, en vez de actuar, conscientemente o no, como enemigos del sistema público. Esto requerirá que la mayoría de agencias de evaluación asigne una cantidad considerablemente mayor de recursos que los que se ha estado invirtiendo en esfuerzos de difusión (tiempo del personal para definir claramente los principales mensajes que deben ser transmitidos al público en general a través de la prensa; tiempo para responder a interpretaciones incorrectas o no deseables de los resultados de las pruebas; así como más dinero).

Aunque recién está empezando, el efecto de los resultados de las pruebas en la renovación curricular y el establecimiento de estándares también parece prometedor – cabe notar que muchos países pusieron en marcha sus sistemas de evaluación al mismo tiempo en que daban inicio a reformas curriculares cuya llegada a las aulas aún tenía que ser asegurada. Sin embargo, el diseño de las pruebas aún necesita mejorarse para permitir investigaciones adicionales de las implicancias pedagógicas de sus resultados. La posibilidad de utilizar los resultados para

²⁸ Ferrer 2005 (borrador)

mejorar la planificación a nivel de las escuelas también está emergiendo como una línea de acción interesante, y requerirá de aplicaciones de tipo censal o de la provisión de instrumentos para la autoevaluación y el análisis colectivo de los resultados.

Otra manera en que las evaluaciones pueden contribuir a mejorar el desempeño docente y los logros de aprendizaje, es la producción de materiales para el desarrollo profesional que estén basados en análisis detallados de los resultados de las pruebas - y la generación de espacios institucionales y eventos de capacitación en los cuales éstos puedan ser apropiados y usados a cabalidad. El análisis de ítems seleccionados de las pruebas que es difundido por las agencias de evaluación, con frecuencia ha facilitado entre los profesores la reflexión sobre sus propias expectativas para el aprendizaje de sus estudiantes, así como la reconsideración por parte de los especialistas de cuán adecuados son los objetivos de aprendizaje que han definido para varios grados. El acceso directo a los instrumentos de evaluación también parece una poderosa manera de aliviar los temores y prejuicios existentes sobre la gama supuestamente limitada de habilidades que las pruebas pueden evaluar. Hoy en día, incluso los ítems de opción múltiple pueden explorar procesos de razonamiento complejos y pueden servir para identificar fuentes de error ligadas a las prácticas de aula - algo que los opositores de las pruebas suelen negar.

Sin embargo, si se quiere asegurar el éxito de las estrategias de difusión y uso de los resultados, se tendrá que prestar mayor atención a las culturas organizacionales -tanto a nivel de las escuelas como en los niveles intermedios de la administración educativa. Las directrices tradicionales de tipo vertical (*top-down*) con frecuencia han fallado en tener los resultados anticipados cuando los intereses burocráticos y profesionales aparecen amenazados por la nueva información disponible. Ninguno de los aportes potenciales de las evaluaciones externas se podrá materializar si no hay terreno fértil en las instituciones que permita cultivar prácticas de evaluación interna y de planificación pedagógica que se nutran de los insumos que aquellas pueden proveer.

Referencias

- Arellano Ortiz, Fernando. "Panorama de la educación pública en Bogotá". Extraído el 10 de abril, 2005, de <http://clubdeanunciantes.com/Educacion%20en%20bogota.htm>
- Argentina. Ministerio de Cultura y Educación. Secretaría de Programación y Evaluación Educativa.
- 1997a **Evaluación y Uso de la información. Manual de Estrategias para el uso e incorporación de la información de evaluación.**
- 1997b **Recomendaciones Metodológicas para la Enseñanza. 3er operativo nacional de evaluación 1995.**
- 1995a **Recomendaciones Metodológicas para la Enseñanza. 2º. operativo nacional de evaluación 1994. Nivel Primario. 7º grado. Lengua**
- 1995b **Recomendaciones Metodológicas para la Enseñanza. 2º. operativo nacional de evaluación 1994. Nivel Medio. 5º año. Lengua.**
- 1994a **Recomendaciones Metodológicas para la Enseñanza. Lengua. Nivel primario.**
- 1994b **Recomendaciones Metodológicas para la Enseñanza. Lengua. Nivel medio.**
- Aristimuño, Adriana y Kaztman, Rubén. 2004. "¿Cómo se usa y qué impacto tiene la información empírica sobre las evaluaciones nacionales de aprendizajes en el mejoramiento de la Educación Primaria pública en Uruguay?" En Cueto, Santiago (ed) **Uso e impacto de la información educativa en América Latina**. Pp. 173-242 Santiago de Chile, PREAL <http://www.preal.cl/Libros/libro19/04-ARISTIMUNO.pdf>
- Arregui, Patricia. 2001. "Sistemas de determinación y evaluación de metas de logros de aprendizaje escolar como instrumento para mejorar la calidad, la equidad y la responsabilización en los procesos educativos en América Latina" en **Análisis de**

- Prospectivas de la Educación en América Latina y el Caribe.** Santiago de Chile, UNESCO, 2001
<http://www.preal.cl/GTEE/docr/Arregui%20-20Systems%20of%20determination.pdf>
- Arregui, Patricia. 2002. **Evaluación de impacto del proyecto MECES.** Informe de consultoría realizado para el Banco Mundial.
- Baeza, Jorge; Fuentes, Ricardo. “Antecedentes y fundamentos de las políticas de gestión y administración en el sistema educativo chileno”
<http://www.mineduc.cl/usuarios/mineduc/doc/200508021319210.antecednets%20y%20fundamentos.pdf>
- Benveniste, Luis. 2000. **La evaluación del rendimiento académico y la construcción de consensos en Uruguay.** Lima: Grupo de Trabajo sobre Estándares y Evaluación del PREAL
<http://www.preal.cl/GTEE/pdf/Benveniste.pdf>
- Bogotá. Alcaldía Mayor. Secretaría de Educación Distrital. Secretaría de Hacienda Distrital. s/f. **Bogotá piensa a futuro. Un marco de gasto de mediano plazo para el sector educativo 2004-2013.**
<http://www.sedbogota.edu.co/prensa/Archivos/bogotapiensaafuturo.pdf>
- Casa Editorial El Tiempo, Fundación Corona y Cámara de Comercio de Bogotá 2003. **Bogotá cómo vamos. Cambios en la calidad de vida de la Ciudad 2000-2002. Memorias del Foro Bogotá cómo vamos.**
- Bonamino, Alicia; Coscarelli, Carla & Franco, Creso. 2002. “Avaliação e letramento: concepções de aluno letrado subjacentes ao SAEB e ao PISA. **Educação & Sociedade**, Dez. vol. 23, no 81: 92-113.
http://www.scielo.php?script=sci_pef&pid=S0101-73302002008100006&lng=en&nrm=iso&tlng=pt
- Brasil. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Exame Nacional do Ensino Médio. Relatório Pedagógico 2001. Brasília-DF, dezembro de 2001**
http://www.inep.gov.br/download/enem/2001/relatorio_final/relatorio2001.pdf
- Centro de Perfeccionamiento Experimentación e Investigaciones Pedagógicas -CPEIP. Plan Complementario de Matemática 2004 orientado al mejoramiento de la educación matemática en segundo ciclo de enseñanza básica (5° a 8°). Santiago: CPEIP
<http://www.cpeip.cl/usuarios/pponce/doc/200408231752530.plan%20mate.doc>
- Chile 2003. Comisión para el Desarrollo y Uso del Sistema de Medición de la Calidad de la Educación. **Evaluación de Aprendizajes para una Educación de Calidad.** Santiago de Chile, Ministerio de Educación
<http://www.preal.org/GTEE/pdf/Simcedoc.pdf>
- Chile 2003a. Ministerio de Educación. Sistema de Medición de la Calidad de la Educación. Orientaciones para la Medición. 2°. Educación media.
http://www.simce.cl/doc/folletos_2003_2medio.pdf
- Chile 2003b. Ministerio de Educación. Sistema de Medición de la Calidad de la Educación. Carpeta Prensa.
http://www.simce.cl/doc/carpeta_prensa_2004.pdf
- Chile. 2004. Ministerio Secretaría General de la Presidencia. División de Coordinación Interministerial. **Gobierno en Marcha.** Logros 2000-2004. Reforma de la Educación.
http://www.gobiernodechile.cl/logros/reforma_educacion.asp
- Chile. 2005. Ministerio de Educación. “Estrategia de asesoría a la escuela para la implementación curricular en LEM: Campaña de Lectura, Escritura y Matemática”. Santiago: Ministerio de Educación de Chile. Extraído el 14 de abril, 2005 de:
<http://www.educared.edu.pe/modulo/upload/1>
- Colombia. Ministerio de Educación Nacional. **Revista Altablero** No. 24. September-October 2003.
- Cueto, Santiago (ed.) 2005. **Uso e impacto de la información educativa en América Latina.** Santiago de Chile, PREAL.

- Elacqua, Gregory y Rodrigo Fábrega. 2004. **El consumidor de la educación: el actor olvidado de la libre elección de colegios en Chile**. Santiago de Chile, Universidad Adolfo Ibáñez. Paper presentado en el Seminario Internacional sobre Accountability Educativa organizado por CIDE y PREAL en Santiago de Chile en abril del 2005.
http://www.uai.cl/p4_home/site/asocfile/ASOCFILE120041030211723.pdf
- Espínola, Viola. **Autonomía escolar: Factores que contribuyen a una escuela más efectiva**. BID. Departamento Regional de Operaciones, División de Programas Sociales, s/f.
- Fernández, Tabaré y Midaglia, Carmen. 2003. “¿Quiénes y cómo usan los informes generados por los sistemas de evaluación de aprendizaje en la educación primaria? Los casos de México y Uruguay”. En Cueto, Santiago (ed) **Uso e impacto de la información educativa en América Latina**. Pp.111-171. Santiago de Chile, PREAL.
<http://www.preal.cl/Libros/libro19/03-FERNANDEZ.pdf>
- Ferrer, J. Guillermo y Arregui, Patricia. 2003. **Las pruebas internacionales de aprendizaje en América Latina y su impacto en la calidad de la educación: criterio para guiar futuras aplicaciones**. Santiago de Chile, PREAL, Documento de Trabajo No. 26 http://www.preal.cl/docs-trabajo/Arregui_et_Al_26.pdf
- Ferrer, J. Guillermo. 2006. **Sistemas de Evaluación de Aprendizajes en América Latina. Balance y Desafíos**. Washington y Santiago, PREAL.
- Forster, Margaret; Valverde, Gilbert. 2003. **La experiencia internacional en sistemas de medición: estudio de casos**. Santiago: Comisión para el Desarrollo y Uso del Sistema de Medición de la Calidad de Chile.
http://www.simce.cl/doc/Comision_Simce_Anexo1_Valverde.pdf
http://www.simce.cl/doc/Comision_Simce_Anexo2_Forster.pdf
- García-Huidobro, Juan Eduardo, et al. 1999. **Educación para Todos: Evaluación en el Año 2000**. Informe de Chile. Santiago: Ministerio de Educación de Chile.
<http://www2.unesco.org/wef/countryreports/chile/contents.html>
- Gilmore, Alison. 2005. **The Impact of Pirls (2001) and Timss (2003) in Low- And Middle-Income Countries. An Evaluation of the Value of World Bank Support for International Surveys of Reading Literacy (Pirls) and Mathematics and Science (TIMSS)**. New Zealand: International Association for the Evaluation of Educational Achievement
http://www.iea.nl/iea/hq/fileadmin/user_upload/WB-report.pdf
- Holdzkom, David. 2001. “Low Performing Schools: So You’ve Identified Them—Now What?” West Virginia: AEL Policy Brief
<http://www.ael.org/rel/policy/pb0202.htm>
- Ingvarson, L., Meiers, M. & Beavis, A. 2005. Factors affecting the impact of professional development programs on teachers’ knowledge, practice, student outcomes & efficacy. En: **Education Policy Analysis Archives**, 13(10).
<http://epaa.asu.edu/epaa/v13n10/v13n10.pdf>
- Kellaghan, Thomas; Greaney, Vincent. 2001. **Using Assessment to Improve the Quality of Education**. Paris: UNESCO/International Institute for Educational Planning.
- Libertad y Desarrollo. 2005. “Calidad de la educación: análisis del fracaso en los resultados”. Santiago de Chile, ILD
<http://www.lyd.com/noticias/fracaso.ppt>
- Libertad y Desarrollo 2002. “Mejoremos la lectura, Escritura y Matemáticas” en **Temas Públicos** N° 576. Santiago de Chile, Libertad y Desarrollo
<http://www.lyd.com/biblioteca/pdf/576mejoremos.pdf>
- Mizala, Alejandra; Romaguera, Pilar. 1998, **Desempeño escolar y elección de colegios: la experiencia chilena**. Documento de trabajo N° 36, Serie Economía. Santiago: Centro de Economía Aplicada.
http://www.webmanager.cl/prontus_cea/cea_1998/site/asocfile/ASOCFILE12003040_2123414.pdf
- Mizala, Alejandra; González, Pablo; Guzmán, Andrea; Romaguera, Pilar. 2000. **Los Maestros en Chile: Carreras e Incentivos**. Research Network Working Paper #R-403. Washington: Banco Interamericano de Desarrollo.

- <http://www.iadb.org/res/publications/pubfiles/pubR-403.pdf>
- Mizala, Alejandra; Romaguera, Pilar. 2004. "Rendimiento escolar y premios por desempeño. La experiencia latinoamericana y el SNED en Chile" En Cueto, Santiago (ed) **Uso e impacto de la información educativa en América Latina**. Pp. 23-60. Santiago de Chile, PREAL.
<http://www.preal.cl/Libros/libro19/01-MIZALA.pdf>
- Moncada, German; Hernández Rodríguez, Russbel; Aguilar, Marina Chávez de Aguilar; Orellana, Dania; Alas Solís, Mario; Hernández, Bessy. 2004. "Uso e Impacto de la Información Empírica en la Formulación y Ejecución de Política de Educación Básica en Honduras en el Período 1990-2002" en Cueto, Santiago (ed) **Uso e impacto de la información educativa en América Latina**. Pp. 243-294. Santiago de Chile, PREAL.
<http://www.preal.cl/Libros/libro19/05-MONCADA.pdf>
- Montoya, Silvia; Cruz Perusia, Juan; Vera Mohorade, Alejandro. 2004. "Evaluación de la calidad educativa: De los sistemas centrales al aula. Estudio del impacto de las políticas de devolución de información personalizada a las escuelas en tres provincias de Argentina" En Cueto, Santiago (ed) **Uso e impacto de la información educativa en América Latina**. Pp. 295-343. Santiago de Chile, PREAL.
<http://www.preal.cl/Libros/libro19/06-MONTOYA.pdf>
- Odden, Allan and Kelley, Carolyn. **Paying Teachers for What They Know and Do. New and Smarter Compensation Strategies to Improve Schools**. California, Corwin Press, Inc.
- Organización de los Estados Americanos. Consejo Interamericano para el Desarrollo Integral. Secretaría Técnica de la Comisión Interamericana de Educación. 2004. **Cuestionario de Prácticas de Divulgación Educativa. Presentación preliminar de resultados**. Washington: CIDI.
<http://www.oas.org/udse/espanol/documentos/Documentoinformativo3esp.doc>
- Perú. Ministerio de Educación. Unidad de Medición de la Calidad.
- 2001a **Análisis de ítemes de las pruebas CRECER 1998. Resultados de lógico-matemática en cuarto grado de primaria**. Boletín 10.
http://www.minedu.gob.pe/gestion_institucional/of_planmedumc/indicadores/medicion/boletines/boletin10/Boletin_Crecer10.pdf
- 2001b **Análisis de ítemes de las pruebas CRECER 1998: Resultados de comunicación integral en cuarto grado de primaria**. Boletín 11.
http://www.minedu.gob.pe/gestion_institucional/of_planmedumc/indicadores/medicion/boletines/boletin11/
- 2001c **Análisis de ítemes de las pruebas CRECER 1998: Resultados de lógico-matemática en sexto grado de primaria**. Boletín 13.
http://www.minedu.gob.pe/gestion_institucional/of_planmedumc/indicadores/medicion/boletines/boletin13/
- 2001d **Análisis de ítemes de las pruebas CRECER 1998: Resultados de comunicación integral en sexto grado de primaria**. Boletín 14.
http://www.minedu.gob.pe/gestion_institucional/of_planmedumc/indicadores/medicion/boletines/boletin14/
- 2001e **Análisis de ítemes de las pruebas CRECER 1998: Resultados de comunicación integral en cuarto grado de secundaria**. Boletín 17.
http://www.minedu.gob.pe/gestion_institucional/of_planmedumc/indicadores/medicion/boletines/boletin17.pdf
- 2002 **Análisis de ítemes de las pruebas CRECER 1998: Resultados de matemática en quinto grado de secundaria**. Boletín 18.
http://www.minedu.gob.pe/gestion_institucional/of_planmedumc/indicadores/medicion/boletines/boletin18.pdf
- 2003a **Cómo rinden los estudiantes peruanos en Comunicación y Matemática: Resultados de la Evaluación Nacional 2001. Cuarto grado de secundaria. Informe pedagógico**. Documento de trabajo UMC 4.

- http://www.minedu.gob.pe/umc/2004/publicaciones_mecep/doctrab/Documento%20e%20Trabajo%2004.pdf
- 2003b **Cómo rinden los estudiantes peruanos en Comunicación – Producción de Textos: resultados de la evaluación nacional 2001. Cuarto y sexto grados de primaria. Cuarto grado de secundaria. Informe pedagógico.** Documento de trabajo UMC 5 http://www.minedu.gob.pe/umc/2004/publicaciones_mecep/doctrab/Documento%20e%20Trabajo%2005.pdf
- Pérez Martínez, Ángel. Educación en Bogotá Período 1992 – 2002. Extraído el 20 de abril, 2005 de:
<http://univerciudad.redbogota.com/bajar-pdf/educacion.pdf>
- Pinkasz, Daniel; Montes, Nancy; Marcalain, Gabriela; Legarralde, Martín; Dussel, Inés; Tiramonti, Guillermina. 2004. “Los usos de la información empírica en el sistema educativo. Un estudio de caso en tres jurisdicciones de la Argentina” En Cueto, Santiago (ed) **Uso e impacto de la información educativa en América Latina.** Pp. 345-394. Santiago de Chile, PREAL.
<http://www.preal.cl/Libros/libro19/07-PINKASZ.pdf>
- Raczynski, Dagmar. **Estudio de evaluación de la jornada escolar completa.** Santiago, Pontificia Universidad Católica de Chile. Dirección de Estudios Sociales.
http://biblioteca.mineduc.cl/documento/informe_jec_UC.pdf
- Ravela, Pedro (ed); Wolfe, Richard; Valverde, Gilbert; Esquivel, Juan Manuel. 2001a. **Los Próximos Pasos: ¿Cómo avanzar en la evaluación de aprendizajes en América Latina?** Santiago de Chile, PREAL, Documento de Trabajo No. 20.
<http://www.preal.cl/doctra20.pdf>
- Ravela, Pedro. 2002. **¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina?.** Santiago de Chile, PREAL Documento de Trabajo No. 2 http://www.preal.org/docs-trabajo/ravela_n22.pdf
- Ravela, Pedro. 2003. **¿Cómo aparecen los resultados de las evaluaciones educativas en la prensa?.** Grupo de Trabajo sobre Estándares y Evaluación del PREAL.
<http://www.preal.cl/GTEE/pdf/prensa.pdf>
- N.N. 2003. “La Revolución llega al aula de clase” en **Altablero.** No. 24 setiembre-octubre. Extraído de
<http://www.mineducacion.gov.co/altablero/articulo.asp?sec=3&id=153&num=24>
- N.N. 2002. “Los mejores colegios” en **Dinero** No. 151: 12-17. Bogotá.
<http://www.dinero.com/>
Schmidt, William; Houang, Richard; Cogan, Leland. 2002. A Coherent Curriculum: The Case of Mathematics. En: **American Educator**, Summer.
http://www.aft.org/pubs-reports/american_educator/summer2002/curriculum.pdf
- Tiramonti, Guillermina; Dussell, Inés; Pinkasz, Daniel; Marcalain, Gabriela; Montes, Nancy. **Un estudio de caso de tres jurisdicciones de la Argentina: grado de conocimiento de la información disponible y su utilidad en los diferentes niveles de gestión del nivel medio de la enseñanza pública.** Buenos Aires, FLACSO
<http://www.preal.cl/FIE/pdf/publicaciones/InformeFinalFlacso.PDF>
- UNESCO Chile. Innovemos. Ficha de registro del Programa de las 900 Escuelas.
<http://innovemos.unesco.cl/red/bgi/?f=F>
- Uruguay—Administración Nacional de Educación Pública. 1996. “Primeros Resultados de la Evaluación Nacional de Aprendizajes en 6to Año de Enseñanza Primaria”. Nota de Prensa. Montevideo: Consejo Directivo Central, mimeo.
- Vega, Emiliana; Umansky, Ilana. 2005. **Mejorar la enseñanza y el aprendizaje por incentivos. ¿Qué lecciones nos entregan las reformas educativas de América Latina?** Washington, D.C., Banco Internacional de Reconstrucción y Desarrollo.
<http://www.preal.org/docs-trabajo/Vegas%20y%20Umansky.pdf>
- Villamil Sivey, Blanca. 2003. Informe Escolar (“State Report Card”) Año Escolar 2002-2003. **Informe sobre el aprovechamiento de las Escuelas Públicas del Estado Libre Asociado de Puerto Rico.** . Departamento de Educación de Puerto Rico

Zibas, Dagmar. 2002. "A reforma do ensino médio no Chile: vitrina pa a América Latina?"
Cadernos de Pesquisa 115: 233-262.