



Programa de Promoción de la Reforma Educativa en América Latina y el Caribe  
Partnership for Educational Revitalization in the Americas

**N° 11**

## **Las Evaluaciones Educativas en América Latina: Avance Actual y Futuros Desafíos**

**Laurence Wolff\***

JULIO 1998.-

(\*) Laurence Wolff es asesor en educación del Banco Interamericano de Desarrollo en Washington, D.C. Anteriormente trabajó durante 22 años en el Banco Mundial, donde participó en proyectos educacionales en América Latina, Asia y el Medio Oriente.

La traducción al español fue realizada por Maritza Blajtrach Roldán.



# INDICE

<b>I. Introducción</b>	<b>4</b>
<hr/>	
<b>II. Recientes avances en las evaluaciones educacionales internacionales y nacionales</b>	<b>6</b>
• Programas internacionales	<b>6</b>
• La experiencia de Estados Unidos	<b>8</b>
• La experiencia francesa	<b>11</b>
• Los países en desarrollo	<b>12</b>
<hr/>	
<b>III. La experiencia latinoamericana</b>	<b>13</b>
• Chile	<b>13</b>
• Costa Rica	<b>14</b>
• Colombia	<b>16</b>
• Brasil	<b>18</b>
• México	<b>19</b>
• Argentina	<b>20</b>
• Programa de Evaluación Regional UNESCO/OREALC	<b>21</b>
<hr/>	
<b>IV. Lecciones aprendidas y futuros desafíos</b>	<b>22</b>
• Resumen de la experiencia en seis países	<b>22</b>
• Generación de consenso y compromiso	<b>23</b>
• Usos de las evaluaciones	<b>24</b>
• Creación de capacidades y competencia técnica	<b>25</b>
• Nuevos enfoques	<b>26</b>
<hr/>	
<b>Bibliografía</b>	<b>27</b>
<hr/>	
<b>Anexo I. Una mirada a los sistemas de evaluación de seis países</b>	<b>29</b>

## I. INTRODUCCION

La educación es fundamental para el desarrollo económico, social y cultural, como asimismo para la estabilidad política, la identidad nacional y la cohesión social. Además, las empresas de hoy, que utilizan tecnología de vanguardia, no pueden avanzar sin personas capaces de incorporar habilidades analíticas, creativas y de cooperación al lugar de trabajo. La disponibilidad de trabajadores con estas características también puede tener un gran impacto sobre la capacidad de un país para atraer inversión extranjera. Los logros económicos de los países del Este de Asia, sólo por dar un ejemplo, pueden atribuirse en gran medida a la calidad y nivel superiores de la educación en la región, la que actualmente se vanagloria de exhibir cuatro de los cinco más altos récords mundiales en matemáticas de octavo grado.

En los últimos años, las naciones del mundo han convenido en la importancia de medir el rendimiento educacional. Estiman que la evaluación de los actuales niveles de logro y la identificación de los obstáculos al progreso les permitirá mejorar el tipo, la profundidad y el alcance de la educación que ofrecen. Las evaluaciones educacionales pueden definirse como medidas del grado en el cual se han logrado los objetivos curriculares, ya sean establecidos por las autoridades gubernamentales o por expertos nacionales e internacionales. Las evaluaciones nacionales miden el avance de las instituciones en todo el país. Difieren significativamente de los exámenes finales o de ingreso destinados a seleccionar a los alumnos para el siguiente nivel de educación. Las evaluaciones internacionales comparan los logros en el aprendizaje entre diversos países. Es obvio que el mero hecho de medir el aprendizaje de los alumnos no aumentará su rendimiento al igual que el hecho de pesar el grano no aumenta la producción agrícola. Sin embargo, es una condición necesaria para establecer metas cuantitativas, evaluando las ventajas comparativas de las estrategias de asignación de recursos y tanto

como combinaciones de insumos, la asignación de recursos y los esfuerzos con miras al logro de las metas establecidas. Para cerciorarse de que las evaluaciones educacionales no sólo mejoren el aprendizaje de los alumnos, sino que también sean eficaces en función de los costos, se requiere de compromiso, financiamiento adecuado, conocimientos técnicos, experiencia administrativa y habilidad política.

Las evaluaciones educacionales son sólo un medio para supervisar el avance en el logro de los objetivos. También es importante evaluar los resultados cuantitativos (por ejemplo, el número de alumnos matriculados, de alumnos que han completado un nivel determinado o de alumnos que han sido promovidos), la suficiencia de los insumos, tales como los textos escolares, los profesores, los coeficientes profesor-alumno y la capacitación docente; las interacciones y la pedagogía en la sala de clases y el desempeño en el mercado laboral (por ejemplo, cuántos egresados obtienen trabajo, con qué sueldo, etc.).

América Latina no es ajena a las evaluaciones educacionales. A fines de los años setenta, el Programa de Estudios Conjuntos para la Integración Económica Latinoamericana (ECIEL) completó un estudio comparativo del aprendizaje en cinco países, utilizando los instrumentos desarrollados por la Asociación Internacional para la Evaluación del Rendimiento de los Alumnos (IEA). Diversos países de América Latina han participado en evaluaciones internacionales patrocinadas por la IEA como asimismo por el Servicio de Pruebas Educativas (Educational Testing Service - ETS). Sin embargo, los resultados de estas evaluaciones han sido decepcionantes; los puntajes obtenidos por los países de América Latina han sido significativamente más bajos que aquéllos de los países de América del Norte, Europa y Asia. En el estudio de 1989 de la IEA (Cuadro1), Venezuela obtuvo un puntaje más bajo que todos los demás países participantes, incluida Indonesia. Trinidad y Tobago obtuvo un puntaje algo mejor que Venezuela, pero aún muy

por debajo de los países desarrollados. En un estudio de matemáticas de 1991 (Cuadro2) implementado por el ETS, las ciudades de Fortaleza y São Paulo, Brasil, obtuvieron un puntaje más bajo que el de todos los países y ciudades participantes, con excepción de Mozambique. Como se observa más adelante en este informe, Colombia, el único país latinoamericano participante, también obtuvo un puntaje bajo en el más reciente estudio de matemáticas de la IEA.

<b>Cuadro 1.</b>		
<b>Estudio de la Habilidad de Lectura de la IEA, 1989</b>		
<b>Países de América Latina y el Caribe</b>	<b>Media</b>	<b>Desviación Standard</b>
Trinidad y Tobago	451	79
Venezuela	383	74
<b>Otros países</b>		
Finlandia	569	70
Estados Unidos	547	74
Hong Kong	517	71
Singapur	515	72
España	504	78
Alemania Occidental	503	84
Indonesia	394	59
Fuente: W.B. Elley, <i>How in the World do Students Read?</i> (Newark, Delaware: International Reading Association, 1992).		

<b>Cuadro 2.</b>		
<b>Evaluación Internacional del Progreso Educacional Prueba de Matemáticas y Ciencias, 1991</b>		
<b>Prueba de Matemáticas</b>		
<b>País</b>	<b>Media</b>	<b>Desviación Standard</b>
Corea	73	0,6
Taiwán	73	0,7
España	55	0,8
Estados Unidos	55	1,0
Portugal	48	0,8
Sao Paulo, Brasil	37	0,8
Fortaleza, Brasil	32	0,6
Maputo y Beira, Mozambique	28	0,3
<b>Prueba de Ciencias</b>		
<b>País</b>	<b>Media</b>	<b>Desviación Standard</b>
Corea	78	0,5
Taiwán	76	0,4
España	68	0,6
Estados Unidos	67	1,0
Portugal	63	0,8
Sao Paulo, Brasil	53	0,6
Fortaleza, Brasil	46	0,6
Fuente: Educational Testing Service, <i>The International Assessment of Educational Progress: Learning Mathematics</i> (Princeton, N.J., ETS, 1992).		

En 1991, Costa Rica, México, Chile y Colombia contaban con sistemas de evaluación. Desde entonces, casi todos los países latinoamericanos han dado inicio a algún tipo de programa. A través de un aporte del BID, la UNESCO ha apoyado un programa regional de evaluación para los alumnos de tercer y cuarto grados en lectura y matemáticas. Además, se ha incrementado la profundidad de las evaluaciones internacionales como asimismo el número de países que participan en ellas. En este informe se analizan los avances en las evaluaciones educacionales, tanto a nivel internacional como en seis países latinoamericanos; también se comentan algunas posibles nuevas orientaciones en cuanto a políticas y programas de evaluación en la región. El análisis se basa en estudios internacionales y nacionales recientes (identificados en la bibliografía), incluidos diversos estudios preparados para una conferencia de PREAL en Río de Janeiro en diciembre de 1996.

## **II. RECIENTES AVANCES EN LAS EVALUACIONES EDUCACIONALES INTERNACIONALES Y NACIONALES**

### **Programas internacionales**

El acontecimiento más importante relacionado con las evaluaciones internacionales en los últimos años ha sido el Tercer Estudio Internacional de Matemáticas y Ciencias (Third International Mathematics and Science Survey - TIMSS) de la IEA. La IEA es un consorcio de evaluación que opera en todo el mundo a través de su oficina central en Amsterdam, Holanda. Es conocida por sus estudios internacionales de programas de matemáticas, ciencias, lectura, alfabetización y ciencias sociales. Los programas de la IEA son financiados por los países participantes.

El TIMSS tiene como objetivo medir, comparar y explicar el aprendizaje en ciencias y matemáticas en 41 países. Actualmente se dispone de los resultados de los exámenes de matemáticas y ciencias para niños de los grados cuarto, séptimo, octavo, y decimosegundo. Además, recientemente, el TIMSS llevó a cabo un innovador análisis sobre la "oportunidad de aprender" de los niños, que

clasifica en categorías y compara los currículos, los textos escolares y la pedagogía en la sala de clases entre los 41 países participantes. Dichos análisis demuestran que los exámenes de rendimiento no deben simplemente concentrarse en la identificación de los países que obtienen los puntajes más altos o más bajos, sino que pueden ser utilizados como una herramienta para medir el progreso educacional de un país, redefinir sus objetivos curriculares y modificar las prácticas en la sala de clases.

En 1991, once países latinoamericanos (Costa Rica, Perú, Argentina, República Dominicana, Colombia, Guatemala, Venezuela, Chile, Ecuador, Brasil y México) asistieron a una reunión regional preliminar del TIMSS, pero sólo Colombia y México participaron hasta el final. Además, sólo se informaron los puntajes de Colombia en la publicación del TIMSS. En el último minuto, México decidió no permitir la divulgación de sus puntajes. Argentina y República Dominicana participaron en el análisis de los currículos. La razón de la baja participación de los países de América Latina (y de todos los países en desarrollo) parece ser una combinación de falta de recursos técnicos y financieros adecuados como, tal vez, la preocupación de algunas autoridades ante la posibilidad de obtener los peores puntajes. En el Cuadro 3 se presentan los puntajes obtenidos en las pruebas de rendimiento de matemáticas de octavo grado del TIMSS. Entre los 41 países que informaron sus puntajes, Colombia obtuvo el puntaje más bajo de todos los países, con excepción de Sudáfrica. Los datos muestran que sólo el 5% de los estudiantes colombianos obtuvieron un puntaje ubicado entre el 50% superior de los estudiantes del mundo. La amplia mayoría de los participantes fueron países desarrollados con un ingreso per cápita mucho más alto que el de Colombia y la mayoría de ellos gastaba significativamente más en su sistema educacional. A su vez, países con un ingreso per cápita similar al de Colombia - específicamente, Bulgaria, Lituania, Filipinas, Rumania, Latvia, Irán, Eslovaquia, Rusia y Tailandia - obtuvieron puntajes más altos que dicho país.

**Rendimiento en matemáticas en el  
Tercer Estudio Internacional de Matemáticas y Ciencias  
Porcentaje de alumnos que lograron los niveles de referencia internacionales, 8° grado**

<b>País</b>	<b>Rendimiento en matemáticas, 8° grado</b>	<b>Nivel del 10% superior</b>	<b>Nivel de la mitad superior</b>	<b>Diferencias de rendimiento entre el 7° y el 8° grado</b>
Singapur	643	45	94	42
Corea	607	34	82	30
Japón	605	32	83	34
Hong Kong	588	27	80	24
Bélgica (Fl)	565	17	73	8
República Checa	564	18	70	40
República Eslovaca	547	12	64	39
Suiza	545	11	65	40
Austria	539	11	61	30
Francia	538	7	63	46
Hungría	537	11	60	35
Eslovenia	541	11	61	43
Federación Rusa	535	10	60	35
Holanda	541	10	63	25
Bulgaria	540	6	58	26
Canadá	527	7	58	33
Irlanda	527	9	57	28
Bélgica (Fr)	526	6	58	19
Austria	530	11	57	32
Tailandia	522	7	54	28
Israel	522	6	56	-
Suecia	519	5	53	41
Alemania	509	6	49	25
Nueva Zelanda	508	6	48	36
Inglaterra	506	7	48	30
Noruega	503	4	46	43
Estados Unidos	500	5	45	24
Dinamarca	502	4	47	37
Escocia	498	5	44	36
Latvia (LSS)	493	3	40	32
España	487	2	36	39
Islandia	487	1	37	27
Grecia	484	3	37	44
Rumania	482	3	36	27
Lituania	477	1	34	49
Chipre	474	2	34	28
Portugal	454	0	19	31
Irán,				
República Islámica	428	0	9	27
Kuwait	392	0	3	—
Colombia	385	0	4	16
Sudáfrica	384	0	2	7

Fuente: Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS) (Chestnut Hill, Mass. Boston College / TIMSS International Study Center, 1996).

La metodología del TIMSS permite que quienes desarrollan los currículos y quienes diseñan las políticas determinen si existe correspondencia entre los currículos, los textos escolares y la enseñanza y aprendizaje en la sala de clases. De este modo, los detalles de los resultados obtenidos por Colombia en una diversidad de temas de matemáticas, como se muestra en el Cuadro 4, permiten formarse una buena idea de las fortalezas y debilidades del currículum en el área de las matemáticas. Así, Colombia puede utilizar sus resultados en cuanto a rendimiento para comparar la calidad de sus objetivos educacionales, currículos y textos escolares con aquéllos de otras naciones y, de este modo, identificar las áreas que deben ser mejoradas.

Cuadro 4.	
¿Cómo varía el rendimiento en Colombia de acuerdo con el currículum que se pretende aplicar y el currículum realmente implementado?	
Contenido	Escala internacional
Perímetro, área y volumen	1
Patrones, relaciones y funciones	41
Geometría: transformaciones	41
Números reales, sus propiedades y subcategorías *	12
Números racionales y sus propiedades*	12
Fracciones y decimales	40
Ecuaciones y fórmulas*	40
* Enfatizados en los textos escolares colombianos y por los profesores.	
Fuente: Variations in Mathematics Performance in Colombia in Accordance with the Intended and Implemented Curriculum," presentación de Gilbert Valverde, Las políticas educacionales nacionales y la participación en estudios internacionales," reunión de PREAL, Río de Janeiro, diciembre de 1996.	

A pesar de que el principal programa del TIMSS está próximo a ser concluido, ya se encuentran en curso una serie de actividades de seguimiento. Los investigadores han comenzado a establecer "puntos de referencia" para los currículos y pruebas actuales de los países sobre la base del examen del TIMSS y sus posteriores análisis. Además, ya se ha llegado al acuerdo de repetir el TIMSS en 1999.

### La experiencia de Estados Unidos

Estados Unidos se encuentra a la vanguardia en el desarrollo de una amplia variedad de evalua-

ciones educacionales destinadas a medir el desempeño institucional, establecer niveles de competencia mínimos para el egreso de la escuela y establecer puntos de referencia para medir el rendimiento educacional en comparación con los estándares internacionales. El Recuadro 1, que presenta información acerca de cuatro programas de evaluación que se aplican en el Condado de Montgomery, Maryland, muestra la extensión, la complejidad y, tal vez, los excesos de los programas de evaluación desarrollados en Estados Unidos. El primer programa consiste en una prueba desarrollada por el condado, que entrega a los padres información acerca de la medida en que



sus hijos que cursan los grados tercero a octavo dominan los principales elementos del currículum a nivel del grado correspondiente. Sirve como una herramienta de diagnóstico que permite identificar las debilidades y fortalezas de un alumno e incluye preguntas de selección múltiple, preguntas de interpretación abierta y preguntas basadas en el rendimiento. El segundo programa, el Programa de Evaluación del Rendimiento Escolar de Maryland, evalúa la medida en que cada escuela cumple con los estándares de rendimiento establecidos por el estado para los grados tercero, quinto y octavo. Utiliza principalmente preguntas de selección múltiple o sólo pide respuestas breves. El tercer programa de evaluación, las Pruebas Funcionales de Maryland, consiste en un "examen de competencia mínima" en lectura, matemáticas, redacción y estudios cívicos, que debe ser aprobado por los alumnos para recibir un certificado válido de educación secundaria. El cuarto programa de evaluación consiste en las Pruebas Globales de Habilidades Básicas, recientemente lanzado y que se aplicará a una pequeña muestra de alumnos que cursan los grados segundo, cuarto y sexto. Servirá como un indicador de los niveles de logro de las habilidades básicas en relación con las normas nacionales e internacionales.

En virtud de su historia y de sus leyes, y en contraste con la mayoría de los países, Estados Unidos no cuenta con un currículum obligatorio a nivel nacional. Y dado que la mayoría de los estados sólo entregan pautas generales, cada distrito escolar tiene considerable libertad para establecer el currículum. De acuerdo con un reciente informe de 1997 del Centro Nacional de Investigaciones de Estados Unidos (U.S. National Research Center), los currículos de los estados generalmente no están definidos adecuadamente y no se los aplica a nivel local. Como resultado, lo que se enseña realmente en la sala de clases varía enormemente de una escuela a otra y de un distrito a otro, y los alumnos y los profesores a menudo tienen una noción exagerada del nivel de logro que obtienen. Debido a estos problemas, las

autoridades han hecho un gran esfuerzo por establecer estándares más sólidos en cuanto a aprendizaje y logros a nivel de los estados, como asimismo estándares nacionales posiblemente voluntarios. Por otra parte, varios estados - Maryland entre ellos - están intentando basar sus currículos y pruebas en los estándares internacionales, especialmente en relación con el TIMSS. Cabe señalar que las dos principales asociaciones gremiales de profesores, especialmente la Federación Norteamericana de Profesores (American Federation of Teachers - AFT), han apoyado permanentemente los esfuerzos a nivel nacional y estatal tendientes al mejoramiento de la evaluación y la calidad de la educación.

**Las evaluaciones educacionales en Estados Unidos**  
Programa de evaluaciones del Condado de Montgomery, Maryland

El sistema de escuelas públicas del Condado de Maryland, una zona suburbana cercana a Washington, D.C., utiliza una amplia batería de pruebas para medir el rendimiento de sus estudiantes. La siguiente descripción del programa de evaluaciones del condado ha sido extraída de un volante informativo enviado a los padres de los niños matriculados en las escuelas del mismo.

**Evaluaciones del Condado de Montgomery**

**Pruebas de Criterio (CRT).** Las Pruebas de Criterio (CRT), desarrolladas por las Escuelas Públicas del Condado de Montgomery, son aplicadas cada primavera a los alumnos que cursan el tercer y octavo grado con el fin de medir la calidad del aprendizaje y el grado en que están aplicando las habilidades enseñadas en los currículos de matemáticas y lectura/idioma/arte.

Además del formato tradicional basado en la selección múltiple para evaluar los conocimientos de los contenidos concretos, las pruebas de criterio incluyen también preguntas de interpretación abierta y preguntas de desarrollo. Estas preguntas reflejan el compromiso, tanto a nivel local como del estado, con respecto a instrumentos más modernos no basados en la selección múltiple que evalúan habilidades más ricas y complejas, tales como la habilidad de comunicarse a través de la escritura y de resolver problemas de la vida diaria.

En el otoño, los padres reciben un informe individualizado que muestra el rendimiento de sus niños en las preguntas de selección múltiple de las pruebas de criterio y ofrece una comparación de dicho rendimiento con los puntajes a nivel de la escuela y del condado. El informe incluye los puntajes de años anteriores para ayudar a estimar el progreso del alumno y de la escuela, identificar las fortalezas y debilidades, y evaluar el efecto de cualquier medida tomada con el fin de mejorar el rendimiento durante el año.

Los planes actuales contemplan el uso de una escala única que combina el rendimiento de los alumnos en las

preguntas de selección múltiple e interpretación abierta para informar a las escuelas y a los padres lo que los alumnos saben y lo que son capaces de hacer.

**Las evaluaciones del Estado de Maryland**

**El Programa de Evaluación del Rendimiento Escolar de Maryland (MSPAP).** Estas pruebas utilizan preguntas de respuesta breve y actividades complementarias para determinar si las escuelas han cumplido con los estándares de rendimiento del estado en las áreas de lectura, redacción, uso del lenguaje, matemáticas, ciencias y ciencias sociales. Los alumnos de tercero, quinto y octavo grado rinden estas pruebas cada primavera. En la actualidad, no se informan los puntajes individuales. Los resultados de las escuelas se publican en el mes de enero.

**Las Pruebas Funcionales de Maryland.** Estas pruebas evalúan las habilidades de competencia mínimas en lectura, matemáticas, redacción y estudios cívicos. La aprobación de estas pruebas es un requisito para recibir el diploma de educación secundaria de Maryland. Los alumnos dan una prueba de lectura y matemáticas en el séptimo grado, de redacción en el octavo grado y de estudios cívicos en el décimo grado. Los alumnos tienen varias oportunidades de aprobar estas pruebas durante el año. Se informan los puntajes individuales obtenidos por los alumnos en cada prueba.

**Las Pruebas Globales de Habilidades Básicas (CTBS).** Maryland utiliza esta prueba como un indicador de los niveles de logro de las habilidades básicas en lectura, idioma y matemáticas en relación con las normas nacionales. A partir de la primavera de 1977, Maryland aplicará esta prueba en los grados segundo, cuarto y sexto cada año. Las Escuelas del Condado de Montgomery administrarán esta prueba a una muestra de 250 alumnos por grado. Los puntajes de cada alumno serán informados a los padres en el otoño.

Fuente: "Celebrating Knowledge through Assessment," Escuelas Públicas del Condado de Montgomery (mimeo), 1996.

## La experiencia francesa

Desde mediados de los años ochenta, en Francia se ha desarrollado un sofisticado sistema de evaluación de los alumnos. Si bien Francia cuenta con un currículum obligatorio a nivel nacional, las escuelas elementales y secundarias tienen un alto grado de libertad en el uso de sus fondos y sus técnicas pedagógicas. La responsabilidad por la supervisión y evaluación de la educación está en manos de los más de 200 técnicos de la Direction de l'Evaluation et de la Prospective (DEP), una división del Ministerio de Educación de Francia.

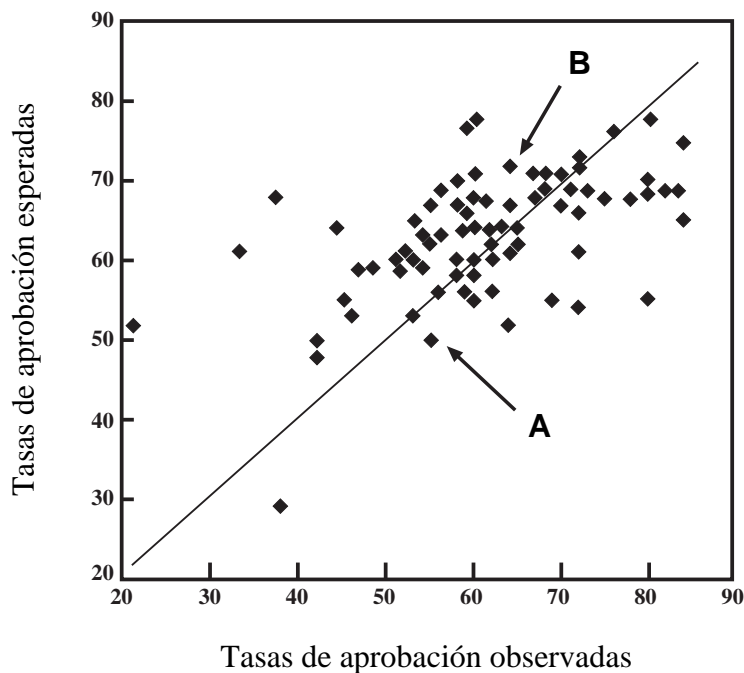
Las principales autoridades de gobierno están firmemente comprometidas con el establecimiento de un sistema de medición de los resultados educacionales basado en estándares claros y en una difusión completa y abierta de los resultados, asegurando de este modo la independencia de la DEP.

Según el sistema francés, los estándares de logro esperados se establecen de manera centralizada y se toman pruebas anuales a todos los alumnos de grados específicos de la educación primaria y secundaria del nivel inferior. El examen de bachillerato (egreso) administrado a los alumnos del nivel superior de la educación secundaria también se utiliza con propósitos de evaluación. La DEP proporciona retroalimentación detallada y fácil de comprender a todas las partes involucradas: la opinión pública, los padres, los alumnos, los profesores y administradores y el

Ministro de Educación. El organismo ha desarrollado también una amplia variedad de materiales de enseñanza y aprendizaje opcionales orientados al rendimiento de los alumnos y a las fortalezas y debilidades de los profesores. Al igual que diversos organismos en Estados Unidos, la DEP ha intentado medir el rendimiento escolar en comparación con los estándares internacionales.

Como ocurre en todo el mundo, las escuelas de las comunidades más pobres de Francia obtienen puntajes más bajos que las escuelas que se encuentran en las áreas de mayor riqueza. La DEP ha intentado identificar las escuelas que se encuentran en áreas más pobres y están exhibiendo un buen rendimiento con el fin de definir de mejor manera el "valor agregado" de la escolaridad. En el Gráfico 1, se presenta un método estadístico de medición del desempeño de las escuelas en relación con el nivel socioeconómico de sus alumnos. Las escuelas que muestran un rendimiento mejor de lo esperado en base a su población escolar se definen como escuelas "eficaces". Estudiando a dichas escuelas, las autoridades debieran ser capaces de determinar por qué son eficaces y luego establecer condiciones similares en las escuelas que presentan un rendimiento más bajo. Este tipo de análisis ya se está llevando a cabo en algunas partes de Estados Unidos (véase, por ejemplo, una descripción del "valor agregado" en el sistema escolar de Dallas, según lo descrito en Alvarez y Ruiz-Casares, 1997).

**Medición del valor agregado en la educación: un ejemplo de Francia**  
Las escuelas “eficaces” producen resultados mejores que lo esperado de una determinada población de alumnos



Fuente: Adaptado de “L’état de l’école,” n° 5, octubre de 1995. Direction de l’Evaluation et de la Prospective (DEP), Ministère de l’Education Nationale, de l’Enseignement Supérieur, de la Recherche, et de l’Insertion Professionnelle (MEN), pág. 51. Los datos utilizados para preparar esta ilustración no corresponden necesariamente a los resultados reales de las pruebas.

En este gráfico se presenta un método utilizado por el gobierno francés para medir el desempeño de las escuelas en relación con el nivel socioeconómico de sus alumnos. Cada punto representa una escuela y el diagrama muestra las tasas de aprobación reales (observadas) y esperadas en el examen de bachillerato. Todos los liceos que están ubicados bajo la línea ofrecen a sus alumnos mayores probabilidades de aprobar el examen comparado con lo que se podría esperar en función de la edad y el origen social de los alumnos.

La escuela A ha obtenido valor agregado en términos de éxito en el examen, debido a que su tasa observada es más alta que su tasa esperada (está bajo la línea). Este no es el caso de la escuela B, que no ha alcanzado su nivel de éxito esperado, a pesar de que su tasa de aprobación observada es realmente más alta que la de la escuela A.

### Los países en desarrollo

Los países en desarrollo también están haciendo esfuerzos por establecer sistemas de evaluación educacional a nivel nacional. Aparte de

América Latina, se han descrito y analizado sistemas de evaluación en Tailandia, Egipto, Corea del Sur y Jordania. A pesar de ser un país pequeño y relativamente pobre, Jordania ha desarrollado un programa de evaluación sistemático y

técnicamente competente. Su éxito en esta área puede explicarse por las siguientes razones: (a) la participación directa de las más altas autoridades (por ejemplo, el príncipe heredero) y su compromiso de largo plazo con el mejoramiento de la calidad de la educación y la información abierta de los resultados; (b) el establecimiento de un organismo independiente y bien financiado, aparte del Ministerio de Educación, responsable de las evaluaciones y (c) un fuerte liderazgo técnico, unido a asistencia del exterior. En 1991, Jordania llevó a cabo varias series de pruebas basadas en las evaluaciones internacionales y desarrolladas por el Servicio de Pruebas Educativas (la prueba "IAEP") y actualmente está proporcionando asistencia técnica en esta área a los países vecinos. Una serie de evaluaciones aplicadas recientemente identificó significativos mejoramientos en el aprendizaje de los alumnos como resultado de los esfuerzos de reforma de los currículos escolares y la capacitación de los profesores que se vienen realizando desde hace una década en Jordania.

### **III. LA EXPERIENCIA LATINOAMERICANA**

Con anterioridad a 1991, Chile y México eran los únicos países de la región que contaban con bastante experiencia en el área de las evaluaciones. El programa de evaluación de Chile estaba siendo aplicado desde 1980 y, a pesar de que México contaba con un programa de evaluación desde hacía varios años, las autoridades hicieron poco o nada por difundir los resultados. Costa Rica llevó a cabo evaluaciones nacionales entre 1986 y 1990. Colombia contaba con un sistema de pruebas nacionales de larga data utilizado principalmente para la selección de los postulantes a la educación superior.

En los años siguientes, otros países latinoamericanos comenzaron también a avanzar en esta dirección. Hacia 1996, su experiencia se había incrementado considerablemente y en la actualidad prácticamente todos los países de América Latina han iniciado algún tipo de programa de

evaluación. En este informe se extraen conclusiones a partir de los progresos hechos en Chile, Costa Rica, México, Argentina, Brasil y Colombia. También se describe el programa regional UNESCO/OREALC.

### **Chile**

Como ya se ha mencionado, Chile ha estado involucrado desde hace largo tiempo en el desarrollo de evaluaciones educacionales. Su programa fue concebido en 1978, cuando el Ministerio de Educación solicitó a la Pontificia Universidad Católica que diseñara e implementara un sistema de información para la educación. En 1988, con la transferencia de las escuelas públicas a las municipalidades, el programa recibió el nombre de Sistema de Medición de la Calidad de la Educación (SIMCE).

La función del SIMCE es ayudar al Ministerio de Educación y a las autoridades regionales y provinciales a supervisar el sistema de educación, a evaluar a las escuelas y a apoyar la capacitación docente en el servicio. El programa evalúa a los niños que cursan los grados cuarto y octavo en las áreas de español y matemáticas y al 10% de ellos en ciencias naturales, historia y geografía. También evalúa el desarrollo personal y las actitudes de los alumnos, las actitudes y antecedentes de los profesores y los padres, y la eficacia de las escuelas. Las evaluaciones de ambos grados se aplican en años alternados. A partir de 1991, el Ministerio de Educación se hizo cargo de la administración del programa.

A contar de 1988, el programa se ha vuelto más eficaz y eficiente, luego de la introducción de mejoramientos en la capacidad técnica, el sistema de procesamiento computarizado y la administración. Actualmente, los puntajes se entregan a las escuelas en forma mucho más rápida y se han simplificado los informes para asegurar la fácil comprensión de los resultados. Estos esfuerzos parecen estar teniendo impacto en la planificación pedagógica en muchas escuelas. Quienes

diseñan los currículos y los materiales pedagógicos, por ejemplo, están dando énfasis al dominio de los objetivos que parecen ocasionar problemas a los alumnos. Algunos padres están utilizando los resultados del SIMCE para seleccionar para sus hijos las escuelas que muestran un mejor rendimiento. El costo del programa es de aproximadamente \$5 dólares por alumno, lo que es comparable con los estándares internacionales.

Los puntajes en el SIMCE también han constituido la base de un programa pedagógico dirigido inicialmente a 900 y posteriormente a 1.200 de las escuelas chilenas que exhiben el peor rendimiento. Las escuelas que participan en este programa han recibido materiales pedagógicos, bibliotecas, libros, infraestructura y capacitación docente en servicio. Dependiendo de su condición inicial, se da preferencia a algunas escuelas en el otorgamiento de subvenciones para programas de mejoramiento locales. Por otra parte, las escuelas que han mejorado sus puntajes de un año al otro han recibido recompensas financieras.

Las pruebas del SIMCE revelan lo siguiente: aquellas escuelas que tienen niños que provienen de familias pobres y con un bajo nivel de educación o de áreas rurales obtienen los peores puntajes en las pruebas del SIMCE; las escuelas municipales públicas y las escuelas rurales obtienen puntajes más bajos que las escuelas privadas, especialmente aquellas instituciones con una larga tradición; las escuelas privadas muestran un rendimiento mejor que las escuelas públicas aun después de controlar el factor del nivel socioeconómico de los padres y, finalmente, los puntajes de las 900 escuelas muestran un mejoramiento significativo durante los últimos años. Entre los actuales problemas se encuentra el hecho de que, si bien muchas escuelas están haciendo uso de los resultados del SIMCE con el fin de mejorar las condiciones locales, aún se espera que sean las autoridades centrales las que inicien las acciones correctivas. Algunas escuelas han informando un aumento del número de alumnos provenientes de hogares en situación desmedrada, en

un aparente esfuerzo por mostrar que sus logros relativos (valor agregado) han mejorado. Las mediciones del área afectiva no han tenido éxito y quizás debieran ser descartadas. Tal vez también sería apropiado reducir la proporción de pruebas al universo y recurrir más al muestreo. Finalmente, debido a problemas técnicos, la comparabilidad de los resultados de un año a otro es inadecuada.

En general, Chile cuenta en la actualidad con el sistema de evaluación más amplio y mejor administrado de América Latina y el SIMCE ha servido como una poderosa herramienta para la implementación de un programa de reforma que promueve la descentralización, la responsabilidad por los resultados y el aumento del aprendizaje y la escolaridad. Las autoridades de gobierno buscan mejorar aun más el sistema. En particular, actualmente están considerando agregar pruebas de desarrollo en las evaluaciones (hoy en día todas las pruebas son del tipo selección múltiple). Dado que aún no se comprenden cabalmente las causas del bajo rendimiento, se encuentra en marcha un esfuerzo de investigación más sofisticado. Finalmente, las autoridades chilenas han decidido participar en el estudio TIMSS de seguimiento programado para 1999 como una manera de establecer puntos de referencia para la comparación del rendimiento de sus alumnos con el resto del mundo.

## **Costa Rica**

Costa Rica ha estado involucrada en las evaluaciones educacionales desde 1986. Su actividad en esta área ha pasado por tres etapas. Durante la primera etapa - entre 1986 y 1990 - los objetivos planteados fueron medir el grado de aprendizaje de los conceptos básicos por parte de los niños y jóvenes, estimular a los padres y profesores a utilizar el tiempo de enseñanza en forma más eficaz, promover una discusión nacional acerca de la calidad de la educación, destacar el hecho de que todos los costarricenses son responsables del mejoramiento de la calidad de la

educación y demostrar la importancia del restablecimiento de los exámenes de certificación nacionales al final de la educación secundaria. El programa fue formulado en su mayor parte por el Instituto de Investigación para el Mejoramiento de la Educación Costarricense (IIMEC), una institución autónoma de la Universidad de Costa Rica. Inicialmente, el programa introdujo pruebas de español y matemáticas para todos los alumnos de los grados tercero, sexto, noveno, undécimo y duodécimo. En 1987 y 1988, las pruebas se ampliaron cubriendo ciencias físicas, ciencias sociales, inglés y francés. En 1988, se estableció un examen de egreso de la escuela secundaria. Este primer programa generó una controversia pública, debido al bajo nivel de rendimiento de los alumnos y se volvió a las pruebas globales objetivas.

Durante el período 1988 a 1993, la concentración se dirigió principalmente al examen de egreso de la escuela secundaria. Las pruebas, preparadas por el IIMEC y aplicadas por el Ministerio de Educación, representaban el 60% de la nota final y los puntajes se asignaron a nivel local. Cuando un estudio de 1991 indicó que el 30% de las pruebas habían sido evaluadas de manera incorrecta, casi en todos los casos en favor de los alumnos, las autoridades decidieron eliminar la mayoría de las preguntas de desarrollo y asignar el puntaje de casi todas las pruebas de manera centralizada utilizando lectores ópticos. Además, los exámenes de sexto grado, preparados por las autoridades regionales, representaban el 50% de la nota final, pero no se tomaron las medidas suficientes como para asegurar la confiabilidad y la validez de los resultados informados por las autoridades regionales.

A contar de 1993, bajo un nuevo gobierno y Ministro de Educación, Costa Rica se comprometió firmemente con las evaluaciones y recibió ayuda en esta área a través de un proyecto financiado por el Banco Mundial. Se volvieron a establecer las pruebas de diagnóstico para los grados tercero, sexto y noveno en cuatro materias,

se efectuaron otros tipos de evaluaciones y se fortaleció el IIMEC mediante nuevo equipo y personal. El nuevo sistema consiste en la aplicación de una diversidad de pruebas y evaluaciones, la mayoría de las cuales son preparadas por el IIMEC.

Entre las evaluaciones formativas se incluyen las siguientes:

- 1) pruebas de rendimiento diagnósticas para tercer grado en las materias básicas. Estas fueron pre-evaluadas en 1995 y aplicadas a una muestra amplia en 1996.
- 2) evaluaciones diagnósticas iniciales, aplicadas al 10% de los niños que ingresan al primer grado para obtener información acerca de su condición física, cognitiva y socio-emocional. Los resultados serán utilizados como una manera de establecer pautas para la entrega de experiencias de aprendizaje apropiadas en la educación preescolar y en primer grado.
- 3) pruebas de habilidades de resolución de problemas, aplicadas en 1996 a una muestra de alumnos del noveno grado. El objetivo de estas pruebas era medir las capacidades cognitivas en esta área como asimismo las variables socioeconómicas y académicas que influyen en el rendimiento.
- 4) evaluaciones de la capacidad física, para medir la aptitud física de una muestra nacional de niños que cursaban los grados tercero, sexto y noveno.

También se han desarrollado diversas pruebas sumativas. Se aplica un examen de egreso preparado a nivel nacional a todos los alumnos que terminan el noveno grado en todas las materias básicas. Este examen consiste en preguntas de selección múltiple cuyo puntaje se asigna por medio de lectores ópticos, junto con preguntas de desarrollo que son evaluadas por profesores especialmente capacitados. Los resultados representan el 25% de la nota del segundo semestre. A contar de 1988, en Costa Rica también se ha aplicado un examen de egreso de la educación

secundaria, preparado y administrado a nivel regional. Representa el 60% de la calificación final y consta tanto de preguntas objetivas como de desarrollo. Además, las oficinas regionales del Ministerio de Educación preparan y administran exámenes de egreso de sexto grado, los que, sin embargo, difieren en calidad, validez y confiabilidad.

En la etapa de prueba inicial (1988), los niños mostraron un rendimiento muy por debajo de las expectativas del currículo nacional. Las escuelas privadas urbanas obtuvieron los mejores puntajes en estos exámenes. Si bien el IIMEC distribuyó información resumida a cada escuela, las instituciones no hicieron ningún intento explícito de utilizar los resultados. Sin embargo, una encuesta hecha a los profesores en 1989 mostró que el 70% de éstos tenían conocimientos acerca del material y aproximadamente el 35% de ellos utilizaban el material de una manera u otra en la sala de clases.

En la segunda etapa, la información acerca de los resultados de los exámenes de egreso de la educación secundaria fue entregada a algunos alumnos como también a algunas escuelas secundarias. El uso de dicha información quedaba a completa discreción de la escuela.

Se espera que la tercera etapa refleje un enfoque más sistemático en lo que respecta a la divulgación, utilización y retroalimentación. Se prepararán informes para los diferentes niveles y se solicitará a las unidades educacionales regionales que incorporen los resultados de las evaluaciones en sus planes educacionales. También se hará el intento de medir el grado en el cual ha mejorado la educación. A pesar de estos avances, los padres no reciben ninguna otra información aparte de los puntajes de los alumnos, no se emiten comunicados de prensa especiales sobre educación, los resultados no se utilizan para preparar materiales específicos destinados a corregir las debilidades en el aprendizaje y las instituciones de capacitación no han recibido ninguna pauta ni informa-

ción específicas acerca de las pruebas. Sin embargo, se entregarán informes a las escuelas, especialmente a nivel de primer grado y educación preescolar, junto con sugerencias para enriquecer el currículo.

En general, en Costa Rica los esfuerzos de evaluación se han visto dificultados por un enfoque "intermitente". Si bien el actual Ministro de Educación apoya decididamente el programa, algunos educadores y líderes políticos no lo hacen, y no está claro si esta situación continuará bajo la nueva administración, que asumió en mayo de 1998. Dado su reducido tamaño, es posible que Costa Rica haya emprendido un programa demasiado ambicioso. Quizás sería más productivo reducir el número de pruebas y concentrar los esfuerzos en la divulgación y utilización de los resultados. El examen secundario de alta exigencia de Costa Rica, que representa el 25% de la nota final como un medio de aumentar el rendimiento en este nivel puede ser de interés para otros países latinoamericanos que deseen aumentar el aprendizaje.

## **Colombia**

En 1986, el gobierno de Colombia estableció al interior del Ministerio de Educación un departamento responsable de la evaluación de las instituciones y programas educacionales. Luego, en 1990, puso en marcha un sistema nacional de evaluación con el propósito de construir un modelo de evaluación operacional que permitiera tomar decisiones tendientes al mejoramiento de la calidad. Además, se esperaba que el sistema evaluara los conocimientos de los alumnos y determinara de qué manera estaban siendo motivados por los profesores, las escuelas y los materiales pedagógicos, generara un mayor sentido de la importancia de la escolaridad, comunicara este sentido a toda la sociedad, proporcionara el marco teórico para los conceptos tales como la calidad y la evaluación del rendimiento educacional y apoyara las investigaciones que permitieran al país lograr estos objetivos. El programa fue llevado a



cabo por un equipo inter-institucional, conformado por profesionales del Ministerio de Educación, la Universidad Pedagógica Nacional, el Centro de Estudios Sociales y el Instituto SER.

A partir de 1968, el Instituto Colombiano para el Mejoramiento de la Educación Superior (ICFES) ha sido responsable de la formulación de un examen nacional de ingreso a las instituciones de educación superior de la nación. En 1980, estos exámenes pasaron a ser obligatorios y, de ahí en adelante, fueron aplicados por el Servicio Nacional de Pruebas del ICFES. El programa, que se centra en las pruebas y el ingreso de los alumnos a las instituciones de educación superior más que a la evaluación del sistema educacional en su conjunto o sus subsistemas, ha ayudado a crear una competencia técnica en el área de la administración de pruebas.

A diferencia de lo que se hace en Chile y Costa Rica, en Colombia se da énfasis al muestreo y a la investigación como una manera de identificar las causas del bajo rendimiento. Por lo tanto, no se entrega información a cada padre, alumno o escuela relativo al rendimiento observado. Colombia está considerando actualmente ampliar la muestra considerada con el fin de entregar retroalimentación a un mayor número de municipalidades y escuelas.

El sistema colombiano comenzó con una prueba a los alumnos de tercer y quinto grado en las áreas de matemáticas y español. Posteriormente, se desarrollaron pruebas en las áreas de ciencias naturales y ciencias sociales para los grados séptimo y noveno. Estas pruebas medían explícitamente habilidades de razonamiento de orden superior tales como el uso de algoritmos y las habilidades de resolución de problemas en matemáticas o la habilidad de extraer el significado a partir de material escrito en español.

Cuando se miden los logros cognitivos de los alumnos de tercer y quinto grado en comparación con el currículum nacional, los resultados son

decepcionantes. Pocos alumnos son capaces de utilizar las operaciones básicas requeridas para resolver problemas concretos y son pocos los que comprenden cabalmente lo que leen, ya sea en un sentido crítico o reflexivo. Los mejores puntajes se obtienen en las escuelas urbanas y privadas y en algunas regiones. Incluso un estudio que controlaba el factor socioeconómico, sugirió que el hecho de asistir a una escuela privada otorga una ventaja en términos de logros. El rendimiento fue más alto en aquellos alumnos que habían asistido a la educación preescolar, no habían repetido, no tenían muchas ausencias, tenían libros en sus hogares y cuyos padres tenían un mayor nivel de educación. Si se contaba con textos escolares y se los utilizaba, el rendimiento era más alto. Los alumnos que tenían profesores mejor capacitados, más textos escolares, docentes de sexo femenino y asistían a escuelas completas también exhibían un buen rendimiento. Las pruebas confirmaron también que los alumnos de la “Escuela Nueva”, un innovador programa para las pequeñas escuelas rurales, mostraban un rendimiento significativamente mejor que los alumnos de las escuelas rurales que no participaban en el programa.

Las conclusiones de estas evaluaciones han influido en los objetivos y el contenido de la nueva ley de educación, aprobada en 1994. Sin embargo, la información obtenida a partir de las actuales evaluaciones basadas en una muestra presta poca utilidad a las municipalidades y escuelas específicas, que se han vuelto cada vez más responsables de sus escuelas bajo el nuevo sistema de educación descentralizada. Si bien las actuales autoridades estiman que sería demasiado caro ampliar la evaluación a todos los alumnos, Colombia pretende trabajar con una muestra más amplia y espera establecer una base de datos y preguntas que pueda ser utilizada directamente por las escuelas y las autoridades educacionales. Como se mencionara anteriormente, Colombia es el único país latinoamericano que participó completamente en el programa del TIMSS. Si bien los resultados fueron decepcionantes, cabe aplaudir

la voluntad de Colombia de aceptar el escrutinio externo. Actualmente, Colombia está intentando utilizar los resultados del TIMSS para reforzar el sistema de evaluación nacional y para reformar el currículum.

El programa de evaluación colombiano ha incluido un fuerte elemento de investigación que ha influido en las políticas globales, específicamente en la redacción de la Ley de Educación de 1994, y ha confirmado el éxito de la "Escuela Nueva". Sin embargo, los resultados aún no han sido utilizados sistemáticamente para mejorar el rendimiento de las escuelas mismas ni para reformar el currículum. En particular, el currículum de matemáticas y ciencias requiere una reforma. Al mismo tiempo, la asociación entre las instituciones públicas y privadas ha funcionado bien.

## **Brasil**

Brasil comenzó tardíamente el desarrollo de herramientas de evaluación a nivel nacional. En 1990, se estableció el Sistema Nacional de Evaluación de la Educación Básica (SAEB) y sólo se comenzó a tomar muestras en 1993 y 1995. Uno de los objetivos del SAEB ha sido estimular a los estados y municipalidades a iniciar sus propias evaluaciones. En particular, los estados de Paraná y Minas Gerais han iniciado programas de evaluación recientemente.

En 1995, el SAEB anunció una serie de innovaciones. Su estudio incluiría tanto la educación secundaria como las instituciones privadas. También adoptó métodos de medición más sofisticados, introdujo instrumentos que entregarían información acerca de los antecedentes de los alumnos y redujo el tiempo de demora de la publicación de los resultados. El estudio de 1995 se centró en los grados cuarto y octavo de la educación básica y en el tercer grado de la educación secundaria.

Los resultados del estudio, publicados en 1995, indicaron que se evaluaron 90.499 alumnos de

cuarto y octavo grado de la educación básica y segundo y tercer grado de la educación secundaria. Esta muestra de alumnos fue tomada de 2.289 escuelas públicas y 511 escuelas privadas. Las preguntas del estudio se basaron en el nivel de aprendizaje esperado por los profesores y especialistas en educación. En el área de lectura, se evaluó el rendimiento de los alumnos en comprensión, extensión y examen crítico del significado. En el área de las matemáticas, el estudio se centró en tres categorías: comprensión de los conceptos, comprensión y aplicación de los procedimientos, y resolución de problemas. En el análisis estadístico, la principal tarea fue determinar los niveles de rendimiento esperados y reales de los niños en diversos grados.

En la prueba de 1995, los niños de todo el país obtuvieron un puntaje significativamente inferior a los niveles que los profesores y especialistas esperaban de ellos. En matemáticas, sólo el 21% de los estudiantes que cursaban el cuarto grado obtuvieron un puntaje sobre el nivel esperado, sólo el 15% de los alumnos de octavo grado obtuvieron un puntaje igual o superior al nivel esperado y sólo el 4% de los estudiantes secundarios obtuvieron un porcentaje superior al nivel esperado. En idioma, los puntajes fueron aun más bajos. Sólo el 22% de los alumnos de cuarto grado, el 14% de los alumnos de octavo grado y el 1% de los alumnos de educación secundaria obtuvieron puntajes superiores al nivel esperado. Los puntajes más altos se observaron en el sur, el sudeste y el centro-oeste del país. En el norte y el noreste se observaron los puntajes más bajos. Los alumnos de las ciudades principales obtuvieron mejores puntajes que aquéllos del interior. Los niños que tenían padres con mayor nivel de educación y aquéllos que asistían a escuelas privadas también obtuvieron puntajes más altos. Los estudiantes que asistían a las escuelas nocturnas obtuvieron puntajes inferiores. Los alumnos de mayor edad obtuvieron puntajes más bajos que los alumnos más jóvenes. Los estudiantes blancos y orientales obtuvieron puntajes más altos que aquéllos de raza mixta o negros. Y los alumnos con

profesores mejor capacitados obtuvieron puntajes más altos que aquéllos con profesores con menor capacitación. Estos resultados tienen importantes implicaciones en el área de las políticas educacionales, pero las autoridades recién están comenzando a tomarlas en consideración. Aun más, todavía no han incorporado técnicas más sofisticadas en sus análisis, tales como el control del factor socioeconómico en los alumnos que ingresan.

Recientemente, en el estado de Sao Paulo se inició un amplio programa de evaluación. Se evalúa a todos los alumnos de grados seleccionados de la educación básica y del nivel inferior de la educación secundaria en las áreas de matemáticas e idiomas, y se envían los resultados a los padres, profesores y escuelas. Se entrega un análisis de las preguntas en las cuales los alumnos han obtenido el peor puntaje. Se incluyen “preguntas ancla” para asegurar la compatibilidad de los resultados de un año al otro. El programa de Sao Paulo se aplica principalmente bajo contrato con organismos de pruebas privados sin fines de lucro. También parece ser que el programa de evaluaciones de Minas Gerais está avanzando en la utilización de sus resultados, ya que se están informando los puntajes individuales a las escuelas.

De manera bastante independiente de lo anterior, en virtud de una ley aprobada en 1995, el Ministerio de Educación desarrolló un sistema para evaluar las instituciones de educación superior. El propósito de este sistema es informar a los estudiantes y a la sociedad en general acerca de la calidad de las instituciones de educación superior. Las escuelas que obtengan un bajo puntaje tendrán que diseñar programas para fortalecer la calidad, los que serían apoyados por el gobierno federal.

En resumen, el programa de evaluación brasileño todavía se encuentra en sus primeras etapas. Las autoridades brasileñas están recién comenzando a considerar de qué manera podrían utili-

zarse las evaluaciones para mejorar las políticas educacionales y el currículum. Existe la necesidad de llevar a cabo análisis más rigurosos para identificar los factores que afectan el rendimiento. Dado que la educación brasileña es altamente descentralizada y no se rige por un currículum nacional, el gobierno está considerando establecer algunos estándares nacionales voluntarios como asimismo una prueba que permita medir si éstos han sido alcanzados. El programa de evaluación de la educación superior, el primero de este tipo en la región, constituye una importante innovación.

## **México**

En 1970, México estableció una oficina en la Unidad de Planificación Educacional de la Secretaría de Educación Pública, que finalmente se transformó en la Subdirección de Evaluación y Acreditación, con el propósito de estudiar las características y la calidad del sistema educacional del país. Los profesionales involucrados llevaron a cabo estudios de aptitud en niños de sexto grado de la educación básica y establecieron un examen de ingreso a la educación secundaria.

Durante el período 1976 a 1982, la subdirección investigó el aprendizaje en una muestra representativa de alumnos que cursaban el cuarto y quinto grado. Los resultados de esta evaluación aparecieron en publicaciones científicas y especializadas, no siendo divulgadas de ninguna otra manera, y las autoridades les prestaron poca atención. En efecto, la información con respecto a la evaluación se transformó en un “secreto de estado” sólo conocido por un reducido número de profesionales de la secretaría. Este enfoque dificultó el desarrollo técnico como asimismo la utilización para el diseño de políticas.

Durante el período 1983 a 1988, en México se desarrolló un examen para egresados de las escuelas de capacitación de profesores. Posteriormente, en 1989, se tomó la decisión de aplicar el

concepto de evaluación de manera más amplia con el fin de mejorar la enseñanza y el aprendizaje y de publicar los resultados. En 1992, el gobierno federal y las Asociaciones Gremiales Nacionales de Profesores acordaron un programa tendiente a modernizar la educación básica a través de su descentralización a nivel de los estados, dejando en manos del gobierno federal la responsabilidad de medir y evaluar el aprendizaje y asegurar la calidad de la educación básica y la capacitación de los profesores. Con este propósito, la Secretaría de Educación Pública se comprometió a apoyar las evaluaciones de los profesores, las evaluaciones en la sala de clases y aquéllas realizadas a nivel nacional.

En 1994, después de cinco años de evaluaciones de la calidad de la educación en México, la Secretaría emitió un informe acerca de los conocimientos y las habilidades de 480.000 profesores como asimismo el rendimiento de 2,8 millones de niños de educación básica y secundaria. Sus principales conclusiones fueron que los niños que habían asistido a la educación preescolar obtenían puntajes más altos que aquéllos que no lo habían hecho; los niños que habían repetido el sexto grado como asimismo aquéllos que estaban trabajando exhibían un menor rendimiento que aquéllos que nunca habían repetido y aquéllos que no trabajaban; los niños que asistían a escuelas urbanas o privadas obtenían mucho mejores resultados que aquéllos que asistían a escuelas rurales y públicas; los puntajes más bajos se encontraban entre los niños que asistían a escuelas indígenas y comunitarias, que contaban con instalaciones inadecuadas y profesores con menor nivel de capacitación, y los niños que obtenían los mejores puntajes asistían a escuelas urbanas y tenían padres que exhibían un mayor nivel de educación. Mientras que los niños de primer y segundo grado obtenían puntajes cercanos a lo que los investigadores y diseñadores de currículos esperaban, sus puntajes bajaban sucesivamente cada año en términos del porcentaje de respuestas correctas. Las autoridades mexicanas informaron también que era imposible medir sistemáticamente el

desempeño de los profesores en la sala de clases, debido a que las poblaciones de alumnos eran extremadamente diversas y aún existían dificultades técnicas.

Como se mencionara anteriormente, México participó cabalmente en el TIMSS, pero en el último momento decidió no publicar los resultados. Esta decisión no debiera sorprender si se considera la renuencia general de México a divulgar los resultados de los exámenes en el pasado, tratándose en ocasiones como secretos de estado. Si bien las actitudes han cambiado desde entonces y actualmente se divulgan los resultados de las evaluaciones correspondientes a la escuela básica y secundaria, esto aún se hace con alguna ambivalencia. La característica más distintiva del sistema de evaluaciones de México es que ha evaluado sistemáticamente los conocimientos y la capacidad de los profesores.

### **Argentina**

Al igual que Brasil, Argentina comenzó a realizar evaluaciones en forma tardía, pero ha avanzado con rapidez en el establecimiento y la utilización de las evaluaciones para mejorar la calidad educacional. De acuerdo con una nueva ley de educación federal que descentralizó la educación, el Ministerio de Educación estableció en 1993 un sistema de evaluación nacional con oficinas en la Secretaría de Programación y Evaluación Educacional. Los objetivos del sistema de evaluación eran promover la descentralización, proporcionar información clave acerca de la situación de la educación, supervisar el avance en el logro de los objetivos de la reforma, identificar las desigualdades e insuficiencias como asimismo las áreas en las que se requieren programas compensatorios para las poblaciones en situación desventajosa y estimular la participación de sectores más amplios en la toma de decisiones en el área de la educación. A partir de 1993, en Argentina se han administrado pruebas a todos los niños que cursan el último año de la educación primaria y secundaria tanto en idioma como en

matemáticas. En 1995, las evaluaciones se ampliaron cubriendo el tercer grado de la educación primaria y el segundo grado de la educación secundaria. También se comenzó a administrar pruebas en las áreas de las ciencias sociales y las ciencias naturales. Se aplicaron cuestionarios complementarios a los profesores, directores, alumnos y padres. Las pruebas se basaron en un análisis detallado de los logros curriculares esperados y un cuidadoso programa de pruebas piloto. Se confirmó la confiabilidad y la validez de estas pruebas a través de análisis estadísticos y revisiones por parte de expertos de la medida en la cual las preguntas correspondían a los objetivos curriculares. Todo el programa es administrado por el Ministerio de Educación. Los resultados de las evaluaciones se incorporan en un sistema de información educacional más amplio, que se utiliza para controlar y supervisar el sistema educativo en todos los niveles.

Entre la información relativa a las evaluaciones se incluyen detalles acerca del rendimiento escolar y los perfiles de los alumnos, profesores y modelos administrativos asociados al rendimiento. Esta información es dada a conocer a las autoridades nacionales, estatales, locales y escolares. Sin embargo, los puntajes individuales de cada escuela no se informan a las escuelas mismas. Diversos informes han abordado también el tema de los contenidos de las materias e incluyen recomendaciones tendientes a introducir mejoramientos en la pedagogía, tanto a nivel central como a nivel de cada escuela. Los resultados de las pruebas también han sido utilizados para desarrollar manuales orientados al mejoramiento del aprendizaje y la capacitación en el servicio, y estas pautas han sido entregadas, junto con asistencia técnica, a los 24 estados del país.

Si bien las preguntas de las pruebas fueron elaboradas en base a los niveles de respuesta mínimos esperados, en promedio los alumnos respondieron sólo el 50% de las preguntas correctamente. Los factores escolares representan el 40% de la varianza asociada a los puntajes en matemáticas y el 28% de la varianza en los puntajes

en idioma. Los niños de nivel socioeconómico más alto obtuvieron los mejores puntajes. Sin embargo, las diferencias entre las escuelas fueron mayores que las diferencias al interior de las escuelas, lo que sugiere que el sistema educacional argentino está altamente segmentado. Los niños pobres que asistían a clases heterogéneas obtuvieron puntajes más altos que los niños pobres de clases homogéneas. En general, los resultados muestran que se requiere introducir mejoramientos significativos en las escuelas.

El programa de evaluación de Argentina, si bien se inició sólo en 1993, está bien concebido y tiene una estrecha relación con su estrategia de descentralización y mejoramiento de la calidad de la educación en general. Al igual que Chile, Argentina está realizando todos los esfuerzos posibles por utilizar los resultados de las evaluaciones. Ya se han distribuido ampliamente informes nacionales y regionales, muchos de los cuales abordan los problemas más significativos y entregan materiales pedagógicos destinados a fortalecer la enseñanza.

### **Programa de Evaluación Regional UNESCO/OREALC**

A través de un subsidio del BID, la Oficina Regional de la UNESCO (Oficina Regional para América Latina y el Caribe, OREALC) se encuentra implementando un programa de evaluación regional bajo la denominación de "Laboratorio Latinoamericano de Medición de la Calidad de la Educación". El programa, iniciado en 1995, busca fomentar la coordinación y fortalecer la capacidad de los organismos de evaluación nacionales en la región. En 1997, se administraron pruebas de matemáticas y lectura, acompañadas de detallados cuestionarios de antecedentes, a una muestra de alumnos de tercer y cuarto grado de 15 países de América Latina. Las pruebas fueron desarrolladas por un comité en el que se encontraban representados todos los países participantes y se basan en un análisis informal de los objetivos curriculares en el área de las matemáticas y la lectura en la región. Los resultados se darán a conocer aproximadamente a mediados de 1998.

En 1993, la UNESCO implementó una prueba similar, en calidad de piloto, para alumnos de cuarto grado de siete países latinoamericanos. La mayoría de estos alumnos obtuvieron un puntaje muy por debajo de los niveles mínimos esperados por los expertos regionales en el área de los currículos y las pruebas. En promedio, los niños pudieron responder correctamente alrededor de la mitad de las preguntas. Venezuela y Costa Rica obtuvieron los mejores puntajes, en tanto que República Dominicana, Bolivia y Ecuador obtuvieron los puntajes más bajos. En el Cuadro 5, se entrega un resumen de los resultados.

Cuadro 5.		
<b>Puntajes obtenidos en los exámenes regionales de cuarto grado en matemáticas e idiomas, 1994</b> Porcentaje de alumnos que respondieron las tres cuartas partes (o más) de las preguntas correctamente		
País	Matemáticas	Idiomas
Venezuela	15,5	25,2
Costa Rica	10,4	23,8
Chile	5,0	23,2
Argentina	1,3	18,5
República Dominicana	1,4	5,0
Bolivia	2,1	4,9
Ecuador	0,4	7,7
Región	5,5	15,8

Fuente: UNESCO/OREALC, "Medición de la Calidad de la Educación: Resultados" (Santiago, Chile, 1994).

La UNESCO tiene proyectado encargar un estudio que compare el grado de dificultad y el contenido curricular de la prueba de matemáticas de la UNESCO con la prueba de matemáticas del TIMSS. Esto entregará evidencia directa del nivel en que se encuentran los países latinoamericanos en comparación con los países de fuera de la región. También se espera que, con el apoyo de la UNESCO, los expertos del TIMSS ayuden a países seleccionados de América Latina a implementar la metodología del TIMSS para medir la "oportunidad de aprender", como por ejemplo, la revisión y comparación del currículum oficial, el currículum aplicado, los textos escolares y la pedagogía utilizada en la sala de clases.

El programa de la UNESCO constituye un paso importante en el logro de una coordinación

regional. Para asegurar el fortalecimiento a largo plazo de las capacidades regionales, los esfuerzos futuros requerirán el apoyo a los centros regionales de excelencia, el establecimiento de relaciones con las instituciones públicas y privadas, y el mantenimiento de vínculos más fuertes que en el pasado con la IEA y los principales centros mundiales de investigación y desarrollo.

#### IV. LECCIONES APRENDIDAS Y FUTUROS DESAFÍOS

##### Resumen de la experiencia en seis países

Chile y Argentina cuentan con los sistemas de evaluación más amplios y mejor administrados de América Latina. Chile ha demostrado también el mayor compromiso de largo plazo con el desarro-

llo de evaluaciones. Si bien Chile y Argentina han hecho los mayores avances en la utilización de las evaluaciones en el área del diseño de las políticas, la reforma curricular y el mejoramiento de las escuelas, aún quedan muchos aspectos que mejorar incluso en estos países. En particular, aún no se han establecido objetivos de aprendizaje claros a nivel nacional y sólo recién se están iniciando esfuerzos sistemáticos por “compatibilizar” el currículum, los textos escolares y la pedagogía utilizada en la sala de clases. Chile y Argentina también han hecho avances en la integración de las evaluaciones en un sistema amplio de supervisión y evaluación.

Entre los principales problemas que requieren atención se encuentran la renuencia de México a divulgar los resultados de las evaluaciones, el enfoque “intermitente” de Costa Rica con respecto a las evaluaciones y el excesivo número de pruebas aplicadas; y el lento progreso de Brasil en la utilización de las evaluaciones para la reforma de los currículos y las políticas. Con excepción de Costa Rica, los seis países analizados en este trabajo han utilizado preguntas de selección múltiple más que preguntas de rendimiento o de interpretación abierta.

Otro punto que debe tomarse en cuenta es que los seis países informaron resultados similares en los exámenes. En particular, los alumnos que cursaban los últimos años de la educación básica y aquéllos que asistían a la educación secundaria obtuvieron puntajes muy por debajo de las expectativas de los educadores e investigadores profesionales. Los estudiantes con los mejores puntajes provenían de escuelas urbanas y privadas y tenían padres con un mayor nivel de educación. Pocos países de la región han emprendido análisis multivariados detallados de estos resultados con el fin de identificar sus causas. Si bien diversos estudios han mostrado una correlación entre los profesores mejor capacitados y un mejor rendimiento de los alumnos, la mayoría de estos estudios no controlaban el factor del nivel socioeconómico de los alumnos o la ubicación de

la escuela lo suficientemente como para confirmar la importancia de la educación de los profesores. Además, aún existe poca información con respecto a si los niveles de rendimiento de los alumnos han cambiado a través del tiempo.

Entre las innovaciones importantes en la región se incluye el intento de Brasil de evaluar las instituciones de educación superior; las evaluaciones de la capacidad de aprendizaje efectuadas en Costa Rica entre los niños que entran a la educación primaria y sus exámenes para los alumnos que terminan la educación secundaria; las evaluaciones aplicadas en México para medir los conocimientos y habilidades de los profesores; el énfasis dado en Colombia a la investigación; el uso que se hace en Chile de las evaluaciones para orientar la asignación de recursos y los esfuerzos de Argentina por utilizar las evaluaciones para la reforma curricular. El programa UNESCO/OREALC también constituye una importante innovación que tendrá mayor impacto a medida que se vaya aumentando la coordinación con las instituciones como la IEA.

En general, entre las principales lecciones aprendidas y desafíos para el futuro se incluyen los siguientes: (a) la importancia de un consenso nacional y un compromiso de largo plazo; (b) la importancia de centrarse explícitamente en el uso de las evaluaciones como herramientas para mejorar el aprendizaje; (c) la necesidad de generar capacidades y competencia técnica y (d) la necesidad de obtener beneficios de los adelantos e innovaciones provenientes tanto de América Latina como de otras regiones, especialmente a través de programas de pruebas internacionales como los de la IEA.

### **Generación de consenso y compromiso**

**Generación de un consenso nacional.** Tal vez la lección más importante a la fecha es que la necesidad de llevar a cabo evaluaciones ya no debiera ser un tema de debate en los países de América Latina. Por el contrario, estos países

necesitan establecer objetivos educacionales, determinar si los niños, las instituciones y los sistemas escolares están cumpliendo sus objetivos y, luego, establecer programas tendientes a asegurar que dichos objetivos se cumplan finalmente. Las evaluaciones no servirán para mejorar la calidad de la educación a menos que todos concuerden en la importancia de mejorar la calidad y en la necesidad de informar los resultados de las evaluaciones de una manera amplia, oportuna y fácilmente comprensible a todos los participantes en el proceso (es decir, de manera “transparente”). Este consenso debe representar una coalición entre los profesores, los padres, los administradores y los líderes empresariales y políticos. Debe persistir en el tiempo y contar con el respaldo de las más altas autoridades de gobierno; sin embargo, sus decisiones no se deben ver influidas por los compromisos políticos. Los programas de evaluación deben ser llevados adelante con decisión y no se debe permitir que sean suspendidos y luego reiniciados en forma intermitente (como en el caso de Costa Rica) ni que sus resultados sean ocultados a la opinión pública (como en el caso de México).

**Integración de los profesores en el proceso de evaluación.** Es preciso incluir a los profesores en el proceso de evaluación desde el comienzo si se desea que las evaluaciones tengan impacto en la educación. En América Latina, las asociaciones gremiales de profesores tienden a ser observadores pasivos o incluso a oponerse a las evaluaciones. Es preciso convencerlas de que el hecho de adoptar un enfoque más profesional con respecto a su trabajo, perfeccionar los procesos de la sala de clases y participar cabalmente en todos los programas de evaluación va en su propio interés. Las asociaciones gremiales no deben temer que las evaluaciones sean utilizadas para recompensar o castigar a los profesores, puesto que el alumnado varía tanto de un año al otro que la calidad de la enseñanza siempre será difícil de medir con precisión, como lo demuestra la experiencia en Francia, Estados Unidos y México. Si las evaluaciones se visualizan como una herra-

mienta para evaluar a las escuelas y a los sistemas escolares locales en lugar de los profesores, tanto los padres como los profesores y los administradores estarán mejor capacitados para trabajar conjuntamente como equipo con el fin de mejorar los puntajes en el rendimiento a nivel de las escuelas.

**Un enfoque sistémico.** Un compromiso con la transparencia también requiere un enfoque sistémico para abordar la supervisión y la evaluación. Al respecto, debe entenderse que las evaluaciones sólo miden el rendimiento de los niños y es poco lo que nos dicen acerca de la magnitud y las causas de la deserción. Otras herramientas para medir el rendimiento de los alumnos y del sistema son las siguientes: estudios de la deserción y la repitencia, mediciones de los insumos escolares y estimaciones de los recursos mínimos que cada escuela debiera tener, observaciones sistemáticas de los procesos escolares y estudios del desempeño de los egresados de las escuelas en el mercado laboral.

### **Usos de las evaluaciones**

El mayor desafío para el futuro en América Latina será asegurar que las evaluaciones se utilicen precisamente como herramientas para mejorar la calidad de la educación. Las evaluaciones pueden influir en muchos elementos del sistema educacional, incluidas las políticas educacionales nacionales, los programas de reforma educativa, los currículos educacionales, las decisiones de los padres, estudiantes y profesores, las políticas educacionales locales y regionales y la pedagogía y los programas a nivel de las escuelas. Tres áreas de particular importancia en América Latina en las cuales pueden utilizarse las evaluaciones son las siguientes: el refuerzo de la “compatibilización”, la orientación de los recursos a las escuelas más necesitadas y el estímulo a la responsabilidad y la recompensa de los logros a nivel de las escuelas.

**Refuerzo de la “compatibilización”.** Por sobre todo, las evaluaciones pueden ayudar a los



países de América Latina a lograr una mayor compatibilización del currículum que se pretende aplicar, el currículum que realmente se aplica, los textos escolares, los conocimientos de los profesores, la pedagogía en la sala de clases y el aprendizaje. Los ajustes resultantes permitirán establecer objetivos más altos pero factibles, mejorar la capacitación docente y la pedagogía, y asegurar que el currículum revisado sea aplicado en la práctica. Esto requerirá un gran trabajo de análisis, difusión y capacitación. Con una buena retroalimentación y recursos, las evaluaciones podrán ser utilizadas directamente como una herramienta para mejorar la enseñanza y el aprendizaje. Argentina ha hecho los mayores avances en este proceso a la fecha.

**Orientación de los recursos.** Dada la gran disparidad en los logros educacionales en América Latina, las evaluaciones deben ser utilizadas para canalizar los recursos adicionales hacia las escuelas más necesitadas de la región, particularmente aquéllas que se encuentran en las áreas rurales y en los barrios más pobres de las zonas urbanas. Dicha actividad puede llevarse a cabo siguiendo el modelo del “Programa de las 900 escuelas” aplicado en Chile, que está explícitamente dirigido al mejoramiento del rendimiento de las escuelas que han obtenido los peores puntajes.

**Estímulo a un mejor rendimiento en las escuelas.** Las evaluaciones deben proporcionar información útil a las escuelas, padres y profesores. Esto requerirá, entre otros elementos, determinar el “valor agregado” de la escolaridad, identificando aquellas escuelas que obtienen un puntaje mayor que el esperado en atención a los antecedentes socioeconómicos de los alumnos, como se hiciera en Francia, o midiendo el grado de mejoría en los puntajes de las escuelas en períodos sucesivos. Estos enfoques conducen a un reconocimiento y recompensa de las escuelas “eficaces” y motivan a las demás escuelas a reproducir las características de dichos establecimientos. Este enfoque elimina también el sesgo en contra de las escuelas que se encuentran en barrios pobres.

Hasta la fecha, Chile es el país que más ha avanzado en el reconocimiento y estímulo a dichas escuelas.

### **Creación de capacidades y competencia técnica**

**Fortalecimiento de las instituciones independientes.** En la mayoría de los países de América Latina, sería muy difícil para el gobierno central implementar un programa de evaluación directamente en el largo plazo, como se hizo en Francia. Entre los obstáculos a un sistema de ese tipo se encuentran los posibles conflictos de intereses, los sueldos insuficientes y los procedimientos de contratación y administración inflexibles. Una estrategia más apropiada para la región sería fomentar la creación de organismos de pruebas competentes independientes del gobierno, manteniendo un organismo gubernamental pequeño, especializado y altamente eficiente en la función de supervisar dichas actividades. Colombia parece haber desarrollado una sólida asociación de este tipo entre el sector público y el sector privado.

**Garantía de conocimientos especializados adecuados e investigación de apoyo.** Si se desea que las evaluaciones sean eficientes como herramientas para medir lo que se pretende que midan, se requiere un alto nivel de conocimientos especializados. De lo contrario, es muy fácil que presenten deficiencias: puede que no discriminen adecuadamente entre los aprendientes; la comparabilidad de un año al otro puede verse afectada; las escuelas y regiones pueden engañar; el marco de la muestra puede ser inexacto, con lo cual toda la evaluación puede resultar inútil; técnicas analíticas más sofisticadas pueden mostrar que las supuestas relaciones “causales” son falsas. Muchos de estos problemas ya han sido detectados en las evaluaciones hechas por los países. Por lo tanto, es esencial capacitar y remunerar adecuadamente a expertos en evaluación de currículos, metodología de estudios muestrales y técnicas analíticas. Como se ha hecho en Jordania, es importante consultar a las

autoridades más conocidas y respetadas del mundo en el área de las evaluaciones, ya que no existe una metodología “regional” de pruebas.

Los gobiernos y fundaciones también necesitan fomentar centros de posgrado de excelencia en el área de la docencia y la investigación en la región. En especial, existe la necesidad de apoyar las investigaciones independientes basadas en los datos de las evaluaciones, que constituyen una rica fuente de información sobre todos los aspectos de la educación relacionados con los alumnos, los profesores y las escuelas. Para obtener los mejores resultados, estas investigaciones deben basarse en las técnicas analíticas más modernas. Otro aspecto igualmente importante, es que los investigadores como asimismo quienes diseñan las políticas deben estar de acuerdo en los objetivos de las investigaciones y estar dispuestos a divulgar los resultados aun si éstos son ambiguos o negativos.

### **Nuevos enfoques**

**Participación en los programas de evaluación internacionales.** El TIMSS ofrece a los países de América Latina una enorme oportunidad para el establecimiento de estándares, la compatibilización de los currículos y los textos escolares, el perfeccionamiento del profesorado y los avances en el aprendizaje. Toda la región puede obtener beneficios de los conocimientos especializados y la metodología del TIMSS.

Actualmente, la IEA tiene planes de duplicar el TIMSS en 1999. Chile y Brasil ya han expresado su intención de participar. Los países latinoamericanos podrían también participar en el nuevo programa de educación cívica de la IEA, que será particularmente importante por los aspectos relacionados con el fortalecimiento de la sociedad civil en la región. Actualmente, Chile y Colombia están comprometidos con el estudio relativo a la educación cívica.

Como alternativa, los investigadores del TIMSS podrían llevar a cabo un análisis de los

currículos en los textos y las pautas curriculares y compararlos con los currículos aplicados en otras regiones del mundo. También podrían hacer el intento de determinar si los exámenes nacionales y otros son compatibles con el currículum nacional que se pretende aplicar. La metodología del TIMSS ya ha sido utilizada para analizar el contenido curricular, los textos escolares y la pedagogía en Argentina, Colombia, México y República Dominicana. Es posible solicitar asistencia técnica en cualquiera de estas actividades tanto a los profesionales internacionales como a los investigadores regionales del TIMSS.

**Método basado en el universo versus método basado en una muestra.** Antes de emprender una evaluación del universo, los países deben estar conscientes de que ésta es una tarea de alto costo y que requiere una administración minuciosa por parte de profesionales competentes. También se necesita un sistema de retroalimentación detallado y apoyo a las escuelas. En Chile, el costo de la aplicación de pruebas al universo es de aproximadamente \$5 dólares por alumno, lo que representa aproximadamente el 2% de los costos unitarios totales. Incluso esta reducida cantidad puede ser difícil de justificar en los países con graves restricciones fiscales y que carecen de fondos para material pedagógico.

En comparación, los estudios muestrales cuestan menos, pero pueden proporcionar información suficiente para el establecimiento de políticas nacionales y para la identificación de los problemas de compatibilización entre el currículum, los textos escolares y la enseñanza. Sin embargo, estos estudios requieren precisas técnicas de muestreo y una administración rigurosa. Siempre que sea posible, los países de América Latina también deben utilizar los exámenes de selección existentes para efectos de evaluación.

**Pruebas de rendimiento.** En los últimos años, Estados Unidos ha ido incorporado cada vez más las preguntas de interpretación abierta en sus pruebas de rendimiento, en tanto que Europa se ha

concentrado en mejorar la confiabilidad y validez de los instrumentos de medición. Por el contrario, la mayoría de los países de América Latina todavía favorecen las evaluaciones del tipo “selección múltiple”. Los países de la región debieran por lo menos explorar los enfoques más nuevos con respecto a las pruebas. Sin embargo, no podrán hacerlo si aquéllas personas que califican las pruebas no cuentan con una cabal capacitación en los procedimientos pertinentes.

**Exámenes de competencia mínima.** Puede que los países de América Latina deseen considerar la aplicación de “exámenes de competencia mínima” para el nivel secundario. Estos ya han sido desarrollados en algunas partes de Estados Unidos y en Costa Rica, como una manera de estimular el aprendizaje. Este tipo de exámenes serían especialmente apropiados para los países descentralizados de gran tamaño, como Brasil.

**Opciones para los países más pequeños y más pobres.** Los países más pequeños y más pobres simplemente no cuentan con los conocimientos técnicos especializados ni los recursos necesarios para implementar un sistema de evaluación en forma sostenida. Como lo sugiere la experiencia de Costa Rica, los conocimientos especializados locales pueden ser insuficientes. Por lo tanto, los países más pequeños deberían emprender sólo un mínimo de pruebas (por ejemplo, sólo matemáticas e idioma) en unos pocos grados. Otra opción es operar en consorcios (compuestos, por ejemplo, por los países centroamericanos), hacer uso de los programas internacionales y aprovechar la asistencia técnica regional (por ejemplo, de Chile y Argentina).

**Cooperación regional.** Para fortalecer el proceso de coordinación regional, será necesario desarrollar vínculos más estrechos con la IEA y con los principales centros de investigación y desarrollo del mundo en el área de las pruebas y evaluaciones, junto con establecer una amplia participación, que incluya las instituciones no gubernamentales, en los foros y comités

regionales. También son necesarios los esfuerzos a nivel regional para: analizar los resultados de las evaluaciones de la UNESCO/OREALC de manera de identificar los factores asociados a un alto rendimiento, comprender las interacciones entre los objetivos curriculares y el aprendizaje a través de la medición de las “oportunidades para el aprendizaje” y, lo que es más importante, identificar y apoyar los centros regionales de excelencia dedicados a las pruebas, las mediciones y las investigaciones educacionales.

## Bibliografía

- Aldana, Eduardo. 1996. “Medición del logro educativo: el caso colombiano”. Seminario de PREAL sobre Evaluación y Estándares en la Educación Latinoamericana, Río de Janeiro.
- Alvarez, Benjamín, y Mónica Ruiz-Casares, eds. 1997. *Assessment and Education Reform: Policy Options*. 1997. Washington D.C.: Academy for Educational Development.
- Capper, Joanne. 1996. “Testing to Learn, Learning to Test”. Newark, Del.: International Reading Association.
- Esquivel, Juan Manuel. 1996. “Medición de logros del aprendizaje y el empleo de los resultados en Costa Rica”. Seminario de PREAL sobre Evaluación y Estándares en la Educación Latinoamericana, Río de Janeiro.
- Horn, Robin, Laurence Wolff y Eduardo Vélez. 1991. *Developing Educational Assessment Systems in Latin America: A Review of Issues and Recent Experience*. Washington D.C.: Banco Mundial.
- Ministerio de Educación y Cultura, Brasil. 1996. “Seminario Internacional de Avaliação da Educação”, octubre de 1995, Río de Janeiro.

- De Moura Castro, Claudio, y otros. 1984. "Determinantes de la Educación en América Latina: Acceso, Desempeño y Equidad". ECIEL (Programa de Estudios Conjuntos para la Integración Económica Latinoamericana). Río de Janeiro: Fundação Getulio Vargas.
- Murphy, Paul, Vincent Greaney, Marlaine E. Lockheed y Carlos Rojas, eds. 1996., National Assessments - Testing the System. Washington, D.C.: Banco Mundial, Economic Development Institute.
- National Research Council. 1995. National Science Education Standards. Washington, D.C.: National Academy Press.
- Rodríguez Schuller, Cristina. 1996. "Sistema de Medición de la Calidad de la Educación: características y uso de los resultados de la evaluación para mejorar la calidad (Chile)". Seminario de PREAL sobre Evaluación y Estándares en la Educación Latinoamericana, Río de Janeiro.
- Schmelkes, Silvia. 1996. "Propuesta para el diseño y desarrollo de un sistema nacional de evaluación integral de la educación básica en México". Seminario de PREAL sobre Evaluación y Estándares en la Educación Latinoamericana, Río de Janeiro.
- Schmidt, William, Curtis McKnight y Senta A. Raizen. 1997. *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education*. Dordrecht: Kluwer Academic.
- Schmidt, William A., y otros. 1997. *Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Science*. Dordrecht: Kluwer Academic.
- U.S. National Research Center for the Third International Mathematics and Science Study. 1997. *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education*. Dordrecht: Kluwer Academic.
- Valverde, Gilbert A. 1996. "Las políticas educacionales nacionales y la participación en estudios internacionales". Seminario de PREAL sobre Evaluación y Estándares en la Educación Latinoamericana, Río de Janeiro.
- Vianna, Heraldo Marelim. 1996. "Programas de avaliação de sistema de educação estadual do estado de São Paulo", Seminario de PREAL sobre Evaluación y Estándares en la Educación Latinoamericana, Río de Janeiro.

## ANEXO: UNA MIRADA A LOS SISTEMAS DE EVALUACIÓN DE SEIS PAISES

A R G E N T I N A	B R A S I L
<p><b>Fecha de la primera evaluación:</b> 1993.</p> <p><b>Fecha del programa de evaluación más reciente:</b> 1995.</p> <p><b>Grados evaluados (programa más reciente):</b> Grados 2,3,5 y 7 en educación primaria.</p> <p><b>Materias evaluadas:</b> Matemáticas y lectura y, en el 7° grado, también ciencias naturales y sociales.</p> <p><b>Muestra o universo:</b> Universo de estudiantes.</p> <p><b>Resultados importantes:</b> Las preguntas se seleccionaron en base a los niveles mínimos esperados, pero los escolares respondieron sólo el 50% de las preguntas correctamente; los factores escolares representaron el 40% de la varianza asociada a los puntajes en matemáticas.</p> <p><b>Tipos de investigaciones realizadas:</b> Correlaciones simples y tabulaciones cruzadas.</p> <p><b>Información disponible acerca de los costos:</b> —</p> <p><b>Elementos técnicos de interés:</b> Selección múltiple e interpretación abierta.</p> <p><b>Estructura institucional:</b> Administrado por el Ministerio de Educación, pero dado en contrato a diversas instituciones privadas.</p> <p><b>Retroalimentación:</b> Programa amplio que entrega información acerca del rendimiento escolar, los perfiles de los alumnos, los profesores y los modelos administrativos asociados al rendimiento; identifica el contenido con mayor dificultad e incluye recomendaciones para mejoramientos en la pedagogía.</p> <p><b>Impacto sobre las políticas:</b> Manuales metodológicos para mejorar el aprendizaje como asimismo capacitación en el servicio basada en los resultados de las pruebas distribuidos a todas las escuelas; asistencia técnica entregada a 24 estados.</p> <p><b>Aspectos políticos y otros:</b> —</p> <p><b>Relación con programas de evaluación internacionales:</b> Participó en el análisis de currículos del TIMSS y emprendió estudios demasiado tarde como para ser incluidos en el informe; ha entregado asistencia técnica a Bolivia y Paraguay.</p>	<p><b>Fecha de la primera evaluación:</b> 1990.</p> <p><b>Fecha del programa de evaluación más reciente:</b> 1995.</p> <p><b>Grados evaluados (programa más reciente):</b> Grados 4 y 8 de la educación primaria y grados 2 y 3 de la educación secundaria (además, Brasil ha iniciado un programa de evaluación de todos los estudiantes universitarios).</p> <p><b>Materias evaluadas:</b> Matemáticas y lectura.</p> <p><b>Muestra o universo:</b> Muestra de 90.499 alumnos de 2.289 escuelas públicas y 511 escuelas privadas.</p> <p><b>Resultados importantes:</b> Los escolares generalmente obtuvieron puntajes muy por debajo de los niveles de aprendizaje esperados definidos por los profesores, investigadores y especialistas; los puntajes más altos se observaron en el sur del país, en las ciudades capitales, en las escuelas urbanas y privadas y en las escuelas diurnas de jornada completa; los estudiantes más jóvenes, los estudiantes blancos y orientales y los estudiantes con profesores con mayor capacitación obtuvieron los mejores resultados.</p> <p><b>Tipos de investigaciones realizadas:</b> Correlaciones simples y tabulaciones cruzadas.</p> <p><b>Información disponible acerca de los costos:</b> —</p> <p><b>Elementos técnicos de interés:</b> Selección múltiple.</p> <p><b>Estructura institucional:</b> Administrado por el Ministerio de Educación, pero dado en contrato a diversas instituciones privadas.</p> <p><b>Retroalimentación:</b> Los informes fueron puestos a disposición de la prensa, pero no fueron utilizados por las escuelas mismas.</p> <p><b>Impacto sobre las políticas:</b> A la fecha no se ha tomado ninguna decisión específica sobre la base del programa de evaluación.</p> <p><b>Aspectos políticos y otros:</b> No se han considerado en forma adecuada los usos de los resultados de las evaluaciones en el área de las políticas, tales como la concentración en las escuelas con peor rendimiento.</p> <p><b>Relación con programas de evaluación internacionales:</b> Participó en el estudio de la IEA en matemáticas y lectura en 1993 y tiene planes de participar en el TIMSS-R.</p>

C H I L E	C O L O M B I A
<p><b>Fecha de la primera evaluación:</b> 1981.</p> <p><b>Fecha del programa de evaluación más reciente:</b> 1991-96.</p> <p><b>Grados evaluados (programa más reciente):</b> Grados 4 y 8.</p> <p><b>Materias evaluadas:</b> Lectura y aritmética en los grados 4 y 8: muestra del 10% en ciencias naturales, historia y geografía.</p> <p><b>Muestra o universo:</b> Principalmente universal, excepto lo observado anteriormente.</p> <p><b>Resultados importantes:</b> Los resultados continúan mostrando un bajo rendimiento en las escuelas públicas municipalizadas y en las escuelas rurales. Alta correlación entre el nivel socioeconómico y el rendimiento; se han producido significativos mejoramientos en los puntajes en las escuelas participantes en el programa de “900 escuelas” que se concentra en las escuelas con peor rendimiento.</p> <p><b>Tipos de investigaciones realizadas:</b> Correlaciones simples y tabulaciones cruzadas; trabajo analítico más sofisticado actualmente en curso.</p> <p><b>Información disponible acerca de los costos:</b> US\$ 5 por estudiante evaluado.</p> <p><b>Elementos técnicos de interés:</b> Selección múltiple hasta ahora; se considera la introducción de más preguntas de rendimiento, interpretación abierta y ensayo; se está revisando el método estadístico para asegurar una mayor confiabilidad de un año al siguiente.</p> <p><b>Estructura institucional:</b> La capacidad de pruebas ha sido creada por el Ministerio de Educación más que por un organismo autónomo.</p> <p><b>Retroalimentación:</b> Los informes están siendo constantemente simplificados y ajustados a receptores específicos, especialmente a nivel de las escuelas; los directores de las escuelas, los profesores y los padres utilizan los resultados de las pruebas; algunos padres utilizan los resultados de las pruebas para seleccionar las escuelas; los diseñadores del currículum nacional están utilizando los informes para revisar el currículum.</p> <p><b>Impacto sobre las políticas:</b> Los diseñadores del currículum están dando énfasis al dominio de las materias en las áreas de debilidad; se ha continuado con los programas pedagógicos dirigidos a las escuelas de peor rendimiento (programa de “900 escuelas”).</p> <p><b>Aspectos políticos y otros:</b> El programa está tan institucionalizado que se están volviendo más importantes aspectos de “segundo orden”; por ejemplo, los profesores podrían estar “enseñando con miras a la prueba” de manera exagerada y las escuelas podrían no estar considerando el nivel de pobreza de los estudiantes.</p> <p><b>Relación con programas de evaluación internacionales:</b> Ninguna directamente hasta el momento, pero existen planes de participar en el TIMSS-R.</p>	<p><b>Fecha de la primera evaluación:</b> 1990.</p> <p><b>Fecha del programa de evaluación más reciente:</b> 1994.</p> <p><b>Grados evaluados (programa más reciente):</b> Inicialmente los grados 3 y 5 y posteriormente los grados 7 y 9.</p> <p><b>Materias evaluadas:</b> Inicialmente matemáticas y lectura; posteriormente, también ciencias naturales y sociales.</p> <p><b>Muestra o universo:</b> Sólo muestra.</p> <p><b>Resultados importantes:</b> Los niveles cognitivos de los alumnos de los grados 3 y 5 son bajos comparados con las expectativas, especialmente en el área de las habilidades críticas y reflexivas. Los mejores puntajes se obtienen en las escuelas urbanas y privadas, entre quienes han asistido a la enseñanza pre escolar, no han repetido, tenían libros en sus casas y tenían padres con niveles de educación relativamente altos; los alumnos de clases con profesores con mayor nivel de educación, disponibilidad de textos escolares, docentes de sexo femenino, de la escuela nueva y de escuelas completas obtuvieron mejores resultados.</p> <p><b>Tipos de investigaciones realizadas:</b> Correlaciones, tabulaciones cruzadas y algún análisis multivariado.</p> <p><b>Información disponible acerca de los costos:</b> —</p> <p><b>Elementos técnicos de interés:</b> Selección múltiple.</p> <p><b>Estructura institucional:</b> Empresa conjunta entre el Ministerio de Educación, el Instituto Ser, la Universidad Nacional de Profesores y el Centro de Estudios Sociales.</p> <p><b>Retroalimentación:</b> Los informes y resúmenes de las investigaciones están disponibles pero no son útiles para la mayoría de las escuelas y municipalidades debido a la base de la muestra.</p> <p><b>Impacto sobre las políticas:</b> La Ley de Educación de 1994 incorpora los resultados de las evaluaciones tales como la concentración en la educación preescolar, la educación de los profesores y los textos escolares.</p> <p><b>Aspectos políticos y otros:</b> Se considera que la expansión al universo sería demasiado cara y superaría las capacidades institucionales, pero se ampliará la muestra de manera de entregar más información a las escuelas y regiones evaluadas.</p> <p><b>Relación con programas de evaluación internacionales:</b> Pleno participante en el TIMSS con información de los resultados; tiene planes de participar en el estudio de ciencias sociales de la IEA.</p>

C O S T A R I C A	M E X I C O
<p><b>Fecha de la primera evaluación:</b> 1987.</p> <p><b>Fecha del programa de evaluación más reciente:</b> 1993-97.</p> <p><b>Grados evaluados (programa más reciente):</b> Grados 1,3,6,9,12.</p> <p><b>Materias evaluadas:</b> Grado 1, capacidad de aprendizaje; grados 3, 9, 12, todas las materias; también muestras sobre las capacidades cognitivas y físicas en diversos grados.</p> <p><b>Muestra o universo:</b> Principalmente universo, con excepción de las capacidades físicas y cognitivas y la capacidad de aprendizaje.</p> <p><b>Resultados importantes:</b> Aún no se encuentran disponibles para las pruebas actuales; los resultados anteriores mostraban que los alumnos tenían un rendimiento muy inferior al esperado en comparación a los objetivos del currículum nacional; los alumnos de las escuelas privadas, urbanas y académicas obtuvieron los mejores resultados. Los alumnos obtuvieron puntajes progresivamente peores en comparación con los objetivos del currículum a medida que aumentaba su tiempo de permanencia en la escuela.</p> <p><b>Tipos de investigaciones realizadas:</b> Tabulaciones cruzadas simples.</p> <p><b>Información disponible acerca de los costos:</b> US\$1,50-US\$7,50 por alumno evaluado en 1991.</p> <p><b>Elementos técnicos de interés:</b> Principalmente selección múltiple.</p> <p><b>Estructura institucional:</b> Se está probando actualmente una empresa de cooperación entre el Ministerio de Educación y el IIMEC en la Universidad de Costa Rica.</p> <p><b>Retroalimentación:</b> Los anteriores programas de evaluación tuvieron poca difusión; en el nuevo programa, los informes serán preparados para múltiples receptores.</p> <p><b>Impacto sobre las políticas:</b> Reintroducción en 1989 de las pruebas nacionales o regionales que representan un alto porcentaje de las notas finales de los alumnos; utilizadas como justificación para el aumento de los fondos para educación.</p> <p><b>Aspectos políticos y otros:</b> Enfoque intermitente con respecto a las evaluaciones, dependiendo de la actitud de los sucesivos ministros de educación; el actual ministro está a favor de las evaluaciones.</p> <p><b>Relación con programas de evaluación internacionales:</b> Ninguna directamente.</p>	<p><b>Fecha de la primera evaluación:</b> 1976.</p> <p><b>Fecha del programa de evaluación más reciente:</b> 1994.</p> <p><b>Grados evaluados (programa más reciente):</b> Todos los grados de la educación básica y secundaria, como también los profesores.</p> <p><b>Materias evaluadas:</b> Matemáticas y lectura en las escuelas; se evaluó a los profesores en las materias y los conocimientos pedagógicos.</p> <p><b>Muestra o universo:</b> Universo en 1994 anteriormente muestras.</p> <p><b>Resultados importantes:</b> A medida que los alumnos avanzan en la escuela, obtienen puntajes progresivamente más bajos que lo esperado; los puntajes más altos fueron obtenidos por los alumnos que habían asistido a la enseñanza preescolar, no estaban repitiendo el año, asistían a escuelas urbanas o privadas y tenían padres mejor educados en áreas urbanas; los niños de las escuelas indígenas y comunitarias con instalaciones escolares inadecuadas y profesores con menor capacitación obtuvieron los peores puntajes.</p> <p><b>Tipos de investigaciones realizadas:</b> Correlaciones simples y tabulaciones cruzadas.</p> <p><b>Información disponible acerca de los costos:</b> US\$35 por alumno en las muestras tomadas en 1990.</p> <p><b>Elementos técnicos de interés:</b> Selección múltiple.</p> <p><b>Estructura institucional:</b> Administrado por el Ministerio de Educación (Secretaría de Educación Pública).</p> <p><b>Retroalimentación:</b> Los informes se entregan a la opinión pública pero no son fáciles de comprender; hasta 1990 los resultados de las evaluaciones tuvieron carácter confidencial.</p> <p><b>Impacto sobre las políticas:</b> Hasta hace poco ningún impacto sobre las políticas; las evaluaciones a los profesores están relacionadas con el convenio con la asociación gremial nacional para mejorar la calidad de la educación básica.</p> <p><b>Aspectos políticos y otros:</b> Difusión y utilización inadecuadas a nivel de las escuelas.</p> <p><b>Relación con programas de evaluación internacionales:</b> Pleno participante en el TIMSS, pero en el último momento solicitó que no se informaran los resultados.</p>

  
**Ve. A. Ce.**

Teléfonos: (56-2)638 5746 - (56-2)285 1837

E-mail: [ve.a.ce@bellsouth.cl](mailto:ve.a.ce@bellsouth.cl)

Santiago - Chile

---