

REVISTA IBERO— —AMERICANA

de Educación

de Educação



Organización
de Estados
Iberoamericanos

para la Educación,
la Ciencia
y la Cultura

NÚMERO 28

Monográfico: Enseñanza de la tecnología / *Ensino da tecnologia*

Enero-Abril 2002 / Janeiro-Abril 2002

TÍTULO: Oportunidades regionales en la participación en estudios comparativos internacionales de evaluación educativa

AUTOR: Alejandro Tiana Ferrer y Guillermo Gil Escudero

OPORTUNIDADES REGIONALES EN LA PARTICIPACIÓN EN ESTUDIOS COMPARATIVOS INTERNACIONALES DE EVALUACIÓN EDUCATIVA¹

Alejandro Tiana Ferrer (*)
Guillermo Gil Escudero ()**

SÍNTESIS: Los estudios para evaluar la calidad de la educación han experimentado, en los últimos años, una clara expansión. Ello se debe al impulso que han dado a los mismos algunos organismos internacionales como la Asociación para la Evaluación de los Logros Educativos (International Association for the Evaluation Achievement - IEA), la Organización para la Cooperación y el Desarrollo Económico (OCDE) y, en el ámbito latinoamericano, la Oficina Regional de Educación para América Latina y el Caribe (OREALC-UNESCO).

Estas iniciativas responden, sin duda, al profundo cambio producido en el último cuarto del siglo pasado, en el que el foco prioritario de las políticas educativas se desplazó desde la expansión de los sistemas para alcanzar adecuados niveles de cobertura, hacia la calidad de los procesos de enseñanza-aprendizaje como mecanismo y garantía de resultados acordes con los requerimientos de las sociedades y de los mercados laborales.

Iberoamérica, y en particular los países latinoamericanos, han sido los grandes ausentes en la mayoría de este tipo de investigaciones, exceptuando las desarrolladas por el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación en trece países de la región.

Sin embargo, los posibles beneficios resultantes de los estudios internacionales, que permiten comparar realidades educativas mediante cortes transversales y sincrónicos, no han tenido oportunidad de verificarse en Latinoamérica.

El presente trabajo, basado en un proyecto promovido por el Banco Interamericano de Desarrollo (BID), muestra, a partir del análisis de cinco estudios internacionales y de los resultados que los mismos ofrecen para

¹ Este trabajo es el resultado de un proyecto promovido y financiado por el Banco Interamericano de Desarrollo, y constituye el complemento de otro anterior, publicado por el propio Banco (Tiana, 2000).

(*) Vicerrector de la Universidad Nacional de Educación a Distancia (UNED), España.

(**) Jefe del Servicio de Evaluación del Instituto Nacional de Calidad y Evaluación (INCE), España.

la elaboración de políticas educativas, las posibilidades, requerimientos metodológicos, condiciones y beneficios que la participación regional en este tipo de investigaciones puede tener para el mejoramiento de la calidad de la educación en América Latina.

SÍNTESE: *Os estudos para avaliar a qualidade da educação têm experimentado, nos últimos anos, uma clara expansão. Isto se deve ao impulso de alguns organismos internacionais, como a Associação para a Avaliação dos Resultados Educativos (International Association for the Evaluation Achievement - IEA), a Organização para a Cooperação e o Desenvolvimento Econômico (OCDE) e, no âmbito latino-americano, o Órgão Regional de Educação para América Latina e o Caribe (OREALC-UNESCO).*

Estas iniciativas respondem, sem dúvida, à mudança profunda ocorrida no último quarto do século passado, no qual o foco prioritário das políticas educativas deslocou-se da expansão dos sistemas para alcançar níveis adequados de cobertura à qualidade de processos de ensino-aprendizagem como mecanismo e garantia de resultados de acordo com as exigências das sociedades e dos mercados de trabalho.

Ibero-América e, em particular, os países latino-americanos têm sido os grandes ausentes na maioria desse tipo de investigações, excetuando as desenvolvidas pelo Laboratório Latino-americano de Avaliação da Qualidade da Educação em treze países da região.

No entanto, os possíveis benefícios resultantes dos estudos internacionais, que permitem comparar realidades educativas por meio de cortes transversais e sincrônicos, não têm tido oportunidade de se verificar na América Latina.

O presente trabalho, baseado num projeto promovido pelo Banco Interamericano de Desenvolvimento (BID), mostra, a partir da análise de cinco estudos internacionais e dos resultados que os mesmos oferecem para a elaboração de políticas educativas, as possibilidades, notificações metodológicas, condições e benefícios que a participação regional neste tipo de investigações pode ter para a melhoria da qualidade da educação na América Latina.

1. INTRODUCCIÓN

Durante la última década se ha producido una notable expansión de la evaluación educativa a escala mundial, que ha pasado de centrarse básicamente en los estudiantes y en el currículo a prestar atención al funcionamiento de los establecimientos docentes y al rendimiento de los sistemas educativos (Tiana, 1998). Ese nuevo interés por la evaluación ha producido diversos efectos. Por una parte, ha impulsado la puesta en marcha de una gran variedad de proyectos de evaluación, que tienen por objeto valorar el funcionamiento de diversos programas e instituciones educativas. Por otra parte, ha determinado el estableci-

miento de diversos organismos destinados a evaluar los sistemas educativos nacionales. Y, también, ha impulsado la participación de muchos países en los estudios internacionales de evaluación, cuyos primeros pasos se remontan a los años sesenta, pero cuya expansión se ha dejado sentir sobre todo en la última década.

La realización de estudios internacionales de evaluación de la educación ofrece indudable interés para los países participantes. De un lado, les permite comparar su rendimiento educativo con el de otros países que les pueden servir de punto de referencia. De otro, les proporciona elementos de reflexión acerca de los objetivos educativos que se han propuesto, los medios que han puesto en marcha y los resultados efectivos que han logrado. Así mismo, les da la posibilidad de experimentar y aplicar nuevas metodologías de evaluación, capaces de mejorar la información que se obtiene sobre el estado y la situación del sistema educativo.

Aunque las ventajas que ofrece la participación en estudios internacionales de evaluación parecen fuera de toda duda, la región latinoamericana y caribeña se ha implicado poco en esa tarea hasta ahora (Tiana, 2000). Han sido escasos los países del área que han participado activamente en los distintos proyectos desarrollados por organizaciones tales como la IEA o la OCDE y pocas las iniciativas de carácter latinoamericano, como las auspiciadas por OREALC-UNESCO. Por dicho motivo, el Banco Interamericano de Desarrollo creyó conveniente promover este trabajo, destinado a poner de relieve los beneficios que podrían obtener los sistemas educativos de la región a través de su participación en estudios internacionales de evaluación, así como las posibilidades que existen para desarrollar enfoques de carácter regional en el seno de aquellos estudios.

2. ESTUDIOS ACTUALES EN EL CAMPO DE LA EVALUACIÓN COMPARATIVA INTERNACIONAL DEL RENDIMIENTO ACADÉMICO

A comienzos de la década de los 70 comenzaron a gestarse los primeros estudios internacionales de evaluación de los sistemas educativos, todavía con unos objetivos limitados, una metodología relativamente sencilla y la participación de un número reducido de países. Sin embargo, aquellos primeros pasos permitieron asentar un desarrollo posterior mucho más ambicioso, que se dejó sentir con fuerza a partir de

finales de los ochenta. La labor pionera de aquellos primeros investigadores sirvió de base para la expansión de lo que hoy es un campo pujante.

En la actualidad existen diversos proyectos internacionales de evaluación de los sistemas educativos, de distinto carácter y con diferentes objetivos. Entre todos ellos merecen especial atención los que se orientan a la evaluación del logro o rendimiento académico de los estudiantes en diversas áreas. Hemos seleccionado aquí los promovidos por tres organizaciones que ejercen una labor de liderazgo en este ámbito. Se trata de los trabajos coordinados por la International Association for the Evaluation of Educational Achievement (IEA), la Organización para la Cooperación y el Desarrollo Económico (OCDE) y la Oficina Regional de Educación para América Latina y el Caribe (OREALC-UNESCO). Cada una de ellas ha desarrollado uno o más estudios internacionales de evaluación, que cuentan con presencia en la región latinoamericana y caribeña. Se han dejado de lado otros proyectos de carácter más amplio, tales como los encaminados a la producción de indicadores educativos o a la cooperación internacional en el ámbito de la evaluación educativa, puesto que sus características son diferentes y plantean otros problemas concretos. No debe deducirse de ello que se minusvalore su interés, sino sólo que se prefiere limitar voluntariamente el campo de estudio para poder profundizar algo más en el análisis.

2.1 ESTUDIOS DE LA IEA

Los primeros trabajos realizados por la International Association for the Evaluation of Educational Achievement (IEA) se remontan a hace más de cuarenta años, ya que fueron iniciados inmediatamente después de su creación en 1959. La IEA es una asociación internacional de carácter no gubernamental, formada por centros y agencias de investigación y de evaluación educativa, que desarrolla estudios comparativos del rendimiento logrado por los sistemas educativos. Hoy participan en ella unos sesenta países de todos los continentes y con diversos grados de desarrollo económico y social, si bien hay que señalar que son muy pocos los países latinoamericanos y caribeños que toman parte en sus trabajos.

A lo largo de estos años, la IEA ha desarrollado estudios en áreas académicas muy diversas, tales como la lectura, la escritura, las matemáticas, las ciencias, la educación preescolar, la educación cívica o las nuevas tecnologías (Degenhart, 1990). Los estudios que ha emprendido

tienen dos propósitos prioritarios. Por una parte, pretenden proporcionar una información significativa acerca de los logros de los sistemas educativos, recogiendo para ello datos procedentes de los estudiantes, de los profesores y de las instituciones escolares. Por otra parte, pretenden analizar los factores que explican las diferencias de rendimiento encontradas. Por lo tanto, no se limitan a establecer comparaciones del rendimiento alcanzado por los países, sino que se esfuerzan por explicar las diferencias halladas, aun cuando haya que reconocer la dificultad de esta tarea.

Para llevar a cabo dichos estudios, la IEA ha tenido que impulsar dos tipos de instrumentos. En primer lugar, ha elaborado pruebas de rendimiento internacionalmente válidas y aceptadas por los países participantes. Se trata de pruebas basadas en análisis curriculares detallados y muy ligadas al tipo de trabajo que se lleva a cabo en las aulas escolares, aspecto en el que se diferencia de otras iniciativas, como las auspiciadas por la OCDE. En segundo lugar, ha confeccionado cuestionarios de contexto y de procesos educativos dirigidos a los alumnos, a los profesores y a los directores escolares. En su dilatada trayectoria ha desarrollado y aplicado unos procedimientos técnicos contrastados, que han servido de base para los proyectos emprendidos por otras organizaciones. La experiencia acuñada por la IEA ha influido mucho en la evolución de los estudios comparativos internacionales de evaluación educativa.

2.1.1 Tercer Estudio Internacional en Matemáticas y Ciencias (TIMSS)

Entre las labores adelantadas durante los últimos años y aún en la actualidad por la IEA cabe destacar el Tercer Estudio Internacional en Matemáticas y Ciencias, más conocido por sus siglas TIMSS. Se trata del estudio más ambicioso desarrollado sobre estas áreas hasta el momento, si bien no es el único que se ha llevado a cabo. Su principal novedad en relación con los anteriores consiste en que adopta la concepción de un ciclo de estudios. Dicho de otro modo, en vez de analizar el rendimiento en estas áreas de una forma transversal, realizando una toma de datos única en una fecha determinada, ahora se pretende realizar un estudio con tomas de datos periódicas, que permitan examinar la progresión de los resultados académicos.

El estudio tiene lugar en tres grados escolares distintos: en el cuarto curso de escolaridad (población 1), en el octavo curso (población 2) y en el grado final de la educación secundaria (población 3). En cada uno de ellos se selecciona una muestra de escuelas y de alumnos a los que se aplican unas pruebas de matemáticas y ciencias y unos cuestionarios de contexto. Las pruebas incluyen preguntas de elección múltiple, cuestiones de respuesta libre y pruebas de habilidades (*performance tasks*). Además, se aplican unos cuestionarios a los profesores de matemáticas y de ciencias y a los directores de las escuelas incluidas en la muestra.

Las pruebas se han elaborado a partir de un marco conceptual previo y de un análisis curricular detallado, en cuya construcción han intervenido especialistas de los países participantes. Con objeto de poder estudiar el progreso educativo, las pruebas han mantenido algunas preguntas comunes de unos años y grados a otros, variando otra parte de las mismas. Para no reducir excesivamente la cobertura curricular de las pruebas, se han utilizado varios cuadernillos de cada materia, siguiendo un diseño riguroso para su elaboración, su combinación y su aplicación.

La primera aplicación del estudio tuvo lugar en 1995, produciéndose una repetición en 1999 y estando prevista una tercera toma de datos en el año 2003. La secuencia de estas aplicaciones cada cuatro años, que es un tiempo igual a la diferencia existente entre los grados escogidos, permite extraer diversas conclusiones acerca de los resultados conseguidos y de su evolución a lo largo del tiempo, que algunos países ya han comenzado a explotar, como es el caso de Estados Unidos (NCES, 2000b). En TIMSS-1995 participaron 41 países, entre los que Colombia fue el único latinoamericano, puesto que México y Argentina lo iniciaron pero no llegaron a finalizarlo. En TIMSS-1999 participaron 38 países, esta vez con dos latinoamericanos, Chile y Argentina.

2.1.2 *Estudio Internacional sobre el Progreso en Comprensión Lectora (PIRLS)*

El PIRLS (*Progress in International Reading Literacy Study*) ha sido diseñado para medir la evolución del rendimiento lector de los estudiantes. La recogida de datos se ha llevado a cabo en el año 2001, centrándose la evaluación en las destrezas lectoras de los alumnos y en las experiencias relacionadas con la lectura que viven en el hogar y en la

escuela. El estudio debe proporcionar una valiosa información comparativa sobre los niveles de comprensión lectora, que puede resultar de gran interés para la mejora de la enseñanza y el aprendizaje de la lectura.

Una de las características más destacadas de esta evaluación consiste en que se ha producido 10 años después de la realización del Estudio sobre Comprensión Lectora de 1991 (*Reading Literacy Study*), lo que permite llevar a cabo un estudio profundo de los cambios y las tendencias que se han efectuado durante ese decenio en los países participantes en lo que se refiere a la adquisición de la lectura. Para ello, el estudio ha ofrecido la opción de utilizar un módulo específico para enlazar los trabajos de 1991 y 2001. Además, el PIRLS se ha diseñado con la intención de que inicie un ciclo de estudios que mida la evolución y las tendencias en los resultados lectores de los alumnos, así como en las políticas y prácticas educativas relacionadas con la lectura, efectuándose tomas de datos en los años 2001, 2005 y 2009. De este modo se acentúa la importancia de los estudios de progreso en la dirección trazada por el TIMSS.

La evaluación está dirigida a los alumnos del curso que tenga matriculados a la mayoría de los alumnos de nueve años. En casi todos los países este curso suele ser el cuarto de la educación primaria o básica, que fue uno de los que se examinaron en el Estudio sobre Comprensión Lectora de 1991. La muestra aleatoria incluye en cada país 150 escuelas, en las que se seleccionan, también al azar, una o dos clases de cuarto curso, lo que supone una participación de aproximadamente 3.500 estudiantes de cada país. Participan en el PIRLS un total de 33 países, entre los que se incluyen Argentina, Colombia y Trinidad y Tobago.

El marco conceptual de la investigación se ha desarrollado a través de un proceso de consenso, e incluye la evaluación de una amplia variedad de procesos de comprensión implicados en la lectura de textos para dos objetivos principales: la lectura como experiencia literaria y la lectura para la adquisición y utilización de información. Sus tres elementos centrales son los procesos de comprensión, los propósitos de la lectura y las conductas y actitudes lectoras.

2.2 ESTUDIOS DE LA OCDE

La Organización para la Cooperación y el Desarrollo Económico (OCDE) inició a través de su Centro de Investigación e Innovación Educativa (CERI) un proyecto para la producción de indicadores internacionales de la educación (Proyecto INES), a finales de los años ochenta. Su propósito ha sido, desde su origen, elaborar, publicar y difundir un conjunto significativo de indicadores de la educación, relativos al contexto, las entradas, los procesos y los productos de los sistemas educativos. Los indicadores así elaborados pretenden dar respuesta a las necesidades de información de los especialistas y de los responsables políticos de la educación.

Mientras que muchos de los indicadores elaborados fueron fruto de la evolución de las estadísticas educativas, existían algunas lagunas notables en ciertas áreas, que impulsaron a los promotores del proyecto INES a desarrollar sus propios instrumentos de recogida de datos. Una de las áreas donde más se dejó sentir dicha carencia fue la correspondiente a los resultados de la educación, lo que estimuló el inicio de varios proyectos específicos.

2.2.1 Programa para la Evaluación Internacional de los Resultados de los Alumnos (PISA)

166

El Programa para la Evaluación Internacional de los Resultados de los Alumnos (*Programme for International Student Assessment - PISA*) de la OCDE (OCDE, 1999, 2000a; OCDE/MEC, 2000) está dirigido a producir sistemáticamente, a lo largo del tiempo, indicadores orientados a la información y a la toma de decisiones políticas sobre los resultados de los alumnos, de modo eficiente y con una adecuada relación entre los costos y los beneficios del proyecto. Aunque los resultados puedan ser útiles para múltiples propósitos y audiencias, el objetivo primario del programa PISA es proporcionar información basada en datos empíricos a los administradores de la educación con el objeto de contribuir a una toma de decisiones políticas más fundamentada (Nisbet, 1997).

El programa PISA se centra en la valoración de los resultados de los alumnos de 15 años, estableciendo una comparación válida y fiable tanto entre sistemas educativos como entre cohortes de alumnos dentro de un mismo sistema educativo. Los resultados, que está previsto que se

publiquen de modo cíclico cada tres años, permitirán a los administradores de la educación comparar los niveles logrados por su sistema educativo con los alcanzados por los sistemas de otros países, así como analizar la evolución de los logros de los alumnos a lo largo del tiempo. Los resultados del primer ciclo trienal fueron publicados a finales del año 2001, obteniendo una acogida muy favorable en la comunidad educativa internacional (OCDE, 2001).

PISA producirá, fundamentalmente, tres tipos de indicadores del funcionamiento de los sistemas educativos: 1. Indicadores básicos, que estarán orientados a proporcionar una línea base del perfil de conocimientos, destrezas y competencias de los alumnos, así como de los distintos subgrupos que sean de interés en función de sus características sociales; 2. Indicadores contextuales, que mostrarán la relación que existe entre los niveles de rendimiento de los sistemas y algunas variables relevantes de tipo demográfico, social, económico y educativo; 3. Indicadores de tendencias, que permitirán valorar el cambio en los resultados de los alumnos a través del tiempo, así como el grado de cambio de un sistema en particular en comparación con los cambios producidos en otros sistemas educativos.

Las pruebas se administran a alumnos de 15 años cada tres años, habiéndose realizado la primera toma de datos en el año 2000 (OCDE, 2001). La edad objetivo del estudio se ha elegido tomando como criterio la más alta en que la escolarización en los países de la OCDE es prácticamente universal. El programa PISA cubre los dominios de la lectura, las matemáticas y las ciencias, aunque su objetivo central es evaluar conocimientos, destrezas y competencias de carácter más amplio que se encuentran incorporadas en el contexto de estas áreas de conocimiento. En consecuencia, las capacidades y conocimientos no se definen en términos de los currículos de los países participantes, o de un currículo internacional común, sino en términos de las destrezas y conocimientos importantes o necesarios para la vida adulta, por lo que concede importancia a la evaluación de las competencias transversales a las materias. Este enfoque que enfatiza la valoración de las competencias transversales ha llevado a que el segundo ciclo (2001-2004) incluya, además de las matemáticas, las ciencias y la lectura, un fuerte componente de evaluación de la capacidad de solución de problemas.

La lectura ha constituido el elemento principal del primer ciclo del PISA, dado que es la destreza básica de la que depende el desarrollo del resto de competencias en otras áreas. Las matemáticas y las ciencias

constituirán, respectivamente, el elemento central del segundo y tercer ciclo del PISA (con tomas de datos en los años 2003 y 2006), dado que son áreas de especial importancia en las sociedades modernas, que cada vez dependen más de la innovación y de los avances tecnológicos. En el primer ciclo del PISA han participado 32 países, 28 de ellos miembros de la OCDE, entre los que están incluidos todos los de la Unión Europea, y cuatro no miembros de esta Organización.

2.2.2 Estudio longitudinal sobre la transición de la escuela al trabajo del PISA (PISA Longitudinal School-Work Transition Survey)

El Estudio Longitudinal del PISA (PISA-L) tiene como objetivo principal desarrollar la comprensión de los cambios que se producen a lo largo del tiempo en los estudiantes a partir de los 15 años, bien en el proceso de su acceso a estudios posteriores, bien en su incorporación al mundo del trabajo (OCDE, 2000b; Statistics Canada Consortium, 2000). Dicha comprensión incluye tanto el conocimiento de los procesos complejos y dinámicos del desarrollo de los estudiantes como el análisis de los factores que influyen en ellos. Para poder llevar a cabo tal estudio es necesario evaluar a los mismos alumnos a lo largo del tiempo y tomar múltiples datos adicionales de su contexto.

168

Este trabajo pretende monitorizar y analizar las relaciones entre el rendimiento académico de los estudiantes, sus capacidades y destrezas, y los resultados educativos posteriores o las características de su incorporación al mercado de trabajo, dentro de una concepción del aprendizaje como proceso continuo a lo largo de la vida. El seguimiento de los alumnos a lo largo del tiempo intenta responder a algunas preguntas para las que el diseño básico del programa PISA no tiene capacidad de respuesta, tales como qué factores están relacionados con el abandono escolar a partir de los 15 años, cuáles son las motivaciones para continuar estudiando en niveles superiores, hasta qué punto están los estudiantes bien informados sobre sus opciones educativas ulteriores y sobre las oportunidades en el mercado de trabajo, cuánto tiempo lleva habitualmente la transición entre el mundo educativo y el empleo y qué factores facilitan esta transición, qué consecuencias tiene la escolarización e inserción laboral a tiempo parcial, o hasta qué punto la inversión en estudios posteriores se revaloriza en términos de salario o de seguridad en el empleo, entre otras.

El PISA-L está estructurado en dos fases: la primera dirigida al seguimiento de los alumnos desde los 15 a los 19 años y la segunda encaminada a los comprendidos entre los 20 y los 24 años. El diseño de la primera fase del PISA-L implica la evaluación anual de un grupo de educandos seleccionado entre los participantes en el proyecto PISA básico, o panel de estudiantes del estudio. A este panel se le evaluará a la edad de 15 años, junto con la totalidad de participantes en el proyecto PISA, pero se tomarán referencias adicionales sobre sus actitudes, aspiraciones y planes educativos y laborales. Todos los años se recogerán datos básicos sobre educación y empleo, y cada dos se recopilará información sobre las transiciones y los factores relacionados. Las tomas de datos a partir de los 16 años se llevarán a cabo mediante entrevistas telefónicas estructuradas utilizando sistemas asistidos por computadora (*Computer Assisted Telephone Interviewing* - CATI). La segunda fase del PISA-L estará orientada al estudio más detallado de las consecuencias a medio plazo de la inversión en educación y formación profesional y de los patrones de inserción laboral.

2.3 ESTUDIOS DE LA OREALC

La Oficina Regional de Educación para América Latina y el Caribe (OREALC) de la UNESCO, creó en 1994 el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, como concreción operativa de uno de sus programas estratégicos. El Laboratorio fue concebido desde su origen como una red de sistemas de medición y evaluación de la calidad de la educación, con carácter gubernamental. Se trataba de la primera iniciativa de carácter estrictamente latinoamericano en este ámbito de los estudios comparativos internacionales del rendimiento académico.

El proyecto importante inicial del Laboratorio fue el Primer Estudio Internacional Comparativo sobre Lenguaje, Matemáticas y Factores Asociados en Tercero y Cuarto Grado. El estudio se basó en la aplicación de unas pruebas de rendimiento en matemáticas y en lengua a una muestra suficientemente amplia de alumnos de los cursos mencionados, así como de unos cuestionarios de contexto a los propios alumnos, a sus familias, a sus profesores y a los directores de sus escuelas. Las pruebas de rendimiento en matemáticas y en lengua fueron elaboradas a partir de un análisis curricular detenido y de unas matrices de objetivos curriculares realizadas por especialistas de los países participantes. La información se recogió a lo largo del año 1997, el primer informe se

publicó en 1998 (Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 1998) y el segundo en el año 2000 (Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 2000). Algunos análisis más detallados de los afectos escolares pueden encontrarse en Willms y Somers (2001).

Como se indicaba más arriba, el estudio tuvo un carácter muestral. La muestra debía ser representativa, no sólo del conjunto nacional, sino también de unos estratos poblacionales previamente definidos (megaciudades, urbano y rural; público y privado), con objeto de enriquecer el análisis posterior. Eso hizo que la muestra fuese ligeramente superior a la habitual en este tipo de estudios. En total, fueron trece los países participantes, aunque sólo se publicaron los resultados de once de ellos. Además de la medición del rendimiento, el estudio concedió una gran importancia al análisis de los factores asociados a éste, con el propósito de llegar a algunas conclusiones que permitiesen la orientación de la toma de decisiones en materia de educación. Entre ellos se contemplaban algunas variables de entrada y de proceso relativas a aspectos tales como la gestión y la política en educación, el currículo planificado y el realizado, los equipos directivos y su gestión, los profesores y su desempeño profesional, las familias y su compromiso con la educación de los hijos, los alumnos y sus características.

Uno de los principales méritos del proyecto consiste en que se trata del primer estudio internacional de rendimiento que incluye un número elevado de países latinoamericanos y caribeños. Como puede

Organización	Estudio	Áreas	Población	Instrumentos	Años	Países
IEA	TIMSS	Matemáticas, Ciencias	4º, 8º, último año de secundaria	Cuestionarios Pruebas de habilidad Videos	Cíclico TIMSS (1995) TIMSS-R (1999) TIMSS-TRENDS (2003)	41 38 50
	PIRLS	Lectura	Curso con mayoría de estudiantes de 9 años (4º)	Cuestionarios	Cíclico PIRLS (1991) PIRLS (2001) PIRLS-TRENDS (2003)	32 35
OCDE	PISA	Lectura, Matemáticas, Ciencias, Solución de problemas	Alumnos de 15 años en el año de obtención de datos	Cuestionarios	Cíclico PISA-I (2000) PISA-II (2003) PISA-III (2006)	32 40 ?
	PISA-L	Lectura, Matemáticas, Ciencias, Solución de problemas	Alumnos de 16 a 24 años	Cuestionarios Entrevistas	PISA-L (2003)	?
OREALC	EIC-LMFA	Lectura, Matemáticas	3º, 4º	Cuestionarios	1997	13

apreciarse, su número es significativamente mayor que el de participantes en los estudios de la IEA y de la OCDE. Sin embargo, no está claro el futuro de su continuación, aunque sería muy positivo que lo estuviera. En caso de seguir adelante, debería asegurar una coordinación más estrecha con otras organizaciones que trabajan en esta área en la región.

3. CONSIDERACIONES SOBRE EL APROVECHAMIENTO REGIONAL DE LA PARTICIPACIÓN EN ESTUDIOS INTERNACIONALES

Como ha podido apreciarse en el apartado anterior, los países latinoamericanos y caribeños participan muy poco en los estudios internacionales de evaluación del rendimiento académico, salvo en las iniciativas exclusivamente regionales, que, por otra parte, han sido muy escasas hasta ahora. La participación en los estudios promovidos por la IEA o la OCDE ha sido muy escasa, reduciéndose a un pequeño grupo de países, casi siempre los de mayor desarrollo social y económico. Como se indicaba en un informe reciente, esa situación se explica por motivos relacionados con el grado de desarrollo, pero tampoco puede ignorarse que existen factores de otro tipo, relativos a la voluntad política o a la prioridad concedida a la evaluación de la calidad de la educación por los responsables de la educación (Tiana, 2000, p. 16). A pesar de esta situación, son pocos quienes consideran que la participación en estudios internacionales del rendimiento sea inútil o poco interesante.

171

La pregunta que se plantea es, pues, cuáles son los motivos de ese interés nacional, frente a aquel aparente desinterés por los estudios internacionales. La respuesta es compleja y no puede ser demasiado simplista. En primer lugar, hay que tener en cuenta que la expansión de los programas nacionales de evaluación educativa es muy reciente, remontándose tan sólo a mediados de la década de los 90. Parece lógico que en esa etapa de construcción y consolidación de los sistemas nacionales de evaluación, se les dedique la mayor parte de las energías y de los recursos disponibles, siempre limitados.

En segundo lugar, también se adivina una cierta desconfianza acerca de los beneficios que se pueden extraer de la participación en estudios internacionales. Y esa desconfianza está relacionada con dos datos innegables. Por una parte, muchos de los países de la región sólo ahora están consiguiendo unos niveles satisfactorios de rendimiento educativo, lo que ha provocado que los resultados obtenidos en los

estudios internacionales hayan sido bastante decepcionantes. Por otra parte, al ser pocos los países del área participantes tampoco puede extraerse el beneficio de la comparación con los iguales, que compensaría el efecto anterior. En consecuencia, la falta de participación regional ha actuado como un factor de realimentación, disuadiendo de participar en nuevos estudios. Que esa apreciación no es insensata parece comprobarse en la alta participación regional en el Laboratorio de OREALC, frente a los estudios de la IEA o de la OCDE.

No obstante, en la actualidad parecen estar cambiando las condiciones generales. Son muchos los países que han construido sistemas sólidos de evaluación educativa y que muestran un interés creciente por la comparación internacional. Además, el atractivo que ejerce la participación en estudios de cobertura cada vez más universal es innegable en un mundo globalizado. Y la actuación de diversas organizaciones, tales como la Cumbre de las Américas, la Cumbre Iberoamericana o la OEI, ha contribuido a reforzar tal tendencia, apoyando a los países que han manifestado su voluntad de avanzar por esta vía. En consecuencia, parece que las circunstancias son más propicias que en el pasado reciente para impulsar la participación latinoamericana y caribeña en las evaluaciones internacionales.

172

Aunque todo indique que la situación actual es más favorable, no se puede dejar de lado la cuestión de fondo acerca de los beneficios que pueden esperarse de la participación en estudios comparativos internacionales. Dichos beneficios son de un doble tipo. Por una parte, los países participantes obtienen una información muy valiosa sobre el rendimiento de su sistema educativo, que puede complementar la obtenida a partir de planes exclusivamente nacionales. Y, además de información, adquieren una experiencia y una formación de sus equipos técnicos que puede repercutir en la mejora de sus propios proyectos de evaluación. Aunque este beneficio sea lateral, no puede menospreciarse su relevancia. Por otra parte, los países participantes aspiran a obtener una información comparativa sobre la situación de su sistema educativo. Y es aquí donde parece que han de darse algunos pasos complementarios.

Para asegurar la comparación regional hace falta que se cumplan dos condiciones. En primer lugar, hay que aumentar el número de países participantes en los estudios internacionales. Esta es una condición indispensable, sin la cual no puede asegurarse una comparación rigurosa. En segundo lugar, hay que reforzar la comunicación entre los países de la región, con el fin de abordar su participación coordinada.

Para poder cumplir ambas condiciones es necesario que se den algunos pasos en el seno de las organizaciones de la región que se muestran más activas en este campo. La Cumbre de las Américas, la OREALC o la OEI pueden desempeñar ese papel dinamizador y coordinador, actuando en cooperación con la IEA o con la OCDE. Pero también es necesario definir con claridad los principales objetivos de esa cooperación regional, así como las necesidades a las que deben dar respuesta. Además del trabajo de las organizaciones mencionadas, habría que estimular la creación de redes de expertos, al estilo de PREAL, que fueran construyendo el marco en el que llevar a cabo dichos proyectos.

La participación en estudios internacionales más amplios, con un acento regional, tiene más ventajas que el desarrollo de iniciativas exclusivamente regionales: permite una comparación amplia con otras regiones mundiales y con otros países, y da la posibilidad de conocer y aplicar los métodos y las técnicas más avanzados para la evaluación del rendimiento. Pero el hecho de fortalecer la dimensión regional refuerza la relevancia y la significación de la comparación y proporciona grandes beneficios a los sistemas educativos participantes. En consecuencia, la combinación de ambas perspectivas no resulta en absoluto desdeñable. Incluso desde el punto de vista de la relación que existe entre el coste y el beneficio de dicha participación, se podría aumentar la eficiencia de la inversión realizada en términos de recursos materiales y humanos.

173

Hasta ahora han sido pocas las iniciativas que se han llevado a cabo intentando combinar ambas dimensiones de análisis, la global y la regional. Sin embargo, todo apunta a que se trata de un desafío fundamental para la expansión de la evaluación educativa en la región latinoamericana y caribeña. Por ese motivo vale la pena dedicarle alguna reflexión. Aunque aquí no pretendemos ofrecer respuestas definitivas, que quizás no existan todavía, no eludimos presentar algunas posibilidades dignas de considerar seriamente en la actualidad. En el apartado siguiente se ofrece un conjunto de posibilidades de ampliación regional de los estudios internacionales, sobre las que no está demás reflexionar y profundizar.

4. OPCIONES DE AMPLIACIÓN REGIONAL EN ESTUDIOS INTERNACIONALES DEL RENDIMIENTO

De acuerdo con lo expuesto en el apartado anterior, existen diversas posibilidades para obtener el máximo aprovechamiento de la

participación regional en estudios internacionales. Entre todas ellas, nos hemos centrado en las que tienen que ver con el diseño, la realización y el tratamiento de los resultados de dichas evaluaciones. Hacemos especial hincapié en las que se refieren a las posibles ampliaciones regionales que pueden incluirse en tales estudios, entre las que hemos seleccionado las seis que estimamos más interesantes.

4.1 AMPLIACIÓN DE MUESTRAS

La ampliación de muestras suele ser el complemento habitual que los países realizan en los estudios internacionales en los que participan. Este hecho se debe a que las investigaciones internacionales sólo demandan de los países participantes la selección de una muestra de elementos (ya sean estudiantes, profesores o escuelas) de un tamaño que permita estimar, con un alto y conocido grado de certeza, los valores reales de las características de la totalidad de la población que se estudia a partir de los datos obtenidos en la muestra de la investigación (Foy, Rust, y Schleicher, 1996; Kish, 1995; Levy y Lemeshow, 1999; Lohr, 2000; Rosier y Ross, 1992; Ross, 1991; Rust, 1994). Ahora bien, dicha muestra permite obtener apreciaciones globales para el conjunto de un sistema educativo determinado, pero no consiente alcanzar, con el mismo grado de confianza y precisión, estimaciones específicas para subgrupos de la población total. Es decir, la definición del tamaño de la muestra y el muestreo que habitualmente se llevan a cabo en los estudios internacionales permiten conocer con fidelidad algunas de las características de un sistema educativo considerado en su conjunto (por ejemplo, los niveles de rendimiento de los alumnos o la formación y el grado de satisfacción del profesorado), aunque no admiten la comparación de estas características, por ejemplo, entre diferentes tipos de escuelas, distintas regiones, distintas ramas educativas, etc.

Sin embargo, para los administradores de la educación y para los investigadores educativos es fundamental el conocimiento de las características diferenciales de los distintos subgrupos que conforman un sistema educativo. Esta información diferencial (ya sea del contexto o la propia de los elementos bajo estudio) de los subgrupos de un sistema de educación aporta datos que permiten no solamente una comprensión más completa de los procesos de funcionamiento de los sistemas, sino una mayor capacidad de toma de decisiones a la vista de las disparidades entre unos grupos y otros.

El estudio de las diferencias entre subgrupos de escuelas, profesores o alumnos, es de especial importancia a la luz del principio de igualdad de oportunidades y del derecho a la educación, según el cual todos los alumnos deberían tener la oportunidad de participar a través de su escolarización en los procesos educativos y de recibir la educación en condiciones similares o equivalentes.

Habitualmente la obtención de la muestra básica de las investigaciones internacionales requiere un sistema de muestreo en el que, con objeto de que la muestra sea lo más representativa posible de un sistema educativo en su conjunto, se seleccionan los sujetos de la encuesta de modo proporcional a su ocurrencia natural en la población. De ese modo, por ejemplo, las regiones con más escuelas y más estudiantes contribuyen a la muestra con mayor número de sujetos, mientras que las de menor tamaño lo hacen con un número menor. Las ampliaciones de muestra para lograr la comparabilidad entre regiones, minorías, distintos tipos de escuelas u otras características de interés, suelen requerir sistemas de muestreo no proporcionales que implican la selección de un número suficiente de sujetos de cada categoría bajo estudio, de modo que los resultados obtenidos para cada una de ellas sean representativos de la misma, válidos, fiables y comparables con el resto de categorías en las que se subdivide la población y entre las que se realizan las comparaciones (Brewer y Hanif, 1983). Estos muestreos no proporcionales no generan ningún problema a la hora de establecer valores útiles para el conjunto del sistema, ya que para ello se lleva a cabo una ponderación de los resultados obtenidos en cada categoría, que permite que cada una de ellas influya sobre los valores de un país en su conjunto en la proporción que les correspondería en un muestreo de tipo proporcional, es decir, en la proporción en la que naturalmente aparecen en la población total objetivo de la investigación. Estas posibilidades de ampliación de muestras son aplicables tanto a los estudios de la IEA como a los de la OCDE, y han sido desarrolladas por diversos grupos de países en algunos de esos proyectos.

4.2 RECOGIDA DE DATOS ADICIONALES SOBRE LOS PATRONES DE CERTIFICACIÓN ESCOLAR Y LAS CALIFICACIONES ESCOLARES POR PARTE DE LOS PROFESORES

Cuando las investigaciones internacionales se centran en la evaluación de un área de conocimiento y en las capacidades asociadas, práctica que es habitual, por ejemplo, en los estudios sobre lectura,

matemáticas o ciencias, resulta especialmente interesante el análisis de la relación entre los patrones de certificación escolar de los países y los resultados objetivos obtenidos en la investigación, así como el estudio de las relaciones entre estos últimos y los patrones de calificación que utilizan los profesores.

Los mecanismos de certificación escolar en los diversos sistemas educativos difieren entre sí y pueden llevar a valoraciones muy distintas en unos sistemas u otros, tomando como punto de partida los mismos conocimientos o destrezas (Postlethwaite, 1996). Es común que la responsabilidad de la certificación escolar corresponda bien a los profesores de modo individual, bien a un grupo de profesores de modo conjunto, a la escuela o a tribunales de exámenes externos al centro, normalmente constituidos también por profesores. En la mayoría de los casos no suele haber una definición estándar objetiva y general para todos los alumnos de los niveles de capacidades, destrezas y conocimientos que están asociados a la concesión de una certificación o titulación determinada. La experiencia de los profesores, de los centros y de los examinadores de los tribunales de certificación, junto con su conocimiento de los niveles de rendimiento habituales de cada nivel o etapa educativa, de las especificaciones curriculares y de las tradiciones establecidas en su país, constituyen el entramado en el que se fundamentan los juicios necesarios para tomar una decisión positiva o negativa sobre la certificación escolar de cada alumno.

En un mundo cada vez más interconectado y globalizado, en el que la movilidad de las personas entre países es creciente, se hace patente la necesidad de establecer estándares internacionales, o al menos comparaciones internacionales, que clarifiquen en qué grado de conocimientos, destrezas y aptitudes están asociados en cada país con sus certificaciones escolares. La investigación comparativa internacional es, sin duda, el mejor instrumento disponible actualmente para llevar a cabo dicha clarificación, paliando la falta de estandarización internacional de las certificaciones escolares. El análisis de la relación entre los patrones de certificación usuales en cada país con los resultados objetivos obtenidos en los estudios internacionales, tiene la capacidad de constituirse en un mecanismo que permite establecer una correspondencia clara entre las certificaciones peculiares de los países y el desarrollo educativo real con el que se corresponden.

Los estudios internacionales ofrecen la posibilidad de recopilar, una vez evaluado el nivel real de los resultados educativos de los

alumnos en el marco del diseño básico del estudio, los datos correspondientes a la certificación de los estudiantes evaluados, ya sea esta certificación explícita (normalmente cuando la edad o curso de los alumnos de la muestra se corresponde con el final de una etapa educativa) o implícita (el paso de los educandos a un determinado curso o grado escolar), y analizar la correspondencia entre el nivel real de conocimientos, aptitudes, destrezas y habilidades de los alumnos y su certificación, de modo que quede definido de modo claro cuál es el significado, en términos de resultados educativos, de dicha certificación y su nivel de equiparación con las titulaciones de otros países (en el ámbito de los que participen de modo conjunto en dicho estudio de la relación entre resultados y patrones de certificación escolar).

Debido a la adecuación de las calificaciones del profesorado a los niveles de desarrollo educativo de cada aula, a los múltiples factores que los profesores tienen en cuenta a la hora de calificar a los alumnos y a las diferentes funciones que cumplen, las calificaciones de los profesores no resultan ser las más adecuadas para llevar a cabo evaluaciones externas y objetivas del rendimiento académico. Y, de modo inverso, las calificaciones externas y objetivas de rendimiento, útiles para el análisis global del rendimiento en los sistemas educativos, resultarían inadecuadas para una evaluación dirigida a la regulación de los procesos concretos de enseñanza y aprendizaje.

4.3 AMPLIACIÓN DE LOS INSTRUMENTOS DE MEDIDA

En la concepción y diseño de los estudios internacionales se tiene en cuenta que todo proyecto es limitado, tanto por las restricciones que impone la propia metodología de investigación como por la necesidad de interferir lo menos posible en el funcionamiento de las escuelas, y por la cortapisa de los presupuestos disponibles en cada país para hacer frente a este tipo de proyectos. Por ello, los estudios internacionales suelen organizarse en dos niveles que se pueden denominar básico y ampliado. En el primero se definen los elementos centrales comunes que constituyen la esencia del estudio y en los que deben participar todos los países para que tenga sentido dicho estudio en su conjunto. En el nivel ampliado se definen los elementos de la investigación, que, aun siendo de interés, son optativos en el sentido de que los países pueden tomar la decisión de participar o no en los mismos. Esta división permite a los países mantener sus gastos en un nivel mínimo mediante la participación exclusiva en el proyecto básico, o bien participar en los aspectos

complementarios del estudio con una inversión mayor de recursos (Loxley, 1992).

Es tradicional que los elementos ampliados de los estudios internacionales se clasifiquen en dos tipos: las opciones internacionales y las opciones nacionales. Las primeras son aquellas en las que se produce la participación de un grupo significativo de países, lo que hace rentable el esfuerzo de coordinación y da sentido a la presentación de sus resultados dentro del conjunto de los internacionales. Se clasifican como opciones nacionales aquellas en las que el número de países participantes es escaso, por lo que no resulta rentable su coordinación ni la presentación conjunta de resultados.

La participación en uno u otro tipo de opción, o en ambos, puede ser útil en función de intereses nacionales o regionales específicos. En este último caso, los países de una región determinada pueden definir un objeto de investigación que les resulte de especial interés dadas sus características peculiares, y que está en condiciones de constituirse en una opción internacional o en una nacional compartida por un grupo restringido de países.

Existen múltiples posibilidades de ampliación de las investigaciones internacionales mediante elementos opcionales. A continuación se comentan tres de las más habituales: la ampliación de los estudios mediante pruebas de complemento curricular, la ampliación a través de la inclusión en el estudio de otras áreas de resultados de los alumnos (en adición a los elementos básicos de la investigación), y la toma de datos y análisis de otros aspectos de la actividad educativa.

Uno de los elementos que puede incorporarse opcionalmente a los estudios internacionales centrados en áreas de conocimiento es el de las pruebas de complemento curricular en esas mismas materias. En efecto, los contenidos de las pruebas de rendimiento, bien se hayan definido a partir de un currículo internacional común, o bien en términos de conocimientos, destrezas y capacidades generales para la vida, pueden distanciarse mucho del currículo específico de un país determinado o de un conjunto de países, como también puede haber diferencias significativas en cuanto a los niveles de dificultad de las pruebas y los niveles de rendimiento esperados para determinadas poblaciones. Esas circunstancias pueden hacer interesante, si no necesario, complementar las pruebas internacionales con otras nacionales o regionales más

adaptadas a los currículos y al nivel académico de los alumnos de un país o región concretos.

El Tercer Estudio de Matemáticas y Ciencias (TIMSS), a través de su componente de análisis curricular conocido como el Estudio de la Oportunidad de Aprender en Matemáticas y Ciencias (*Study of Mathematics and Science Opportunity to Learn* - SMSO), ha establecido una metodología muy desarrollada y eficaz para el análisis curricular detallado y la concordancia entre currículos, así como para las pruebas de rendimiento (Howson, 1995; Schmidt, Jorde, *et al.*, 1996; Schmidt, McKnight, *et al.*, 1996; Schmidt, Raizen, *et al.*, 1996).

Las distorsiones que pueden producirse debido a las diferencias entre los niveles de rendimiento esperado de los alumnos y los niveles de exigencia de las pruebas de rendimiento constituyen un tema complejo. En el caso de que las pruebas no se correspondan de modo notorio con los niveles reales de rendimiento de los alumnos pueden producirse dos efectos, conocidos como efectos «techo» y «suelo». El efecto «techo» se produce cuando las pruebas de rendimiento están muy por debajo del nivel de rendimiento medio de los alumnos, lo que produce que un número significativo de ellos responda correctamente a todas o a casi todas las preguntas de las pruebas, siendo la prueba incapaz de medir los niveles altos de rendimiento, y, en consecuencia, de discriminar entre los distintos niveles de rendimiento de los estudiantes con mejores resultados. El efecto «suelo» es inverso, produciéndose cuando las pruebas son notablemente difíciles para una población determinada, de modo que un gran número de estudiantes no es capaz de responder bien a ninguna o a casi ninguna pregunta de las pruebas, siendo en este caso las pruebas incapaces de estimar los niveles de rendimiento inferiores y de diferenciar entre los distintos niveles de rendimiento de los alumnos con resultados académicos bajos. Ambos efectos afectan en parte a la relevancia de las comparaciones internacionales, pero, sobre todo, alcanzan a la utilidad de los resultados para su interpretación y uso en el ámbito nacional.

Resulta difícil establecer a priori si un conjunto de pruebas va a manifestarse como inadecuado para una población de estudiantes o para un país determinado. No obstante, los grandes estudios internacionales antes comentados implican siempre la realización de una prueba piloto, que usualmente se lleva a cabo un año antes de la obtención de datos definitiva, y que permite –con un alto grado de fiabilidad– establecer hasta qué punto los niveles de exigencia de las pruebas se

corresponden con los niveles esperables de rendimiento real en la población de estudiantes de un país. El análisis de los resultados de las pruebas piloto es, sin duda, la mejor guía para decidir si en un país concreto, o en un conjunto de países determinado, es probable que se produzcan efectos «suelo» o «techo» de modo significativo y, en consecuencia, si es adecuado o no complementar las pruebas internacionales comunes con pruebas nacionales o regionales específicas.

La ampliación de las áreas de resultados de los alumnos puede centrarse en áreas curriculares o en temas transversales al currículo. La ampliación en temas curriculares en el nivel nacional resulta de interés para aquellas materias peculiares de una nación –por ejemplo, la historia o la geografía nacionales–, o en las que es difícil establecer unos conceptos, perspectivas y contenidos internacionales comunes –por ejemplo, las ciencias sociales y especialmente la historia–, pero que, no obstante, son de especial atención y relevancia para los objetivos educativos del sistema educativo.

Sin embargo, también puede realizarse una ampliación en diversos ámbitos transversales al currículo. En efecto, gran parte de los aprendizajes que los estudiantes llevan a cabo en las escuelas no están ligados directamente a materias curriculares específicas, sino que su aprendizaje se realiza a través del conjunto de materias escolares en el proceso general de escolarización, por lo que se pueden considerar como transversales al currículo (Trier, 1994). La motivación y las actitudes hacia la enseñanza y el aprendizaje, las técnicas de estudio, las habilidades sociales y de comunicación, las capacidades de trabajo en equipo y de liderazgo, la utilización de nuevas tecnologías de la información y la comunicación y la capacidad de solución de problemas son ejemplos, entre otros, de capacidades que los alumnos desarrollan de forma transversal a los contenidos curriculares comúnmente establecidos.

El estudio sobre la Educación Cívica (*Civic Education Study - CivEd*) de la IEA (Torney-Purta, Schville y Amadeo, 1999; Torney-Purta, Lehmann, Oswald y Schulz, 2001; Lehmann, Schulz, Husfeldt y Sibberns, 2001), los componentes de alfabetización informática y de autorregulación del aprendizaje del primer ciclo del programa PISA o el componente de solución de problemas de su segundo ciclo, son exponentes del interés creciente que los administradores de la educación manifiestan por la evaluación de los resultados de los alumnos en los aspectos no curriculares del proceso de aprendizaje.

Del mismo modo que se pueden rentabilizar las estructuras y dinámicas creadas para la participación en proyectos internacionales para llevar a cabo estudios de los resultados de los alumnos en otras áreas diferentes de las centrales en la investigación, también se pueden aprovechar para el análisis de otros aspectos de la actividad educativa. Tres temas muy distintos pueden ejemplificar el amplio marco de posibilidades que se presentan con la ampliación de las investigaciones a otros ámbitos de interés educativo: la relación de las familias con la educación, los procesos educativos que se desarrollan en los centros y el estudio del profesorado.

La relación que establecen las familias de los alumnos con la educación en todas sus facetas es un factor determinante de los resultados educativos y del funcionamiento que presentan los sistemas educativos. Las actitudes hacia la educación, en especial la importancia que las familias le atribuyen como mecanismo de desarrollo e integración social, el grado y los modos de participación de los padres y tutores, tanto en el proceso de aprendizaje de sus hijos como en la gestión del funcionamiento de las escuelas, las expectativas educativas que se plantean con respecto al nivel final de educación de sus hijos, son factores de tipo general que condicionan en alto grado las propias percepciones y expectativas de los estudiantes en relación con su aprendizaje, de las que dependerán a largo plazo sus niveles de logro educativo.

El examen de los procesos educativos que tiene lugar en las escuelas es una alternativa que también debe tenerse en cuenta, especialmente en los casos en que se pretende orientar la toma de decisiones relacionadas con la enseñanza y con la organización escolar. Dentro de este ámbito una posibilidad relevante consiste en la profundización en el conocimiento de la organización escolar y de la práctica educativa de cada escuela, así como el análisis de los posibles efectos de los patrones de organización sobre el clima escolar resultante y sobre los resultados educativos de los alumnos.

Son muy diversos los aspectos de interés que se pueden considerar en el ámbito de la organización escolar, y la importancia de su análisis depende en gran medida de las características estructurales y organizativas de cada sistema educativo. No obstante, se pueden destacar como más relevantes el papel de los objetivos educativos de carácter general que definen la orientación educativa de las escuelas, las concepciones y el estilo pedagógico de sus profesores, la selección de los contenidos escolares y su secuenciación en el tiempo, las normas de

convivencia y estudio, las políticas de selección, contratación y formación permanente de los profesores, los mecanismos y el uso de la evaluación del proceso de enseñanza y aprendizaje, el tipo de relaciones que se establece con las familias y el grado en que estas pueden participar, y participan efectivamente, en el funcionamiento y gobierno de las escuelas, la información que se les proporciona sobre el progreso educativo de sus hijos, los mecanismos de conexión y cooperación entre padres y profesores, y la satisfacción del profesorado, los alumnos y las familias con el funcionamiento de la escuela.

El análisis de los patrones organizativos de las escuelas y de sus prácticas pedagógicas y la consideración de su relación con los resultados educativos de los alumnos configura la base sobre la que estudiar la eficacia y la eficiencia de las escuelas, al permitir precisar los patrones y factores organizativos que están asociados a un mayor éxito escolar o a una mayor ganancia aportada por las escuelas –o valor añadido–, en relación con los resultados esperados en función de las características iniciales de su alumnado.

El análisis de los procesos de instrucción en el seno de las clases y de su influencia sobre los resultados educativos de los alumnos presenta importantes limitaciones y plantea serias dificultades. Por un lado, la información obtenida a partir de encuestas que cumplimentan los alumnos y los profesores suele ser la única vía para el análisis de los procesos de instrucción que se llevan a cabo en el entorno restringido de las clases. Notorias excepciones a esta norma son el estudio adicional del TIMSS-video, en el que se observaron mediante grabación en video las prácticas educativas reales en un número significativo de clases (Stigler, Gonzales, Kawanaka, Knoll y Serrano, 1999; NCES, 2000a), el análisis de casos del Segundo Estudio sobre las Tecnologías de la Información en Educación –SITES, *Second International Information Technology in Education Study*–, y el Estudio sobre el Entorno de las Clases –*Classroom Environment Study*– de la IEA (Anderson, Ryan y Shapiro, 1989). Pero no pueden ignorarse las limitaciones que presentan las encuestas para obtener informaciones subjetivas, debido tanto a los fallos de memoria y a las percepciones distorsionadas por las expectativas como a los efectos de la deseabilidad social. Por otro lado, los datos obtenidos sobre las prácticas didácticas de un profesor concreto en un momento determinado del proceso educativo de sus alumnos no pueden explicar los efectos acumulados de las experiencias educativas a lo largo de todos los años de escolarización con otros profesores, ni es posible atribuir a un único profesor los logros educativos de los alumnos. Es también notorio que no

son los profesores de lengua los únicos que contribuyen al aprendizaje de la lectura y la escritura, ni los profesores de matemáticas los únicos que contribuyen a su aprendizaje.

4.4 AMPLIACIÓN CON METODOLOGÍA CUALITATIVA

Las evaluaciones internacionales han tenido, hasta el momento, una orientación básicamente cuantitativa, dada la necesidad de establecer una comparabilidad clara entre los sistemas educativos de los países participantes. No obstante, ha habido excepciones a este hábito tales como los estudios antes mencionados del TIMSS-video, del Estudio sobre el Entorno de las Clases y de los estudios de casos del Segundo Estudio sobre las Tecnologías de la Información en Educación (SITES).

La metodología de investigación con encuestas tiene indudables ventajas, entre las que cabe destacar su gran capacidad de generalización a la población de la que se extraen las muestras estudiadas, la posibilidad de analizar los fenómenos educativos tal y como se presentan en la realidad sin necesidad de intervenir a priori en los mismos, y el potencial de análisis del efecto simultáneo de múltiples factores que afectan, o pueden afectar, a la dinámica educativa. Pero no se puede ignorar que también presenta una serie de limitaciones derivadas de las características de los encuestados –sobre todo en lo que se refiere a la veracidad y fiabilidad de sus respuestas–, del propio diseño y de los instrumentos de la encuesta –originados por las restricciones de tiempo y de recursos, por la representatividad de sus preguntas y por la precisión de su medida, entre otros factores–, y del muestreo utilizado –esencialmente debido al error de medida causado por la utilización de una muestra de sujetos en vez de la totalidad de la población. Estas restricciones son la causa de su escasa capacidad explicativa a nivel individual, así como de limitaciones significativas para el análisis de los procesos educativos internos en las escuelas (Gómez, 1990; Martínez Arias, 1995; Rojas, Fernández y Pérez, 1998).

Las restricciones asociadas a la disponibilidad de recursos en términos de tiempo, financieros y de personal, imponen límites a la longitud de los cuestionarios, que a su vez influyen sobre la precisión, amplitud y exhaustividad de los estudios. Conviene señalar que una excesiva longitud de los cuestionarios implica una mayor tendencia a no responder a los mismos, o a responder de modo poco preciso, por lo que

las restricciones prácticas de tiempo limitan sensiblemente la cantidad de información que puede recogerse en una investigación.

Por otro lado, la metodología observacional de carácter cualitativo es muy adecuada para el estudio de fenómenos educativos en su entorno natural, aunque resulta muy costosa, en términos de recursos y de tiempo, para analizar un número de casos que sea representativo de una determinada población. Sin duda, la metodología observacional puede aportar datos e informaciones de alto interés explicativo, al poder considerar múltiples variables y la dinámica que se establece entre las mismas en situaciones específicas, y tener una mayor capacidad para profundizar en el análisis de los procesos y para generar hipótesis explicativas que pueden verificarse de modo más general mediante medidas cuantitativas más precisas (Stallings y Mohlman, 1990).

La metodología de investigación con encuestas que habitualmente se utiliza en las evaluaciones internacionales puede complementarse con metodologías cualitativas, dado que las encuestas también presentan limitaciones para obtener información detallada y fiable sobre situaciones naturales de carácter dinámico. Por ello, merece la pena considerar la posibilidad de ampliación de los estudios internacionales con la incorporación de estudios cualitativos complementarios (Anderson, 1992; Stenhouse, 1990).

Una posibilidad de especial interés sería la de combinar las evaluaciones del rendimiento educativo con una evaluación de las escuelas de carácter cualitativo, de modo que se pudieran analizar con detalle las características de las escuelas y del profesorado, que facilitarían tanto el éxito como el fracaso académico de los estudiantes, contribuyendo así a una clarificación de los factores de eficacia escolar (Willms y Raudenbush, 1997).

Un estudio de este tipo supondría, en un primer paso, identificar las escuelas eficaces en función de su rendimiento absoluto, tomando como base los resultados obtenidos por las mismas en el cuerpo central de la evaluación internacional (Postlethwaite y Ross, 1992), y clasificarlas en categorías de escuelas muy eficaces, de eficacia media y significativamente no eficaces, en relación con la media del conjunto de escuelas participantes en la investigación. En un segundo paso, debería analizarse con detalle la información recopilada por los cuestionarios centrales de la evaluación, con objeto de descubrir las características básicas que diferencian a las escuelas clasificadas como más o menos

eficaces y que pueden orientar las dimensiones y líneas principales de observación cualitativa. En un tercer paso, se procedería a la investigación cualitativa diferencial de las categorías de escuelas eficaces, delimitando el ámbito de las observaciones y el marco de características a observar de forma tan abierta como para integrar aquellos aspectos en los que la metodología de encuestas con cuestionarios encuentra dificultades para su observación y análisis.

De modo similar al del análisis de la eficacia de las escuelas, también podría estudiarse mediante la combinación de los estudios de rendimiento y la metodología de investigación cualitativa la eficiencia de las escuelas. Para la consideración de la eficiencia escolar es necesario tener en cuenta los resultados académicos de los alumnos de un modo relativo y no de un modo absoluto. Esto implica el análisis inicial de las diferencias entre el rendimiento esperado para los alumnos, en función de las características iniciales de su entorno familiar y social, y el rendimiento realmente obtenido por los mismos. Esta diferencia puede considerarse como el «valor añadido» por la escuela al nivel de resultados esperado para los alumnos. Se pueden considerar como escuelas eficientes aquellas que producen un mayor «valor añadido» –o «ganancia»– en los resultados educativos de sus alumnos, con independencia de su situación de partida, mientras que las escuelas eficaces serían aquellas que producen los mejores resultados de modo absoluto sin tener en cuenta ningún otro tipo de consideración (Goldstein y Thomas, 1995; Goldstein y Lewis, 1996; Goldstein, Huiqi, Rath, y Hill, 2000).

Estos estudios cualitativos podrían incluir observaciones sobre la estructura y la dinámica organizativas de las escuelas en un nivel global, considerando los aspectos de gestión de aquellas –como las políticas de personal y utilización de recursos–, los aspectos didácticos –como la definición de los programas escolares o la selección de los materiales didácticos–, y los aspectos relacionales –como la participación en las escuelas o en las normas y en el clima de convivencia en las escuelas–. También se podrían incluir informaciones cualitativas más detalladas sobre el funcionamiento interno de las clases y las prácticas instruccionales de profesores específicos, ámbitos en los que los enfoques cualitativo y observacional pueden desarrollar con amplitud su potencial frente a las limitaciones de los cuestionarios para obtener información relevante y fiable sobre los procesos que tienen lugar en tales entornos.

Este enfoque que utiliza la fusión de las perspectivas cuantitativa y cualitativa, al generar una sinergia positiva entre ambas metodologías de investigación, quizás será capaz de producir resultados de investigación ricos sobre aspectos en los que previamente los estudios comparativos internacionales han tenido dificultades para obtener conclusiones claras.

4.5 ANÁLISIS TEMÁTICOS EN PROFUNDIDAD

Otro enfoque que puede incrementar la rentabilidad de la inversión realizada en los estudios internacionales consiste en la realización de análisis en profundidad de temas específicos de interés. Son múltiples las áreas de relevancia educativa en las que es posible llevar a cabo estudios más detallados de lo que permiten los informes de resultados internacionales. La presentación de los resultados en los informes generales que producen las investigaciones internacionales es limitada, al centrarse en la exposición de los resultados globales que incluyen a todos los países y que ponen de manifiesto relaciones entre factores que son comunes a todos ellos, por lo que siempre queda lugar para elaborar indagaciones complementarias que aporten reflexiones, explicaciones e interpretaciones más profundas de algún aspecto de interés.

186

Aunque se podría elaborar una lista muy extensa de posibles temas de indagación ulterior, en este trabajo se van a comentar cinco temas que parecen apropiados para subsiguientes análisis en el marco de las evaluaciones internacionales.

El análisis de la articulación del currículo en los sistemas educativos ha sido un tema de particular interés en la investigación educativa, sobre todo en aquellos casos en los que la indagación pedagógica ha estado asociada a la generación e implantación de reformas de los sistemas de educación (Alkin, 1990; Sanders, 1990a y b; Walberg y Haertel, 1990; Willis, 1990). La IEA elaboró una distinción importante en el ámbito curricular al diferenciar entre el currículo intentado –aquel que está prescrito por las regulaciones educativas vigentes en cada país–, el currículo implementado –aquel que realmente desarrollan en sus clases los profesores y que suele ser muy similar al que se ha plasmado en los libros de texto y otros materiales didácticos complementarios–, y el currículo aprendido –aquel que los alumnos realmente adquieren e incorporan a su estructura de conocimientos–

(Bloom, Hastings y Madhaus, 1971). El análisis detallado de cómo se articulan estos tres niveles en la práctica educativa de los países merece una atención detallada, especialmente cuando los rendimientos obtenidos por un sistema distan de lo esperado de acuerdo con el currículo oficial.

En relación con la equidad en la educación destacan tres dimensiones que merecen una atenta consideración. En primer lugar, el estudio de la influencia del entorno social y familiar de los alumnos sobre sus resultados educativos tiene gran trascendencia para entender hasta qué punto el principio de igualdad de oportunidades se traduce en una equidad real en los logros educativos de los alumnos. La influencia del entorno familiar y social se constituye casi siempre en un factor modulador de la eficacia de las medidas de política educativa orientadas a la consecución de la equidad en educación. En segundo lugar, también pueden ser útiles los análisis dirigidos al examen de los efectos de la inversión en educación (OCDE, 2000c) –tanto sea esta de carácter diferencial y compensatorio para sectores educativamente desfavorecidos como inversión habitual en educación– y su traducción en términos de equidad de recursos en las escuelas y en los logros educativos de los alumnos con diferente procedencia social, cultural y económica. En tercer lugar, la consideración de los patrones de selección en el sistema educativo –es decir, el grado y los modos en que un sistema es selectivo– y los factores tanto educativos como sociales que están asociados a dicha selección, son aspectos indispensables para conocer el funcionamiento real de los sistemas y para poder generar nuevos modos de articulación de las normas de selección y promoción de los estudiantes que logren una mayor igualdad en todos los sectores de la población.

Otro asunto de la mayor relevancia consiste en la monitorización del funcionamiento de los sistemas de educación a lo largo del tiempo. Este suele ser el caso que se produce cuando se quiere llevar a cabo un seguimiento de las reformas educativas, siendo fundamental tomar datos del sistema antes de introducir los cambios, así como durante su proceso de implantación y de consolidación. Los estudios longitudinales y los análisis de tendencias constituyen las herramientas adecuadas para implantar sistemas de monitorización en el tiempo (Goldstein, 1979; Keeves, 1997a). Dado que el ciclo de estudios de la IEA y el propio diseño del proyecto PISA están concebidos para permitir la monitorización de los cambios en el rendimiento y en los patrones de funcionamiento de los sistemas educativos a lo largo del tiempo, una opción de interés consiste en analizar en detalle las evoluciones y el desarrollo tanto de cada sistema

educativo particular como de sistemas educativos de una región concreta.

Ya se mencionó en el apartado anterior, al tratar el tema de la ampliación de los estudios con enfoques cualitativos y observacionales, el interés del análisis de la eficacia y de la eficiencia escolar y se comentó un enfoque para su estudio en el nivel de las escuelas, con el objeto de clarificar la importancia de la influencia de los factores escolares relacionados con la calidad educativa. Parece claro que la profundización en el tema de la calidad escolar, a través de la consideración de la eficacia y la eficiencia de los profesores, las escuelas y los sistemas, constituye un área que producirá resultados de investigación útiles para la toma de decisiones políticas.

Por último, la incorporación de las mujeres al mundo laboral así como los cambios en la estructura familiar y en los roles de los sexos de las sociedades desarrolladas, incluido el ámbito educativo y la aspiración a la equidad educativa para mujeres y hombres, hacen que el estudio de las diferencias de sexo en los resultados educativos sea importante. En estos análisis deberían tener una consideración central las preferencias de los sexos por determinadas áreas, orientaciones curriculares y carreras educativas, así como las diferencias en la velocidad de maduración de los estudiantes, lo que implica una dificultad añadida y la necesidad de estudios minuciosos y particulares que justifiquen el desarrollo de exámenes en profundidad de este tema complejo.

4.6 PRESENTACIÓN DE RESULTADOS

En cuanto a la presentación de los resultados obtenidos en los estudios internacionales, también existen opciones que complementan a nivel nacional o regional los informes internacionales de resultados. Una posibilidad es la de elaborar informes propios de los países participantes, o bien utilizar los resultados de los estudios internacionales para la elaboración y difusión de indicadores nacionales o regionales como contribución a las iniciativas de producción de indicadores en el contexto regional.

Los informes propios pueden circunscribirse al análisis de la situación del tema central del estudio o profundizar en algún tema de especial interés, como se comentó en el apartado anterior. Dichos informes propios son el vehículo más adecuado para llevar a cabo

comparaciones regionales, comparaciones de las subpoblaciones de alumnos que conforman el sistema, y comparaciones de tipos de escuelas o de ramas de educación propias del sistema educativo concreto.

Estos informes, nacionales o regionales, tienen la cualidad de enmarcar los resultados en un entorno más próximo a la realidad política, económica y social de los países, a diferencia de los informes internacionales habituales, que presentan resultados de países de muy diferentes entornos y circunstancias. Tal hecho es importante, dado que la interpretación de resultados e indicadores depende en gran medida de los distintos contextos en los que se realizan las comparaciones, el marco de referencia interpretativo. Por ello, los informes nacionales y de ámbito regional permiten una interpretación más accesible y tienen más posibilidades de que sus hallazgos ejerzan alguna influencia sobre las prácticas y políticas educativas propias.

En este campo, las organizaciones internacionales existentes de carácter regional (el Programa de Evaluación de la Calidad de la Educación de la OEI, el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación de la OREALC, la Línea 2 –Evaluación e Indicadores Educativos– de la Cumbre de las Américas, etc.) pueden tener un papel especialmente relevante, dadas las limitaciones de las organizaciones internacionales que trabajan en el campo de la evaluación educativa y en la elaboración de indicadores (la IEA y la OCDE) para la realización de informes de carácter nacional o regional y para su posterior difusión.

ÁREA DE AMPLIACIÓN	OBJETIVO
Muestras	Estudio de las diferencias entre subgrupos
Datos sobre certificación y/o calificaciones escolares	Significado educativo y comparabilidad de las certificaciones y los niveles escolares
Instrumentos de medida	Análisis de áreas de contenido complementarias, temas transversales y temas educativos de carácter general
Metodología cualitativa	Análisis de aspectos difícilmente observables con la metodología de encuestas
Análisis temáticos	Estudios en profundidad sobre temas educativos concretos
Presentación de resultados	Presentar de modo combinado resultados obtenidos en un contexto más amplio en un contexto regional

BIBLIOGRAFÍA

ALKIN, M. C. (1990): «Curriculum evaluation models», en H. J. Walberg y G. D. Haertel (eds.): *The International Encyclopedia of Educational Evaluation*. Oxford, Pergamon.

ANDERSON, L. W. (1992.): «Observation in classrooms», en J. P. Keeves (ed.): *Methodology and measurement in international educational surveys*. La Haya, IEA.

ANDERSON, L. W.; RYAN, D. W., y SHAPIRO, B. J. (eds.) (1989): *The IEA classroom environment study*. Oxford, Pergamon.

BLOOM, B. S.; HASTINGS, J. T., y MADHAUS, G. F. (1971): *Handbook of formative and summative evaluation of student learning*. Nueva York, McGraw-Hill.

BREWER, K. R. W., y HANIF, M. (1983): *Sampling with unequal probabilities*. Nueva York, Springer Verlag.

DEGENHART, R. E. (ed.) (1990): *Thirty years of international research. An annotated bibliography of IEA publications (1960-1990)*. La Haya, IEA.

FOY, P.; RUST, K., y SCHLEICHER, A. (1996): «Sample Design», en M. O. Martin y D. E. Kelly (eds.): *TIMSS Technical Report, Volume 1*. Boston, Center for the Study of Testing, Evaluation and Educational Policy, Boston College.

GOLDSTEIN, H. (1979): *The design and analysis of longitudinal studies: their role in the measurement of change*.

GOLDSTEIN, H., y THOMAS, S. (1995): «School effectiveness and “value-added” analysis», en: *Forum*, 37, 2, pp. 36-38.

GOLDSTEIN, H., y LEWIS, T. (eds.) (1996): *Assessment: problems, developments and statistical issues*. Nueva York, Wiley.

GOLDSTEIN, H.; HUIQI, P.; RATH, T., y HILL, N. (2000): *The use of value added information in judging school performance*. Londres, Institute of Education.

GÓMEZ, J. (1990): «Metodología de encuesta por muestreo», en J. Arnau, M. T. Anguera y J. Gómez: *Metodología de la investigación en Ciencias del Comportamiento*. Murcia, Universidad de Murcia.

HOWSON, G. (1995): *Mathematics textbooks: a comparative study of grade 8 texts*. Vancouver, Pacific Educational Press.

KEEVES, J. P. (1997): «Longitudinal research methods», en J. P. Keeves (ed.): *Educational research, methodology, and measurement: an international handbook*. Oxford, Pergamon.

KISH, L. (1995): *Survey sampling*. Nueva York, Wiley.

LABORATORIO LATINOAMERICANO DE EVALUACIÓN DE LA CALIDAD DE LA EDUCACIÓN (1998): *Primer estudio internacional comparativo sobre Lenguaje, Matemáticas y factores asociados en tercero y cuarto grado*. Santiago de Chile, OREALC-UNESCO.

— (2000): *Primer estudio internacional comparativo sobre Lenguaje, Matemática y factores asociados para alumnos del tercero y cuarto grado de la educación básica*. Santiago de Chile, OREALC-UNESCO.

LEHMANN, R.; SCHULZ, W.; HUSFELDT, V., y SIBBERNS, H. (2001): *IEA civic education study: technical report*. Amsterdam, IEA.

LEVY, P. S., y LEMESHOW, S. (1999): *Sampling of populations: methods and applications*. Nueva York, Wiley.

LOHR, S. L. (2000): *Muestreo: diseño y análisis*. México, International Thompson.

LOXLEY, W. (1992): «Managing International Survey Research», en: *Prospects*, 23, 3, pp. 289-296.

MARTÍNEZ ARIAS, R. (1995): «El método de encuestas por muestreo: conceptos básicos», en M. T. Anguera, J. Arnau, M. Ato, R. Martínez Arias, J. Pascual y G. Vallejo: *Métodos de investigación en Psicología*. Madrid, Síntesis.

NCES (2000a): *Highlights from the TIMSS Videotape Classroom Study*. Washington, National Center for Education Statistics.

— (2000b): *Pursuing excellence: comparisons of international eighth-grade Mathematics and Science achievement from a U.S. perspective, 1995 and 1999*. Washington, National Center for Education Statistics.

NISBET, J. D. (1997): «Policy-oriented research», en J. P. Keeves (ed.): *Educational research, methodology, and measurement: an international handbook*. Oxford, Pergamon.

OCDE/MEC (2000): *PISA. La medida de los conocimientos y destrezas de los alumnos: un nuevo marco para la evaluación*. Madrid, Ministerio de Educación y Cultura.

OCDE (1999): *Measuring student knowledge and skills: a new framework for assessment*. París, OCDE.

— (2000a): *Measuring student knowledge and skills: the PISA 2000 assessment of Reading, Mathematical and Scientific Literacy*. París, OCDE.

— (2000b): *Call for tender for the international option of a longitudinal school-work transition survey as part of the OCDE/PISA 2003*. OCDE/DEELSA/SID/BPC(2000)5A. París, OCDE.

— (2000c): *The appraisal of investments in educational facilities*. París, OCDE.

— (2001): *Knowledge and skills for life. First results from PISA 2000*. París, OCDE.

POSTLETHWAITE, T. N. (ed.) (1996): *International encyclopedia of national systems of education*, segunda edición. Oxford, Pergamon.

POSTLETHWAITE, T. N., y ROSS, K. N. (1992): *Effective schools in Reading: implications for educational planners*. La Haya, IEA.

ROJAS, A. J.; FERNÁNDEZ, J. S., y PÉREZ, C., (eds.) (1998): *Investigar mediante encuestas: fundamentos teóricos y aspectos prácticos*. Madrid, Síntesis.

ROSIER, M. J., y ROSS, K. W. (1992): «Sampling and administration», en J. P. Keeves (ed.): *Methodology and measurement in international educational surveys*. La Haya, IEA.

ROSS, K. N. (1991): *Sampling manual for the IEA international study of Reading Literacy*. Hamburgo, IEA.

RUST, K. (1994): «Issues in sampling for international comparative studies in education: the case of the IEA Reading Literacy study», en M. Binkley, K. Rust y M. Winglee (eds.): *Methodological issues in comparative educational studies*. Washington, U.S. Department of Education, Office for Educational Research and Improvement.

SANDERS, J. R. (1990a): «Curriculum Evaluation», en H. J. Walberg y G. D. Haertel (eds.): *The International Encyclopedia of Educational Evaluation*. Oxford, Pergamon.

— (1990 b): «Curriculum Evaluation Research», en H. J. Walberg y G. D. Haertel (eds.): *The International Encyclopedia of Educational Evaluation*. Oxford, Pergamon.

SCHMIDT, W. H.; JORDE, D.; COGAN, L. S.; BARRIER, E.; GONZALO, I.; MOSER, U.; SHIMIZU, K.; SAWADA, T.; VALVERDE, G. A.; MCKNIGHT, C.; PRAWAT, R. S.; WILEY, D. E.; RAIZEN, S. A.; BROTTON, E. D., y WOLFE, R. G. (1996): *Characterizing pedagogical flow: an investigation of Mathematics and Science teaching in six countries*. Dordrecht-Boston- Londres, Kluwer Academic Publishers.

SCHMIDT, W. H.; MCKNIGHT, C. C.; VALVERDE, G. A.; HOUANG, R. T., y WILEY, D. E. (1996): *Many visions, many aims: a cross-national investigation of curricular intentions in school Mathematics*. Dordrecht-Boston-Londres, Kluwer Academic Publishers.

SCHMIDT, W. H.; RAIZEN, S. A.; BRITTON, E. D.; BIANCHI, L. J., y WOLFE, R. G. (1996): *Many visions, many aims: a cross-national investigation of curricular intentions in school science*. Dordrecht-Boston-Londres, Kluwer Academic Publishers.

STALLINGS, J. A., y MOHLMAN, G. G. (1990): «Observation techniques», en H. J. Walberg y G. D. Haertel (eds.): *The international encyclopedia of educational evaluation*. Pergamon, Oxford.

STATISTICS CANADA CONSORTIUM (2000): *A proposal submitted for the international option of a longitudinal school-work transition survey as part of the OECD/PISA 2003: Technical and Business Proposal*. Ottawa, Statistics Canada.

STENHOUSE, L. (1990): «Case Study Methods», en H. J. Walberg y G. D. Haertel (eds.): *The international encyclopedia of educational evaluation*. Oxford, Pergamon.

STIGLER, J. W.; GONZALES, P.; KAWANAKA, T.; KNOLL, S., y SERRANO, A. (1999): *The TIMSS videotape classroom study: methods and findings from an exploratory research project on eight-grade mathematics instruction in Germany, Japan and the United States*. Washington, U.S. Department of Education, National Center for Education Statistics.

TIANA, A. (1998): «The evaluation of education systems: the view at the end of the 1990s. Introduction to the open file», en: *Prospects*, vol. XXVIII, n.º 1, pp. 23-29.

— (2000): *Cooperación internacional en evaluación de la educación en América Latina y el Caribe. Análisis de la situación y propuestas de actuación*. Washington, Banco Interamericano de Desarrollo–Departamento de Desarrollo Sostenible.

TORNEY-PURTA, J.; SCHWILLE, J., y AMADEO, J. (1999): *Civic education across countries: twenty-four national case studies from the IEA Civic Education Project*. Amsterdam, IEA.

TORNEY-PURTA, J.; LEHMAN, R.; OSWALD, H., y SCHULZ, W. (2001): *Citizenship and education in twenty-eight countries: civic knowledge and engagement at age fourteen*. Amsterdam, IEA.

TRIER, U. (1994): *Establishing indicators for cross-curricular competencies in a comparative setting*. Paper presented at the annual meeting of the American Educational Research Association, Nueva Orleans, 4-8 Abril.

WALBERG, H. J., y HAERTEL, G. D. (1990): «Curriculum Evaluation: Introduction», en H. J. Walberg y G. D. Haertel (eds.): *The International Encyclopedia of Educational Evaluation*. Oxford, Pergamon.

WILLIS, G. (1990): «Qualitative curriculum evaluation», en H. J. Walberg y G. D. Haertel (eds.): *The International Encyclopedia of Educational Evaluation*. Oxford, Pergamon.

WILLMS, J. D., y RAUDENBUSH, S. W. (1997): «Effective schools research», en J. P. Keeves (ed.): *Educational research, methodology, and measurement: an international handbook*. Oxford, Pergamon.

WILLMS, J. D., y SOMERS, M. A. (2001): *Schooling outcomes in Latin America*. Report. Canadian Research Institute for Social Policy, Universidad de Nueva Brunswick.