

Catalina Wainerman
María Mercedes Di Virgilio
(Compiladoras)

EL QUEHACER DE LA
INVESTIGACIÓN EN EDUCACIÓN

MANANTIAL
Buenos Aires

 Universidad de
San Andrés

Diseño de tapa: Eduardo Ruiz

Catalina Wainerman

El quehacer de la investigación en educación / Catalina Wainerman y María Mercedes Di Virgilio. - 1a ed. - Buenos Aires : Manantial, 2010.
328 p. ; 16x23 cm.

ISBN 978-987-500-145-9

I. Investigación Educativa. I. Di Virgilio, María Mercedes II. Título
CDD 370.1

Hecho el depósito que marca la ley 11.723
Impreso en la Argentina

© 2010, Ediciones Manantial SRL
Avda. de Mayo 1365, 6° piso
(1085) Buenos Aires, Argentina
Tel: (54-11) 4383-7350 / 4383-6059
info@emanantial.com.ar
www.emanantial.com.ar

Derechos reservados

Prohibida la reproducción parcial o total, el almacenamiento, el alquiler, la transmisión o la transformación de este libro, en cualquier forma o por cualquier medio, sea electrónico o mecánico, mediante fotocopias, digitalización u otros métodos, sin el permiso previo y escrito del editor. Su infracción está penada por las leyes 11.723 y 25.446.

Índice

Prólogo.....	11
Introducción. Problemas que enfrenta la producción de investigación científica en educación en la Argentina, <i>Catalina Wainerman</i>	17
Parte I ¿QUÉ Y CÓMO SE INVESTIGA?	
Prefacio a la parte I	31
<i>Sección 1</i> <i>Las políticas educativas como objeto de investigación</i>	
Capítulo 1. La influencia de las agencias internacionales en las reformas de la formación docente en la Argentina y Brasil (1985-2002): un análisis comparado, <i>Jason Beech</i>	39
Capítulo 2. Investigando políticas educativas: reflexiones en torno a un estudio sobre la reforma del nivel secundario, <i>Jorge M. Gorostiaga</i>	57
Capítulo 3. Los horizontes de la investigación en políticas educativas: los niveles intermedios y la ampliación del campo, <i>Silvina Gvirtz</i>	67

Sección 2

La evaluación del sistema, las instituciones y los actores como objeto de investigación

Capítulo 4. Las políticas de evaluación universitaria en la Argentina: consideraciones metodológicas en torno a las instituciones, los actores y sus prácticas, *Antonio Camou y Marcelo Prati* 81

Capítulo 5
La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas, *Pedro Ravela* 107

Capítulo 6. La problemática de la investigación didáctica: el caso de una investigación sobre formatos de evaluación de los aprendizajes y sus relaciones con las modalidades de estudio de los alumnos universitarios, *Alicia Camilloni y Estela Cols* 127

Sección 3

La escuela como objeto de investigación

Capítulo 7. La educación física escolar en la Argentina: conformación, permanencias y cambios de una matriz disciplinar (1880-1960), *Ángela Aisenstein* 145

Capítulo 8. La biografía escolar en el desempeño profesional de docentes noveles. Proceso y resultados de un trabajo de investigación, *Andrea Alliaud* 163

Capítulo 9. Eficacia escolar y etnografía en escuelas de la ciudad de Buenos Aires, *Cristina Carriego* 179

Sección 4

Condiciones de producción de la investigación científica en educación

Capítulo 10. Interacciones entre campo académico, investigación educativa y la construcción de agendas de política educativa, *Claudio Suasnábar* 199

Capítulo 11. La producción y el uso de los datos para la investigación educativa, *Irene Oiberman* 225

Parte II

REFLEXIONES ACERCA DE CÓMO SE INVESTIGA

Prefacio a la parte II 247

Capítulo 12. La(s) política(s) educativa(s) como objeto de investigación: formas de construcción y de abordaje, *María Mercedes Di Virgilio* 251

Capítulo 13. La evaluación como objeto de investigación y las investigaciones sobre evaluación como objeto de reflexión, *Jorge Fasce* 259

Capítulo 14. La escuela como objeto de estudio: tres posibles miradas, *Alejandra Navarro* 265

Capítulo 15. Condiciones de producción de la investigación científica en educación: restricciones, dilemas, decisiones, problemas y desafíos, *Gustavo Dufour* 275

Parte III

PROPUESTAS PARA EL DESARROLLO DEL CAMPO

Prefacio a la parte III 291

Capítulo 16. ¿Educación o problemas educativos? Reflexiones sobre la construcción del campo educativo, *Ana García de Fanelli* 293

Capítulo 17. La investigación en educación y la toma de decisiones en políticas educativas, *Claudia Jacinto* 303

Capítulo 18. Sugerencias para el desarrollo de la investigación científica en educación, *Ruth Sautu* 309

Sobre los autores 321

Capítulo 5

La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas

Pedro Ravela

El contenido del presente capítulo responde al propósito de este libro: ofrecer una mirada a “la trastienda de la investigación”, es decir, a los aspectos vinculados al itinerario de decisiones, de marchas y contramarchas, de dudas y problemas reales que debe enfrentar y resolver quien hace investigación. Por esta razón, para escribir este texto he optado por utilizar un estilo narrativo en primera persona, a veces en singular, la mayor parte en plural (la investigación es sobre todo una tarea colectiva). Considero que esto es lo más apropiado para la finalidad del libro, ya que una presentación formal de la investigación supondría presentar un producto acabado, que enfatice la coherencia entre todos los aspectos. Normalmente, un informe de investigación no da cuenta de las vicisitudes enfrentadas durante el proceso de su realización.

LA BIOGRAFÍA DEL TEMA DE INVESTIGACIÓN

El primer aspecto que quiero abordar es el relativo a cómo llegamos a definir el tema de la investigación. En general, el tema de una investigación es presentado y justificado en términos teóricos: se presenta una revisión de la literatura y la investigación en el campo de que se trate y se explica la relevancia del tema elegido para el estado del arte. Si bien esto es imprescindible, no da cuenta de la totalidad de factores que intervienen en la construcción de un tema de investigación. La elección y construcción del tema tiene relación con al menos otros dos aspectos relevantes:

- a. la acumulación previa del investigador (algo así como su “biografía temática”); y
- b. el modo en que el investigador concibe su papel en la sociedad y su contribución a la misma (sus intereses y propósitos).

En el caso de la investigación que nos ocupa, el tema tiene una larga historia de acumulación previa en mi trayectoria personal y en la del equipo de investigación. Hace unos veinticinco años, mientras cursaba la Maestría en Ciencias Sociales y Educación de FLACSO con Cecilia Braslavsky, uno de los trabajos finales de curso que realicé fue acerca del uso por parte del docente de la evaluación como instrumento de poder en el aula. Durante los años noventa, luego de finalizada mi maestría, me dediqué a trabajar en temas de evaluación educativa a gran escala, a través de pruebas estandarizadas. Estuve a cargo de la construcción y el desarrollo del sistema de evaluación en mi país. Entre 2002 y 2006 coordiné la participación de Uruguay en el Programa Internacional para la Evaluación de Estudiantes (PISA). Entre 2005 y 2006 participé de la coordinación del Segundo Estudio Regional Comparativo y Explicativo (SERCE) de la UNESCO, que evaluó aprendizajes de escolares de primaria en dieciséis países de la región. Desde hace diez años formo parte del Grupo de Trabajo sobre Estándares y Evaluación (GTEE), una red latinoamericana de intercambio en torno al tema, que me ha dado la oportunidad de conocer y acompañar experiencias en distintos países, así como de investigar y escribir sobre esta cuestión. He realizado además actividad docente sobre temas de evaluación educativa. A lo largo de esta peripecia fuimos estableciendo relaciones con colegas, compartiendo ideas y trabajos, conformando un equipo de trabajo en Uruguay y redes de intercambio con muchos colegas de la región, cada cual con su propia trayectoria personal en el tema de la evaluación.

Un segundo aspecto relevante es el que atañe a la perspectiva que uno tiene sobre su rol en la sociedad. La peripecia brevemente reseñada está relacionada con la evaluación a gran escala. Sin embargo, a lo largo de este camino una de nuestras principales preocupaciones fue siempre la utilidad de dichas evaluaciones para los docentes. De allí que, por ejemplo, en Uruguay hayamos desarrollado un sistema fuertemente orientado a la devolución de resultados a las escuelas (con carácter confidencial y sin "rankings"), así como a la producción de textos de orientación didáctica y al desarrollo de programas de formación en servicio, a partir de los resultados de las evaluaciones nacionales estandarizadas (Ravela, 2008, págs. 74-86). En la base de la concepción de la investigación que estoy reseñando hay ciertas convicciones personales que contribuyeron a dar forma a dicha investigación.

En primer término, la convicción de que el investigador y el evaluador deben asumir responsabilidad por los efectos que la propia tarea tiene en la sociedad y en el sistema educativo. Su labor no se limita a producir conocimiento o información, sino que incluye también el involucramiento en acciones que aseguren su divulgación y uso. En segundo término, que uno de los desafíos más importantes para mejorar la educación es mejorar el trabajo en el aula. Por lo tanto, que la investigación y la evaluación educativas deben encontrar la forma de adquirir significado para el trabajo docente.

De esta manera, a lo largo de la historia personal y colectiva en torno a la evaluación, ciertas preocupaciones fueron tomando forma. ¿Qué información reciben los maestros de las evaluaciones a gran escala? ¿Llegan a leer sus resultados? ¿Los comprenden? ¿Les resultan de alguna utilidad para reflexionar sus prácticas y modificarlas? (Ravela, 2006a).

Hacia fines del año 2006, en el seno del GTEE fue surgiendo la idea de realizar un estudio acerca de las prácticas de evaluación de los maestros dentro del aula y de ampliar el foco de trabajo del grupo, más allá de la evaluación estandarizada. Los ejes principales de la temática se fueron plasmando en los siguientes términos:

- a. Si bien los sistemas estandarizados de evaluación de los aprendizajes aportan información relevante para el debate ciudadano, para la rendición de cuentas y para la discusión de las políticas educativas, la evidencia derivada de los estudios de seguimiento de dichos sistemas muestra que su impacto sobre las prácticas de enseñanza y de evaluación en las aulas ha sido limitado.
- b. Buena parte de los maestros no están preparados en forma adecuada para comprender los resultados de las evaluaciones externas, ni para diseñar evaluaciones apropiadas para su uso en el aula.
- c. Las prácticas de evaluación en el aula suelen ser rutinarias y estar centradas en la calificación.
- d. De lo anterior se desprende la relevancia de estudiar:
 - Los enfoques y propuestas que los maestros utilizan para evaluar a sus estudiantes.
 - El modo en que emplean los resultados de sus propias evaluaciones para dar orientación a sus estudiantes y ayudarles a progresar.
 - El modo en que emplean dichos resultados para comunicarse con las familias acerca del progreso educativo de sus hijos.
 - Los criterios que utilizan a la hora de tomar decisiones respecto a la promoción o reprobación del curso por parte de sus estudiantes.
- e. Un estudio de enfoques y prácticas de la evaluación en el aula permitiría indagar qué tipo de desempeños esperan los maestros que sus alumnos alcancen y en qué medida incorporan en sus evaluaciones lo que estipulan los currículos nacionales (currículo intencionado *versus* currículo implementado y alineamiento entre currículo, evaluaciones externas y evaluaciones internas).

LA BIOGRAFÍA DEL DISEÑO DE LA INVESTIGACIÓN

Al igual que el tema de la investigación, su diseño y metodología se nutren de una acumulación previa de ideas. Para esta investigación deseo destacar

dos. En primer término, como coordinador de evaluaciones a gran escala (nacionales, PISA, y SERCE), estuve vinculado a los denominados estudios de "factores asociados", trabajos de investigación cuyo propósito es analizar qué variables escolares tienen influencia sobre los aprendizajes de los estudiantes.

La mayor parte de los estudios de este tipo se limitan a analizar correlaciones entre variables, empleando técnicas estadísticas cada vez más sofisticadas. Sin embargo, su poder explicativo suele ser limitado, por una razón muy sencilla: muchos de los factores más importantes no pueden ser captados a través de cuestionarios y encuestas. Una muestra de ello es la calidad del trabajo del docente. Los estudios de factores asociados sólo se aproximan a este factor a través de variables que pueden ser relevadas a partir de encuestas como, por ejemplo, la titulación del docente o la cantidad de cursos de capacitación en servicio que realizó. Como resulta obvio, cualquiera de estas dos variables aporta muy poca información, dado que las propuestas de formación inicial y en servicio para docentes incluyen una enorme heterogeneidad en cuanto a su calidad. No debe, pues, sorprender que el poder explicativo de estas variables resulte limitado.

Como consecuencia de lo anterior, los denominados estudios de factores asociados suelen reiterar "hallazgos" muy similares en torno a la importancia de algunos factores que pueden ser medidos con razonable validez y confiabilidad a través de encuestas: el clima en el aula y la existencia de altas expectativas y visiones compartidas en relación con los aprendizajes de los estudiantes, entre otros.

A partir de lo anterior, en el grupo de evaluación de Uruguay fuimos comprendiendo la importancia de complementar la mirada cuantitativa y la cualitativa. Un estudio a gran escala proporciona un mapa de situación del sistema, en cuyo marco es posible realizar una selección más inteligente y relevante de casos en los cuales realizar estudios cualitativos. Esta visión se plasmó en un trabajo publicado en 1999, que incluyó el análisis estadístico de los factores escolares que tenían incidencia en los aprendizajes en escuelas de contextos sociales desfavorecidos, junto con un análisis en profundidad de diez casos de escuelas de dichos contextos con distinto tipo de resultados en las pruebas (Unidad de Medición de Resultados Educativos, 1999). El estudio cualitativo aportó una visión más comprensiva sobre aspectos que el estudio estadístico no podía captar.¹

1. El problema de los estudios de factores asociados es en realidad más complejo. Incluye tanto la cuestión de cuál es la variable dependiente (una sola medida de logro educativo o el "valor agregado", que implica más de una medición en el tiempo) como la cuestión del diseño (diseños de tipo experimental o cuasi experimental son más potentes para construir conocimiento en este campo) (Martínez *et al.*, 2009; Carnoy, 2008; Shadish *et al.*, 2002).

La segunda idea fuerza que se fue conformando a partir de la historia personal y que subyace a este estudio fue la importancia de los estudios internacionales comparados como fuente de una comprensión más rica de las situaciones y las políticas nacionales. La participación en PISA fue muy importante para esto. Pero, un estudio especialmente inspirador fue el realizado por TIMSS a partir de la filmación de clases de matemática en los Estados Unidos, Japón y Alemania (National Center for Education Statistics, 2000 y 2003). Estos trabajos muestran que es posible encontrar un patrón de enseñanza característico de un país, con una muestra relativamente reducida. Hay elementos de la cultura profesional y de las prácticas establecidas social e históricamente por el grupo docente, que conforman una estructura básica de los modos de enseñar en un país, más allá de las diferencias individuales que, por cierto, no desaparecen.

Con estas ideas como trasfondo, en los dos años que estuve en la coordinación técnica del SERCE intenté impulsar la idea de realizar, además del clásico estudio de factores asociados, dos estudios en profundidad en países latinoamericanos, seleccionados en función de sus resultados y de sus contextos culturales: uno en torno a las políticas educativas implementadas a lo largo de una década y otro en torno a las prácticas de enseñanza en las aulas de primaria. Como se verá enseguida, en estas ideas estaba el germen de la investigación que nos ocupa aquí.

LA CONCRECIÓN Y FORMULACIÓN DEL PROYECTO

Una cuestión crucial para la definición final, tanto del tema como del diseño de una investigación, es la relativa al acceso a los recursos económicos indispensables para llevarla adelante. Pocas veces se da la situación de conseguir el financiamiento para avanzar con un proyecto tal como uno lo concibió, sin restricciones. Normalmente conseguir recursos implica restricciones para la construcción del tema y del diseño del estudio.

Conseguir fondos para investigación requiere establecer vínculos, formar parte de redes, presentarse a concursos públicos y/o buscar oportunidades (por ejemplo en Internet), entre otras cosas. Supone formular proyectos que muchas veces no se concretan, tener ideas en mente y ser capaz de formularlas de una manera apropiada a las oportunidades que se presentan. Esto último implica una cierta capacidad de "negociación" entre lo que uno querría hacer, lo que le interesa a quien financia y las restricciones que imponen los fondos disponibles. Una capacidad importante del investigador es la de formular una propuesta de investigación que sea relevante, a la vez, para el investigador, para el que financia y para la sociedad, y que se adecue a los recursos disponibles sin perder su calidad técnica.

Para un estudio regional de tipo comparado, el primer problema son los costos de viajes. De allí que la primera disyuntiva que enfrentamos fue la siguiente: ¿invertimos en cubrir mayor cantidad de países, destinando menos

tiempo al trabajo de campo en cada país o, por el contrario, intentamos un trabajo de campo de mayor amplitud en un número de países más reducido? Idealmente podríamos haber pensado en ir a la totalidad de los países participantes del SERCE, para tener un panorama de toda la región. Pero esto no era viable en términos económicos, a menos que contratásemos a una persona en cada país para hacer el trabajo de campo. Esta opción fue descartada porque, o bien corríamos el riesgo de perder calidad en el relevamiento de información (si el entrenamiento de las personas se hacía a distancia), o bien había que gastar en viajes de entrenamiento. De modo que optamos porque el trabajo de campo fuese realizado por los integrantes del equipo de investigación (tres de cinco) y definimos que el estudio se realizaría en cinco países que representarían situaciones diversas en términos culturales y de resultados educativos (analizamos las alternativas de ir a cuatro o a seis países).

La tercera decisión, tomada en forma paralela a las dos anteriores, fue que el trabajo de campo en cada país tendría una duración de una semana (cinco días hábiles). Esto optimizaba el costo de viaje (se viajaba un domingo y se regresaba el sábado siguiente) y permitía pensar en trabajar con veinte maestros en cada país: dos escuelas cada día, una por la mañana y otra por la tarde, y dos maestros en cada escuela ($2 \times 5 \times 2 = 20$), lo que suponía contar con un par de horas para trabajar con cada maestro. Optamos además por focalizar el estudio en 6° grado y en las áreas de Lengua y Matemática. Incluir más grados hubiese implicado demasiada diversidad para veinte casos por país e incluir más áreas hubiese significado un exceso de temáticas para dos horas de relevamiento. Una cantidad mayor de maestros hubiese implicado un volumen de información demasiado grande para procesar.

¿Qué era posible hacer en esas dos horas? ¿Qué tipo de relevamiento de información podíamos realizar para optimizar ese tiempo? ¿Tenía sentido? ¿Sería posible recoger información relevante en este marco?

A los efectos de maximizar la recogida de información relevante, decidimos emplear tres técnicas:

- a. una entrevista en profundidad a cada maestro;
- b. un registro fotográfico de trabajos de evaluación realizados por los estudiantes;
- c. un cuestionario autoadministrado que el maestro respondería mientras se realizaba el registro de trabajos de los alumnos.

La entrevista en profundidad tendría alrededor de una hora de duración y estaría destinada a preguntar y escuchar acerca de las perspectivas del maestro en cuanto a la evaluación: ¿qué sentido tiene para él la evaluación? ¿Cómo la realiza y registra? ¿Qué tipo de consignas e instrumentos emplea? ¿Qué tipo de devolución hace a los estudiantes? ¿Qué aspectos prioriza a la hora de evaluar la lectura y la escritura? ¿Qué aspectos prioriza a la hora de evaluar matemá-

tica? ¿Qué criterios tiene en cuenta a la hora de calificar a sus estudiantes? Las entrevistas se realizaron siguiendo una pauta flexible, pero que asegurara cubrir los temas principales con cada maestro. El registro se hizo mediante grabación digital, de modo que se conformó una biblioteca de archivos de audio para que pudieran ser escuchados directamente en la computadora.

Por este medio podríamos obtener información sobre las creencias, convicciones y percepciones de los docentes sobre la evaluación. Adicionalmente, una primera aproximación a sus prácticas, al menos, a lo que los docentes creen que hacen o creen que está bien hacer. Pero, dado que estas percepciones tienen carácter subjetivo y a veces normativo (el docente responde en función de lo que considera "correcto"), el relevamiento de trabajos de los estudiantes aportaría una fuente de evidencia adicional contra la cual contrastar el discurso docente.

Nuevamente, esta idea estuvo inspirada en otros trabajos, en particular, en un trabajo de Santiago Cueto, en Perú, en el que se hace un interesante análisis de cuadernos de alumnos (Cueto *et al.*, 2003). Los cuadernos permiten ver qué tareas propone el docente a sus estudiantes. En nuestro caso, decidimos solicitarle al docente que él mismo seleccionase algunas tareas o consignas de evaluación que hubiese propuesto a los estudiantes. Le pedimos además que seleccionase, para cada tarea o consigna, tres ejemplos de trabajos realizados por buenos alumnos, tres intermedios y tres pertenecientes a alumnos de bajo rendimiento. También se les solicitaría muestras de planillas de registro u otro tipo de instrumento que emplease en la evaluación de sus estudiantes. El registro de estos materiales se realizó utilizando fotografía digital.

El cuestionario autoadministrado estuvo destinado únicamente a recoger información sobre el grado de conocimiento y uso de los resultados de evaluaciones estandarizadas por parte de los docentes. Se tomó esta decisión como forma de optimizar el uso del tiempo en la entrevista.

LOS PRODUCTOS DE LA INVESTIGACIÓN

Otro aspecto que debe ser considerado en la etapa de construcción del proyecto de investigación es el relativo a los productos de ésta. Es importante preguntarse desde el comienzo: con estas preguntas de investigación, estos objetivos y este trabajo de campo, ¿qué vamos a poder producir que consideremos relevante como aporte al campo educativo, que sea de interés para los financiadores y acorde con nuestra perspectiva acerca de los problemas educativos? Es decir, es importante pensar no solamente en términos de temas, preguntas, hipótesis y objetivos de la investigación, sino también en términos del tipo de productos que van a resultar de ella, a quiénes estarán dirigidos y por qué pueden resultarles de interés.

En el caso que nos ocupa, definimos cinco productos que consideramos relevantes para distintas finalidades y realizables en el marco del financiamiento

disponible (desde ya, la información recogida es suficientemente rica como para continuar investigando y produciendo conocimiento).

1. Una colección de propuestas "ejemplares" de evaluación en Lengua, elaboradas por maestros de 6° de primaria de distintos países, analizadas y seleccionadas por especialistas en didáctica de la lengua, presentadas con un formato adecuado para ser puestas a disposición de los docentes en la web.
2. Lo mismo para matemática.
3. Un documento de unas veinticinco páginas con un análisis comparativo de las formas de evaluar en el aula entre países y su relación con los marcos curriculares del SERCE, el currículo prescripto y las pruebas nacionales.
4. Un documento de unas veinticinco páginas en el que se analiza el modo en que las evaluaciones son utilizadas por los maestros para orientar a sus alumnos, informar a las familias y tomar decisiones de calificación y aprobación.
5. Un documento de unas veinticinco páginas con recomendaciones acerca de cómo las evaluaciones nacionales pueden usarse para enriquecer las evaluaciones en el aula y, viceversa, cómo las evaluaciones estandarizadas pueden enriquecerse con la experiencia de los maestros en las aulas.

Los productos definidos denotan una preocupación por la comunicación de los resultados y su accesibilidad para diferentes usuarios, lo que no significa que hayamos descartado otras opciones más orientadas a los formatos académicos. Algunos productos están más orientados a la construcción de conocimiento conceptual, en tanto otros están más orientados al desarrollo del conocimiento práctico de los docentes (Tenti Fanfani, 2001). También en este aspecto intentamos plasmar ideas adoptadas de otros trabajos en torno a la acumulación e intercambio de conocimiento práctico entre docentes a través de Internet (Hiebert *et al.*, 2002).

LA SELECCIÓN DE PAÍSES, ESCUELAS Y MAESTROS

En un estudio cualitativo en el que se trabaja con una cantidad reducida de casos y con muestras seleccionadas en forma intencional, es muy importante definir criterios sustantivos para la selección de los casos, de modo de maximizar la relevancia de la información que se recogerá.

Para este estudio seleccionamos los países teniendo en cuenta los resultados en el SERCE y la diversidad regional (cono sur, área andina, Centroamérica y Caribe) (véase cuadro 1). La intención nunca fue cubrir todas las casillas

del cuadro 1, dadas las restricciones económicas. Por otra parte, en función del interés específico del Programa de Promoción de la Reforma Educativa de América Latina y el Caribe (PREAL), se incluyó una mayor cantidad de países de Centroamérica.

CUADRO I
SELECCIÓN DE PAÍSES SEGÚN REGIÓN Y RESULTADOS EN EL SERCE

Resultados en el SERCE	Región		
	Cono Sur	Andina	Centroamérica y Caribe
Altos	Uruguay		Costa Rica (antes Cuba)
Medios			El Salvador
Bajos		Perú	Guatemala

En cuanto a las escuelas, se optó por trabajar en contextos urbanos desfavorecidos, de preferencia, en escuelas públicas. Las escuelas rurales fueron descartadas por razones prácticas. Adicionalmente, se estableció el criterio de que las escuelas debían tener resultados educativos en el promedio del país o por encima (en principio en el SERCE, pero luego nos vimos obligados a flexibilizar este criterio y aceptar el uso de datos de las pruebas nacionales como forma de selección). El propósito de este criterio fue evitar ir a escuelas en las que se trabajase mal o en las que directamente no hubiese enseñanza. Deliberadamente buscamos un conjunto de maestros que, en cada país, reflejase lo que hace el maestro medio o bueno, y las formas en que son evaluados los niños de sectores populares por dichos maestros.

Finalmente, se optó por 6° grado y no por 3° (los grados evaluados en el SERCE), porque en la mayor parte de los países de la región esa evaluación corresponde al final de la educación primaria y suelen aplicarse pruebas estandarizadas.

LA REALIZACIÓN DEL TRABAJO Y LOS CAMBIOS SOBRE LA MARCHA

Una vez confirmado el financiamiento para realizar el trabajo, el primer paso fue constituir un equipo de trabajo. El equipo tuvo un núcleo central de tres investigadores y dos especialistas asociados, uno en Lengua y otro en Matemática. Estos últimos tuvieron a su cargo el análisis de las propuestas de los maestros y sus devoluciones a los estudiantes, desde la perspectiva de la didáctica de la disciplina, con el fin de construir las colecciones de propuestas de evaluación definidas en los productos 1 y 2.

Los integrantes del equipo central nos hicimos cargo de la construcción de un marco conceptual, de la realización del trabajo de campo, de la organización de la información y del análisis de la misma para la elaboración de los productos 3, 4 y 5. También nos ocupamos de la interacción con los especialistas en la producción de sus análisis.

El inicio del trabajo (enero de 2008) estuvo marcado por dos tareas perentorias:

- a. La realización de una revisión de la literatura sobre la temática del estudio y la construcción de un marco conceptual que orientara el conjunto del trabajo, en particular la construcción de instrumentos y los criterios para la organización y el análisis de la información.
- b. La toma de contacto con los países seleccionados para asegurar el acceso a los mismos y a las escuelas. Ésta es una de las cuestiones más complejas en un estudio de tipo comparado, en especial si no se cuenta con vínculos con los países.

En nuestro caso, el acceso a los países y a las escuelas fue relativamente sencillo en la mayoría de los casos, en función de los vínculos previos con las unidades nacionales de evaluación. En El Salvador, Guatemala y Perú fue muy fácil. Nos contactamos con los responsables de las unidades de evaluación, a quienes conocíamos a nivel personal. Nos solicitaron una carta formal para presentar ante el ministro y automáticamente nos ofrecieron el acceso a las escuelas, datos y apoyo para seleccionarlas en función de los criterios estipulados y, en algunos casos, un transporte del Ministerio para toda la semana.

Con Cuba y Uruguay tuvimos mayores dificultades. En Cuba el pedido de autorización para viajar a la isla y visitar escuelas estuvo varios meses a la espera de resolución del ministro, hasta que finalmente, dado que se nos terminaron los plazos para la realización del trabajo de campo, debimos sustituir Cuba por Costa Rica, país centroamericano con buenos resultados en el SERCE. En Costa Rica contamos con total apoyo, al igual que en el resto de Centroamérica.

En Uruguay, nuestro propio país, tuvimos problemas parecidos. El expediente de autorización para acceder a las escuelas y realizar el estudio se perdió varias veces y conseguirlo llevó varios meses. También en este caso estuvimos a punto de cambiar de país, pero dado que el trabajo de campo no implicaba viaje y no era necesario hacerlo en una sola semana, finalmente pudimos realizarlo.

El otro cuello de botella en un estudio de este tipo es establecer el contacto con las escuelas y lograr que sean adecuadamente informadas y que acepten participar. También en este caso contamos con el apoyo invaluable de las unidades de evaluación de todos los países participantes, que establecieron un contacto inicial para sondear la disposición a participar, hicieron llegar a cada escuela una carta, elaborada por el equipo de investigación, explicando el sentido del estudio, qué se haría en la escuela y qué tipo de colaboración

necesitaríamos. En buena parte de los casos, a partir de este primer contacto se nos facilitó una dirección de correo electrónico de la escuela o de su director. También recibimos ayuda en la organización de los itinerarios para poder ir a dos escuelas cada día, en función de su ubicación.

LA CONSTRUCCIÓN DEL MARCO CONCEPTUAL Y LOS INSTRUMENTOS

En el proceso real de implementación de todo trabajo de investigación, los procesos de trabajo se entremezclan. No se sigue un orden rigurosamente lógico, del estilo de pasos o etapas de la investigación. La construcción del marco conceptual, la organización de la logística para el trabajo de campo y la construcción de los instrumentos, se superponen. En este caso, era necesario construir un documento conceptual que desarrollara las ideas básicas contenidas en las preguntas y propósitos del estudio. En paralelo, trabajamos en la construcción de los tres instrumentos principales del estudio: la pauta para la entrevista en profundidad, el protocolo para la selección y registro fotográfico de consignas y trabajos de evaluación, y el cuestionario autoadministrado. Y, simultáneamente, iniciamos los contactos con los países.

El documento conceptual se estructuró en torno a los siguientes ejes temáticos:

1. Funciones y usos de la evaluación en la educación básica.
 - 1.1. La evaluación con fines de diagnóstico.
 - 1.2. La evaluación para reorientar el proceso educativo.
 - 1.3. La evaluación para calificar y certificar aprendizajes.
 - 1.4. La evaluación como herramienta de comunicación con las familias.
2. Políticas de uso de resultados de las evaluaciones nacionales estandarizadas.
3. Evaluación estandarizada y evaluación en el aula.
4. Currículo prescrito vs. currículo implementado.
5. La evaluación de la lectura en la escuela primaria.
6. La evaluación de la escritura en la escuela primaria.
7. La evaluación de la matemática en la escuela primaria.

En paralelo con la escritura de este documento marco se definieron los instrumentos por emplear, según se consigna en los cuadros 2 y 3. El cuadro 2 incluye el itinerario temático por seguir en la entrevista en profundidad. Obviamente, la pauta fue aplicada en forma flexible, dando cierto grado de libertad al docente para hilvanar su discurso, no necesariamente en el orden planteado. Pero la pauta sirve como orientación y como lista de control para cubrir toda la temática. Las preguntas relativas a la matemática se reiteraban luego para lenguaje.

CUADRO 2
PAUTA PARA LA ENTREVISTA EN PROFUNDIDAD

- ¿Qué tipo de evaluaciones realiza en su clase?
- ¿Con qué propósito las realiza?
- ¿Con qué frecuencia realiza estas evaluaciones?
- ¿Qué hace usted con los resultados que obtiene de cada una de estas evaluaciones?
- ¿Cómo comunica los resultados de las evaluaciones a las familias?
- ¿Comunica a los docentes de las otras aulas, de años anteriores o a la dirección los resultados de sus alumnos?
- ¿Qué aspectos del programa de matemática usted prioriza en sus evaluaciones?
- ¿En qué se basa para establecerlos?
- ¿Además de lo establecido por los programas oficiales, qué otros aspectos fuera del currículo evalúa?
- (Si contestó sí a la anterior) ¿Por qué considera necesario evaluar estos aspectos que no están incluidos en el currículum?
- ¿En qué se basa al elaborar las evaluaciones en matemática que aplica en su clase? ¿En ejercicios propuestos en libros, en propuestas creadas por usted?
- ¿Además de usted, otras personas participan en la elaboración de las actividades de evaluación?
- ¿En qué momentos del curso evalúa a sus alumnos? (al término de cada tema, unidad, bimestre, ciclo escolar). ¿Por qué?
- ¿Cómo evalúa más frecuentemente en matemática, en forma oral o en forma escrita?
- ¿Por qué usa más frecuentemente esta forma (la forma oral o escrita, según lo que conteste)?
- (Sólo si contesta oral, continuar preguntando) Cuénteme cuál/cómo fue la última evaluación oral realizada en matemática.
- Usualmente ¿cuánto tarda en entregarles los resultados de las evaluaciones a los niños?
- ¿Cómo les comunica los resultados a sus alumnos? (notas, comentarios escritos, comentarios orales, frente al grupo, individualmente).
- ¿Por qué lo hace de esa manera?
- ¿Qué actividades les propone a los alumnos a partir de los resultados de una prueba?
- ¿Qué acciones implementa con los niños que obtienen bajos resultados en las pruebas de matemática?
- ¿Cómo registra (o sistematiza) los resultados de cada evaluación?
- ¿Para qué utiliza específicamente los resultados de las evaluaciones en matemática?
- ¿Cada cuánto propone evaluaciones a sus alumnos con el objetivo de calificarlos? ¿Por qué?
- ¿Qué aspectos determinantes tiene en cuenta para decidir si un niño aprueba o no el curso en matemática?
- ¿Existe algún otro aspecto de sus prácticas de evaluación que considere importante que no le haya preguntado y quiera agregar?

El cuestionario autoadministrado estuvo limitado al relevamiento de información acerca del grado de conocimiento de las evaluaciones estandarizadas nacionales e internacionales por parte de los maestros, y su eventual uso. Se trató básicamente de un cuestionario precodificado, que sirvió para comprobar lo que suponíamos: que la gran mayoría de los docentes desconoce las evaluaciones estandarizadas, tanto en cuanto a sus enfoques como en cuanto a sus resultados.

CUADRO 3
PROTOCOLO PARA EL REGISTRO FOTOGRÁFICO DE
PROPUESTAS DE EVALUACIÓN

Se tomarán registros fotográficos de:

1. **Consignas propuestas por los maestros** para evaluar a los alumnos en los últimos tres meses en:
 - a. Lectura.
 - b. Escritura.
 - c. Matemática.
2. **Trabajos realizados por los alumnos** a partir de las consignas de los maestros en:
 - a. Lectura.
 - b. Escritura.
 - c. Matemática.
3. **Registros de los resultados de las evaluaciones** realizadas por los maestros. Para seleccionar los registros fotográficos se tendrá en cuenta que ellos incluyan:
 - a. Evaluaciones en las tres áreas mencionadas.
 - b. Diferentes niveles de desempeño de los alumnos a consideración del maestro: dos de nivel superior (buen desempeño), dos de nivel intermedio (aceptable) y dos de nivel inferior (insuficiente).

Llamar a las escuelas el día previo para recordar la visita y la selección de materiales.

LA ORGANIZACIÓN Y EL ANÁLISIS DE LA INFORMACIÓN

El volumen de información recogido (158 archivos de audio digital, 4.360 registros fotográficos de propuestas de evaluación y trabajos de estudiantes, unos 150 cuestionarios autoadministrados) requería, como primera tarea de envergadura, almacenarlo de manera organizada y de fácil acceso. Para eso se protocolizó el modo de archivar toda la información, en los niveles indicados en el cuadro 4.

CUADRO 4
ESTRUCTURA DE ARCHIVOS

	Entrevistas	Nombre del maestro	
País	Fotos	Lenguaje	Nombre del maestro
		Matemática	Nombre del maestro
		Registros del maestro	Nombre del maestro
	Cuestionarios	Nombre del maestro	

CUADRO 5
FICHA RESUMEN DE CADA RELEVAMIENTO

1. País.
2. Fecha del trabajo de campo.
3. Nombres de las escuelas visitadas.
4. Nombres de los maestros entrevistados.
5. Áreas en las que trabaja el maestro entrevistado: lengua, matemática o ambas.
6. Registros fotográficos. Cantidad de registros tomados, discriminados por lengua, matemática y registros de los maestros.
7. Registros en fotocopias discriminados por lengua, matemática, registros de los maestros.
8. Cuestionario que indique si fue completado o no.
9. Autorización del maestro para publicar los materiales relevados, indicando si firmó o no la autorización.
10. Texto narrativo para recoger impresiones y anécdotas de la experiencia vivida durante el trabajo de campo en el país.

Asimismo, se solicitó a cada responsable del trabajo de campo que elaborara una ficha con los datos principales del relevamiento realizado, incluyendo además un registro narrativo de sus impresiones y experiencias más relevantes durante la visita al país en cuestión, como puede verse en el cuadro 3.

Para el procesamiento de la información tomamos una decisión que luego complicó el trabajo: no transcribir las entrevistas sino trabajar directamente sobre los archivos de audio digital. La decisión fue tomada con el fin de ahorrar tiempo, recursos económicos y papel. En lugar de transcribir todas las entrevistas, se generaron planillas que indicaban los tramos de tiempo del archivo destinados a cada conjunto temático de la pauta de entrevista. A la larga, esta decisión implicó mayor dedicación de tiempo de los investigadores principales, pues se buscaban pasajes en los archivos de audio.

El análisis de la información fue organizado en función de las categorías emergentes del marco conceptual y de los productos definidos para el estudio.

A los especialistas en lenguaje y matemática se les entregó discos compactos con la totalidad de las fotografías relativas a su disciplina, organizadas y etiquetadas según se explicó antes, así como los archivos de audio con indicaciones de los pasajes donde los maestros se referían a la evaluación en la disciplina. A poco de avanzar en su trabajo, que consistía principalmente en identificar "buenas prácticas", nos encontramos con que casi no las había. La mayor parte del material presentaba visiones desactualizadas de la didáctica de la disciplina, casi no hallamos ejemplos de devolución formativa, sino calificaciones, la mayor parte de las actividades resultaron ser típicamente "escolares" y de baja exigencia cognitiva.

Como consecuencia, nos vimos obligados a cambiar sobre la marcha el sentido de las colecciones de propuestas: de ejemplos de "buenas prácticas" pasaron a ser "colecciones de propuestas de evaluación comentadas por especialistas en didáctica". La consigna para los especialistas fue seleccionar un

cierto número de propuestas, las más interesantes de cada país, aun cuando tuviesen limitaciones, y efectuar un análisis que destacase las virtudes de cada propuesta, sus limitaciones y los modos en que podría ser mejorada.

Para los tres productos restantes se seleccionaron los tramos de las entrevistas pertinentes a las temáticas por considerar en cada uno de ellos, así como registros relevantes de trabajos de los estudiantes, teniendo en cuenta la complejidad cognitiva de las tareas propuestas por el maestro y el tipo de devolución o corrección escrita realizada.

Para el tercer producto, referido a la comparación de las "formas de evaluar entre países y su relación con los marcos curriculares de L SERCE, así como de la relación, en el interior de cada país, entre la evaluación en el aula, el currículo prescripto y las pruebas nacionales", se recurrió además al marco curricular del SERCE, que describe y resume los diseños curriculares de todos los países participantes, así como a la documentación existente de las pruebas nacionales de cada país y los ejemplos de ítem utilizados.

Para el cuarto producto, referido al "modo en que las evaluaciones son utilizadas por los maestros para orientar a sus alumnos, informar a las familias y tomar decisiones de calificación y aprobación o reprobación", el análisis de la información recogida se estructuró en torno a los siguientes ejes:

- a. tipos de instrumentos de evaluación utilizados por el maestro;
- b. usos de los resultados y modalidades de devolución a los alumnos;
- c. modalidades de comunicación de resultados a las familias;
- d. criterios y prácticas relativas a la calificación de los alumnos.

En el momento actual no se ha comenzado aún con la elaboración del quinto producto.

Otra cosa que ocurre en la mayor parte de las investigaciones, aunque no suele ser reportado, es que uno enriquece el marco conceptual a medida que desarrolla el trabajo de análisis, porque continúa buscando material y realizando nuevas lecturas. En nuestro caso, por ejemplo, resultó de sumo interés un texto de Grant Wiggins (1998) sobre evaluación auténtica. Por ejemplo, a los efectos de nuestro análisis, tomamos un interesante conjunto de reflexiones sobre las devoluciones en la evaluación. Wiggins identifica distintos modos de llevarla adelante:

- a. bajo la forma de juicios de valor ("muy bien", "no satisface", "incorrecto"), que le indican al estudiante si su trabajo está cerca o lejos de lo que el maestro espera, pero no le da pistas concretas que le permitan analizar su trabajo por sí mismo;
- b. bajo la forma de indicaciones más concretas que le permitan comprender qué es lo que no ha logrado y qué debe mejorar;
- c. bajo la forma de modelos de desempeño con los cuales el estudiante pueda contrastar su propio trabajo y corregirlo por sí mismo.

También incorporamos a nuestro análisis las categorías de Wiggins para establecer el grado de autenticidad de las consignas de evaluación.

LA INCORPORACIÓN DE NUEVOS PAÍSES AL ESTUDIO

Otro gran imprevisto ocurrido durante el desarrollo de la investigación que merece ser destacado fue la incorporación de nuevos países y equipos de investigación que se sumaron al estudio. Como resultado de la participación en seminarios y encuentros con otras instituciones de la región y de la mención al trabajo que estábamos realizando, se despertó el interés de otros grupos de investigación. Fue el caso, en primer lugar, del Grupo de Investigación sobre Evaluación de la Calidad de la Educación de la Universidad de San Buenaventura de Cali, Colombia. El grupo expresó su interés en replicar el estudio en su país haciéndose cargo de los costos del trabajo de campo. El acuerdo que establecimos consistió en que nosotros les aportaríamos todos los instrumentos necesarios para realizar el trabajo y todos los antecedentes que teníamos. Ellos llevarían a cabo el trabajo de campo, organizarían la información y realizarían los análisis primarios del caso nacional de su país. Nos entregarían copia de toda la información recogida. Nosotros incorporaríamos sus análisis a nuestros productos regionales, con los créditos del caso. Ellos realizarían aportes al documento de marco conceptual.

Algo similar ocurrió con el Instituto para la Evaluación de la Educación (INEE) de México y con el Departamento de Educación de la Universidad Católica Argentina, con quienes se establecieron acuerdos de trabajo similares. Esto permitió, por un lado, ampliar la cantidad de casos nacionales en el estudio, que quedó como se indica en el cuadro 6. Por otra parte, nos permitió iniciar una red de intercambio entre equipos de investigación de la región.

CUADRO 6
TOTAL DE PAÍSES INCLUIDOS EN EL ESTUDIO

Resultados	Región		
	Cono Sur	Andina	Centroamérica y Caribe
Altos	Uruguay		México - Costa Rica
Medios	Argentina	Colombia	El Salvador
Bajos		Perú	Guatemala

Esta ampliación de los países por incluir en los productos del estudio implicó una negociación con el PREAL para ampliar el plazo de la entrega de los productos y un leve incremento en el financiamiento (del orden del 10%) con el fin de cubrir el trabajo adicional.

LA PROYECCIÓN AL FUTURO

Cada trabajo de investigación contiene dentro de sí la semilla de nuevos emprendimientos. En nuestro caso, varios proyectos y posibilidades se abren hacia el futuro a partir de este estudio:

- Como fruto de las preocupaciones sobre el tema y enriquecidos por su estudio y por el material recogido, elaboramos un proyecto de Diploma en Evaluación de Aprendizajes, de carácter semipresencial, dirigido a docentes de aula. El proyecto fue aprobado por el Consejo Directivo de nuestra Universidad y ya se ha puesto en marcha. En el momento de escribir este texto se han completado los cupos disponibles (70).
- A partir de las colecciones de propuestas de evaluación analizadas y comentadas por especialistas, proyectamos poner en marcha un sitio web de recursos para la evaluación en el aula, dirigido a docentes.
- Estamos en conversaciones iniciales con el GTEE-PREAL para realizar diversos eventos de difusión de los resultados del trabajo.
- Estamos iniciando el intercambio con los otros grupos para concebir nuevos análisis del *corpus* de información recogido, que puedan dar lugar a nuevos trabajos de investigación o que puedan servir de base para la realización de tesis o memorias de grado.

REFLEXIONES FINALES PARA JÓVENES INVESTIGADORES

Quiero cerrar esta presentación con cinco reflexiones que pueden resultar orientadoras para quienes están en las primeras etapas de su trayectoria profesional en la investigación educativa, dado que entiendo es uno de los propósitos de los organizadores de este libro.

En primer lugar, es importante elegir dentro del campo educativo algunos temas que a uno le interesen especialmente y acumular en torno a ellos lecturas y trabajos. Esto no significa (no se me malinterprete) desconocer o desinteresarse por el resto de las temáticas del campo educativo. Muy por el contrario, tener una visión sobre diversos temas (formación docente, enseñanza para la comprensión, políticas educativas, debates sobre didáctica de la lengua, a modo de ejemplo), me permitió hacer más relevantes los aportes desde mi campo de especialización, la evaluación educativa. Lo importante, en mi opinión, es evitar la situación opuesta: acercarse un poco a muchos temas sin especializarse en ninguno. En resumen, que sin dejar de tener una visión global sobre el campo educativo (lo que se logra principalmente a través de la interacción con colegas con distintas perspectivas y especializaciones), es muy importante elegir uno, dos o tres temas de acumulación y especialización.

En segundo lugar, en la medida en que uno trabaje y lea en torno a ciertos temas, se irá encontrando con trabajos inspiradores, tanto desde el punto de

vista conceptual como desde el punto de vista práctico. Es importante acostumbrarse a registrar esas ideas y ponerlas como en un germinador: darles tiempo, mirarlas de vez en cuando, dejar que den vueltas en la cabeza, escribir unas líneas adicionales. Estarán allí cuando surja una oportunidad.

Tercero, el trabajo bien hecho es la base de la confiabilidad que uno va adquiriendo ante los otros. En mi experiencia, es mejor hacer pocas cosas muy bien que muchas cosas más o menos bien. Lamentablemente, haciendo muchas cosas, aunque no sean buenas ni relevantes, uno puede lograr mejores puntajes en ciertos sistemas de concursos. Pero haciendo muy bien su trabajo, los otros construirán la confianza en ustedes y, a largo plazo, esto les abrirá puertas y posibilidades.

Cuarto, es muy importante construir vínculos con otros profesionales y equipos, así como formar parte de redes de intercambio. Es fundamental elegir con sabiduría. Por ejemplo, en el largo plazo puede ser mejor permanecer formando parte de un equipo de gente competente que hace cosas relevantes, en vez de irse a otro trabajo de menor calidad en el que pagan mejor. Cuando les toque a ustedes conformar un equipo, sepan elegir a la gente adecuada para realizar el trabajo, que comparta un mismo compromiso e inspiración.

Por último, es importante esclarecer y explicitar cuáles son las convicciones que uno tiene sobre el cambio en la educación y sobre su propio compromiso con el mismo, así como un conjunto de ideas centrales acerca de qué es importante para propiciar el cambio desde nuestro ámbito de trabajo. Esas convicciones son las que luego permiten que uno decida sobre los trabajos que uno realiza y el modo de hacerlos. Por oposición, hay quienes simplemente se orientan por el interés de acceder a fondos o de acumular curriculum y antecedentes, sin importarles si su trabajo tiene algún efecto valioso para la sociedad y para los demás. Un corolario de lo anterior, especialmente relevante en las ciencias sociales, es la importancia de preocuparse por la comunicación de la investigación a los distintos agentes sociales, de modo que el conocimiento que se genera sea socialmente útil y contribuya al desarrollo del sistema educativo.

REFERENCIAS BIBLIOGRÁFICAS

- Carnoy, M. (2008): "Valor agregado: por qué usarlo, cómo medirlo y posibles sesgos", mimeo.
- Cueto, S.; Ramírez, C.; León, J. y Pain, O. (2003): *Oportunidades de aprendizaje y rendimiento en matemática en una muestra de estudiantes de sexto grado de primaria de Lima*, Lima, Grupo de Análisis para el Desarrollo (GRADE). Documento de Trabajo n° 43.
- Hiebert, J.; Gallimore, R. y Stigles, J. (2002): "A Knowledge Base for the Teaching Profession: what Would it look like and how can we get one?", en *Educational Researcher*, vol. 31, n° 5, págs. 3-15.

- Martínez, R.; Gaviria, J. L. y Castro, M. (2009): "Concepto y evolución de los modelos de valor añadido en educación", en *Revista de Educación* (Madrid), n° 348, págs. 15-45.
- National Center for Education Statistics (NCES) (2003): *Teaching Mathematics in seven countries. Results from the TIMSS 1999 Video Study*, Washington D.C., NCES 2003-013.
- National Center for Education Statistics (NCES) (2000): *Highlights from the TIMSS Videotape Classroom Study*, Washington D.C., NCES 2000-094.
- Ravela, P. (2008): "Políticas educativas y evaluación", en P. Ravela et al. (coords.), *La evaluación externa en seis países de América Latina: Balances y retos*, Cali, Bonaventuriana, págs. 54-87.
- Ravela, P. (2006a): "¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina?", en P. Arregui (ed.), *Sobre estándares y evaluaciones en América Latina*, Santiago de Chile, PREAL, págs. 181-248.
- Ravela, P. (2006b): *Fichas didácticas. Para comprender las evaluaciones educativas*, Santiago de Chile, PREAL.
- Shadish, W.; Cook, T. y Campbell, D. (2002): *Experimental and quasi-experimental designs for generalized causal inference*, Boston-Nueva York, Houghton Mifflin.
- Shepard, L. (2006): "La evaluación en el aula", México D.F., Instituto Nacional para la Evaluación de la Educación (INEE). Traducción autorizada del capítulo 17 "Classroom assessment", en R. Brennan (ed.), *Educational Measurement*, Westport, ACE/Praeger.
- Tenti Fanfani, E. (2001): *En casa de herrero, cuchillo de palo: la producción y uso de conocimiento en el servicio educativo*, Buenos Aires, Instituto Internacional de Planeamiento de la Educación (IPE/UNESCO).
- Unidad de Medición de Resultados Educativos (UMRE) (1999): *Estudio de los factores institucionales y pedagógicos que inciden en los aprendizajes en escuelas primarias de contextos sociales desfavorecidos en el Uruguay*, Montevideo, Administración Nacional de Educación Pública.
- Wiggins, G. (1998): *Educative Assessment. Designing assessments to inform and improve student performance*, San Francisco, Jossey-Bass Publishers.