

## EN EL PERÚ: BALANCE Y PERSPECTIVAS

*Santiago Cueto\**

### INTRODUCCIÓN

Son pocas las áreas en educación en el Perú en las que se ha acumulado un cuerpo de investigaciones empíricas suficientes como para hacer un balance del conocimiento acumulado. La evaluación del rendimiento escolar en educación básica es sin duda una de ellas. El propósito del presente artículo es presentar y discutir el conocimiento acumulado en cuatro evaluaciones nacionales (EN) y dos internacionales (EI) del rendimiento escolar realizadas en los últimos diez años, conducidas por la Unidad de Medición de la Calidad Educativa (UMC). La UMC es una oficina del Ministerio de Educación del Perú creada en 1996, aunque el trabajo de preparación para la primera EN se inició en 1994. El balance incluye no solamente los informes principales de cada evaluación, sino además los estudios secundarios realizados con las bases de datos generadas.

El presente documento incluye algunos antecedentes de estas evaluaciones, las características y los resultados principales de las seis evaluaciones, los factores asociados al rendimiento en Lenguaje y Matemática y una discusión sobre los retos y las oportunidades del sistema de evaluación del rendimiento estudiantil en el Perú.

El apéndice 1 contiene una lista de todas las publicaciones que se han podido recolectar, editadas por la UMC o de otros autores, que han usado los datos de las evaluaciones nacionales o internacionales mencionadas. Todas estas publicaciones están disponibles en el portal de la UMC.<sup>1</sup>

---

\* Con la valiosa colaboración de Juan León, Laura Chang y Álvaro Zevallos, asistentes de investigación en GRADE. Agradezco los comentarios de Andrés Burga, Giuliana Espinoza, Liliana Miranda, Martín Benavides y José Rodríguez a una versión preliminar del presente documento.

1 Ver <<http://www.minedu.gob.pe/umc/>>.

## 1. ANTECEDENTES

En los últimos veinticinco años prácticamente todos los países de América Latina han creado un sistema nacional de evaluación del rendimiento estudiantil. Ferrer (2005) presenta un balance de los propósitos asumidos en diecinueve países de la región. Una de las formas de clasificar los sistemas de evaluación es observar las implicancias de los resultados: por un lado están los de bajas implicancias, que generan información para fines formativos de diverso tipo, sin consecuencias directas para los actores involucrados; por otro, los sistemas de altas implicancias, que usan los resultados de pruebas para diversos fines, como definir incentivos para los docentes (Chile, Sao Paulo en Brasil y México, por ejemplo), definir la promoción educativa de los estudiantes (Sao Paulo en Brasil, República Dominicana y Costa Rica, por ejemplo) o informar a la población sobre el rendimiento de los estudiantes en los centros educativos (por ejemplo, Chile). Los sistemas de altas implicancias a menudo utilizan marcos de evaluación censales y están ligados a la corriente internacional que aboga por la rendición de cuentas en educación (identificada internacionalmente por el término en inglés *accountability*). El sistema peruano es de bajas implicancias y, como muchos de su tipo, evalúa sobre la base de muestras. Las implicancias de un sistema de evaluación son cruciales para su definición y desarrollo, por lo que serán retomadas al final del presente capítulo.

Ravela, Wolfe, Valverde y Esquivel (2001) presentan otra dicotomía para los sistemas de evaluación del rendimiento escolar, referida a la orientación basada en normas o criterios. En un primer momento, las pruebas referidas a normas fueron dominantes en América Latina. Este modelo proviene de la psicología, ámbito en el que a menudo las variables de interés (inteligencia, por ejemplo) se miden en niveles de intensidad y no por su carencia o dominio absoluto. Las evaluaciones basadas en normas permiten definir cómo se ubica el rendimiento de un estudiante o grupo de estudiantes frente a otros (de la misma edad, grado de estudios u otro rasgo en común). Aplicado a la educación, este modelo llevó al desarrollo de pruebas que procuraban lograr una distribución normal en el rendimiento y presentaban los datos analizando la posición relativa de un grupo de estudiantes frente a otro (estudiantes de centros públicos *versus* privados, por ejemplo). Así, se presentaban resultados indicando en qué percentil se encontraba el rendimiento de un estudiante o grupo de estudiantes y con esto se podía saber qué porcentaje se encontraba por encima y por debajo del grupo descrito. En este tipo de evaluación es difícil establecer de manera absoluta, sin embargo, cuánto saben los estudiantes en relación con un determinado estándar o nivel mínimo aceptable. El modelo de criterios se basa, en cambio, en establecer claramente qué se está midiendo y a partir de qué nivel de rendimiento

se puede fijar un nivel de logro aceptable. En ambos modelos, la base para las evaluaciones es por lo general el currículo escolar, pero el modelo de criterios implica una mayor preocupación por la representatividad de los ítems respecto del objeto de evaluación (verbigracia, validez de contenido, que en pruebas criteriosales se hace sobre la base de detalladas especificaciones del contenido de la prueba). Una forma de entender la diferencia entre ambos modelos es en cuanto a su carácter relativo (normas) versus absoluto (criterios) en la medición del rendimiento; esto implica que las pruebas de normas se usan para interpretar los resultados de un grupo en relación con otro o al promedio nacional, mientras que las pruebas de criterios se usan para interpretar los resultados de un grupo en relación con puntos determinados en la escala (este modelo también permite hacer comparaciones entre grupos). Como se verá más adelante, la información que puede brindar un modelo de normas en un sistema de evaluación del rendimiento escolar (en concreto, el peruano) es sumamente limitada en comparación con el modelo de criterios.<sup>2</sup>

Un error común en la evaluación del rendimiento es considerar que basándose en cualquier prueba se pueden establecer puntos de corte simplemente calculando cuál sería el equivalente del puntaje 11 (o 51%) del máximo posible para la prueba. Esta práctica es errada porque el punto de corte debe ser establecido sobre la base de la dificultad de la prueba y del uso que se le piensa dar a los resultados. Por ejemplo, con todo lo demás equivalente, los contenidos de las pruebas y el puntaje de corte para otorgar el título profesional a los docentes deberían ser más exigentes que para admitir estudiantes a la carrera de educación en una universidad o un instituto superior pedagógico, aunque algunos contenidos de ambas pruebas podrían ser similares. Cizek (2001) presenta varios métodos para establecer puntos de corte en evaluaciones criteriosales estandarizadas, uno de los cuales fue adaptado por la UMC para definir los puntos de corte en la EN 2001 y la EN 2004. Las evaluaciones nacionales de docentes del decenio de 1990 y principios de la presente década han manejado, en cambio, el modelo “simple” de usar el 51% de la prueba administrada (o incluso puntajes menores) como equivalente a un puntaje aprobatorio.<sup>3</sup> Esta práctica es claramente inaceptable de acuerdo con estándares internacionales sobre evaluación del rendimiento escolar (ver AERA, APA y NCTM 1999).

2 Para una mayor discusión sobre los modelos de evaluación, ver Popham (1980).

3 Para el procedimiento utilizado en la evaluación nacional de docentes del 2002, ver Piscoya 2005.

## 2. EVALUACIONES NACIONALES DE RENDIMIENTO EN EL PERÚ

En 1980 la Dirección de Educación Básica Regular del Ministerio de Educación realizó una evaluación del rendimiento de los estudiantes peruanos, con una muestra de estudiantes de sexto grado de primaria en Lenguaje, Matemática, Ciencias Naturales y Ciencias Sociales. La muestra incluía a 1.600 estudiantes de centros públicos y particulares de las tres regiones naturales del Perú, aunque no queda claro si era representativa a escala nacional y de los estratos que se reportan en el informe. Tampoco queda clara la metodología para el desarrollo de las pruebas. Los resultados fueron convertidos a una escala vigesimal, que arrojó puntajes de 07 para Matemática, 12 en Ciencias Histórico-Sociales, 13 en Ciencias Naturales y 14 en Lenguaje. Sin embargo, de acuerdo con lo mencionado antes sobre el establecimiento de puntos de corte en pruebas de rendimiento, sería errado interpretar estos resultados usando el puntaje 11 como indicador de un rendimiento aceptable. Es interesante observar, empero, que el promedio de los estudiantes de Lima Metropolitana fue el más alto, seguido por el de otros de estudiantes de la costa, la sierra y la selva, en ese orden; asimismo, que los estudiantes de centros educativos particulares obtuvieron rendimientos superiores a los de sus pares públicos. Como se verá más adelante, estos patrones se repiten en las evaluaciones de la UMC.

En el Perú se han realizado cuatro evaluaciones estandarizadas nacionales del rendimiento escolar. Todas fueron diseñadas e implementadas por la Unidad de Medición de la Calidad Educativa (UMC), instancia del Ministerio de Educación creada en el marco de un acuerdo y con fondos del Banco Mundial (BM). El BM siguió apoyando las actividades de la UMC en posteriores préstamos. Además del apoyo del BM a las actividades de la UMC, principalmente para la evaluación de la educación primaria, se debe destacar el apoyo del Banco Interamericano de Desarrollo (BID) principalmente para las evaluaciones de la educación secundaria.

El cuadro 1 presenta las principales características de las cuatro evaluaciones nacionales del rendimiento administradas en la última década en el Perú. Como se puede notar, han ido evolucionando en cuanto al modelo y la complejidad de lo evaluado: desde un modelo de normas basado casi exclusivamente en preguntas de opción múltiple hasta uno basado en criterios, que incluye diferentes tipos de preguntas y evalúa a estudiantes con lengua materna indígena en su propia lengua. En todos los casos la evaluación fue muestral; ninguna, censal. La primera prueba en tener verdadera representatividad nacional fue la del 2001, pues las de 1996 y 1998 excluyeron parcial o totalmente a la población rural. Las muestras fueron diseñadas de modo que permitieran reportar resultados con márgenes aceptables de error muestral a escala nacional y por algunos desagregados. Todas las evaluaciones nacionales incluyeron encuestas

**Cuadro 1**  
**Evaluaciones nacionales de rendimiento en el Perú**

Año	Representatividad de la muestra*	Áreas evaluadas	Grados evaluados	Modelo de evaluación	Tipo de ítems e instrumentos**
1996	A nivel nacional, pero la muestra es representativa solo de escuelas polidocentes completas y para departamentos, costa, sierra y selva, área geográfica (urbano y rural) y por tipo de gestión (estatal y no estatal).	Lenguaje y Matemática.	4.º de primaria	Normas	Opción múltiple, comprensión de lenguaje oral y expresión escrita.
1998	A nivel nacional, pero la muestra es representativa solo para zonas urbanas y para departamentos, costa, sierra y selva, tipo de gestión (estatal y no estatal) y Lima Metropolitana.	Lenguaje, Matemática, Personal Social (en primaria) y Ciencias Naturales (en primaria).	4.º y 6.º de primaria y 4.º y 5.º de secundaria	Normas	Opción múltiple y expresión escrita.
2001	La muestra es representativa a nivel nacional y para escuelas polidocentes y multigrado, gestión estatal y no estatal y urbano-rural en el caso de primaria. En secundaria es representativa, además, por departamentos.	Comunicación y Matemática. Comunicación incluye pruebas para estudiantes de escuelas bilingües interculturales con quechua del sur y aimara. Escalas de actitudes hacia la matemática, lectura, escritura y respeto por las lenguas nativas.	4.º y 6.º de primaria y 4.º de secundaria	Criterios	Opción múltiple, solución de problemas con materiales concretos en matemática, expresión escrita, expresión oral. Escalas Likert para medir actitudes.
2004	La muestra es representativa a nivel nacional y para gestión estatal y no estatal, urbano y rural, regiones y escuelas polidocentes y multigrado	Comunicación, Matemática y Formación Ciudadana (solo en 6.º de primaria y 5.º de secundaria). Comunicación incluye pruebas para estudiantes de escuelas bilingües interculturales con quechua del sur y aimara.	2.º y 6.º de primaria, 3.º y 5.º de secundaria	Criterios	Opción múltiple, preguntas de respuesta abierta corta y preguntas de respuesta abierta larga (incluye producción de textos y reflexiones sobre formación ciudadana).

\* Todas las EN permiten comparar rendimiento de mujeres y hombres.

\*\* Todas las EN han sido acompañadas de encuestas a directores, docentes, padres de familia y estudiantes.

sobre factores asociados al rendimiento respondidas por estudiantes, padres de familia, docentes y directores (en algunos casos comprendieron, además, observaciones de infraestructura y servicios disponibles en el centro educativo). Finalmente, todas aplicaron pruebas piloto de los instrumentos, aunque la orientación fue diversa: mientras que las pruebas de normas usaron principalmente los resultados de las pruebas piloto para decidir las bondades de cada ítem,<sup>4</sup> en las pruebas del 2001 y el 2004 se utilizaron criterios de discriminación de acuerdo con el modelo de Rasch, junto a otros vinculados a la correspondencia entre los ítems y los objetos de medición; así, ítems con buen ajuste al modelo podrían ser descartados si en la evaluación piloto se encontraba que algunos sectores de la población no comprendían bien la pregunta o incluso, antes del piloto, si expertos en el currículo determinaban que un ítem no era válido para medir una determinada competencia.

## 2.1. Evaluación Nacional 1996

La evaluación de 1996 es especialmente interesante, pues al ser la primera experiencia de este tipo hubo problemas de concepción e interpretación de datos ligados sobre todo a lo siguiente: de acuerdo con las corrientes internacionales prevalecientes en ese momento (sobre todo en Chile, con mucha influencia en los orígenes de la UMC), el modelo adoptado fue de normas; pero una vez obtenidos los resultados, se los quiso interpretar como de criterios. En otras palabras, al seguir el modelo de normas (que fue correctamente implementado), los resultados en Lenguaje y Matemática finalmente arrojaron resultados alrededor de 50% del puntaje máximo posible (45,4% en Matemática y 49,7% en Lenguaje).<sup>5</sup> Esto fue interpretado por las autoridades educativas del momento como equivalente al puntaje 10 (desaprobado) y por tanto inaceptables. Se pensó que el gobierno, que asumió el Ejecutivo en 1990, quedaría mal parado con estos resultados. Se trató sin duda de una interpretación peculiar, pues no existían antecedentes similares con los cuales comparar el rendimiento. Esto generó que el informe, ya impreso y listo para distribuir, fuera finalmente secuestrado y no divulgado.

Las únicas otras publicaciones resultantes de esta evaluación fueron cuatro breves manuales de Lenguaje y dos de Matemática con recomendaciones técnico-pedagógicas orientadas a docentes. Los manuales incluyen sugerencias para mejores prácticas de los docentes en el aula, basadas —según se explicita en

4 En la prueba final se prefirieron ítems que, sin ser demasiado fáciles ni difíciles, permitían una mayor discriminación entre estudiantes de rendimiento global alto y bajo.

5 Ver Ministerio de Educación (1998).

cada publicación— en los resultados de las pruebas; sin embargo, no se incluyen ejemplos de ítems ni resultados globales. Con base en estos manuales se programaron decenas de capacitaciones de docentes a escala nacional, a cargo de especialistas de la UMC, en las que se repartieron los materiales aludidos. Tales actividades generaron un notable recelo de otra oficina del Ministerio de Educación, cuya función principal era justamente la capacitación de docentes en ejercicio. Pensamos que todo esto deja una primera lección: establecer con claridad que los reportes de la UMC deberían alimentar las decisiones de otras instancias; como oficina, no debería incursionar en actividades que corresponden a otras dependencias ministeriales.

## 2.2. Evaluación Nacional 1998

La EN 1998 hacia el final, y la del 2001 desde su concepción hasta la administración de la prueba, tuvieron un ingrediente adicional: la colaboración de un órgano asesor externo. Este órgano asesor externo fue contratado con base en una licitación, que fue ganada por GRADE (Grupo de Análisis para el Desarrollo). La iniciativa de contar con esta asesoría externa nació del Banco Mundial, que —como se dijo— prestó el financiamiento para las actividades iniciales de la UMC. La colaboración entre GRADE y la UMC resultó en la invitación a una serie de especialistas internacionales con experiencia en evaluaciones estandarizadas. Ellos trabajaron con especialistas de la UMC,<sup>6</sup> así como con consultores nacionales contratados por la UMC y el órgano asesor coordinadamente.

Como se mencionó antes, la evaluación nacional de 1998 también fue normativa. La publicación de las pruebas se hizo en sucesivos números del *Boletín UMC*, los primeros de los cuales reportaron los resultados de las encuestas. El *Boletín UMC* 5/6 incluye los resultados obtenidos en las pruebas de Lenguaje y Matemática; el 7, en Ciencias Sociales y Naturales. Los números 10 al 20 ofrecen ejemplos de los ítems de las diferentes pruebas y grados y sus resultados, a partir de los cuales se discuten las implicancias pedagógicas. En este periodo se publicaron además dos números de la revista *Crecer*, que incluyen una presentación de los objetivos y fundamentos de la EN 2001, un análisis de resultados y artículos de opinión de especialistas de la UMC, GRADE y otras instituciones.

---

6 Los jefes de la UMC durante este periodo fueron María Esther Cuadros, quien dirigió la UMC desde su fundación hasta noviembre 1999, y José Rodríguez, quien la dirigió entre diciembre de 1999 y diciembre del 2002.

### 2.3. Evaluación Nacional 2001

La evaluación nacional 2001 fue la primera basada en un modelo de criterios. El plan general de la evaluación fue presentado en un documento público que orientó el desarrollo de los demás instrumentos específicos (Rodríguez y Cueto 2001). Este documento especificaba las preguntas de la evaluación y el uso que se daría a los resultados. En la EN 2001 se tuvo un gran cuidado para fundamentar y definir el objeto de evaluación, que fue publicado en las especificaciones de contenidos de las pruebas. A diferencia de las evaluaciones anteriores, estas especificaciones incluían una justificación de la orientación de la evaluación para cada área y de los aspectos del currículo vigente incluidos en la evaluación, una definición de cada uno de estos aspectos y ejemplos de los ítems a ser incluidos en las pruebas. Estas especificaciones eran bastante más detalladas que las contenidas en las EN 1996 y 1998, que se habían basado principalmente en la taxonomía cognoscitiva de Bloom y en los contenidos curriculares vigentes en cada momento, presentados sin especificaciones. Las detalladas especificaciones de la EN 2001 requirieron un mayor número de ítems que las EN anteriores, para poder evaluar los contenidos. Dado que no se podía someter a todos los estudiantes a resolver todos los ítems, a partir de la EN 2001 se usó la metodología de “formas rotadas”. Este diseño permitió dividir el número total de ítems en varios cuadernillos, asignándose aleatoriamente a cada estudiante uno o dos de estos cuadernillos. Luego, a partir del cuadernillo resuelto por los estudiantes era posible estimar la probabilidad de acertar los ítems en los cuadernillos que los estudiantes no habían respondido, y calcular un puntaje total en toda la prueba para cada estudiante. Para esto se usaron procedimientos estadísticos vinculados al modelo de Rasch, que en general se enmarca en los modelos de teoría de respuesta al ítem. Este avance de formas rotadas es fundamental para el sistema, pues permitió aumentar considerablemente el número de ítems de la evaluación, afinando de este modo el análisis que se podía hacer de los aprendizajes de los estudiantes. Por lo demás, el sistema de formas rotadas es común en los países más desarrollados, aunque no conocemos instituciones —además de la UMC— que lo usen en el Perú.<sup>7</sup> La EN 2001 incluyó asimismo justificaciones de las encuestas y escalas de actitudes utilizadas. Todos estos documentos se encuentran en el portal de la UMC; además, algunos han sido impresos y distribuidos por el Ministerio de Educación.

---

7 Con la excepción de GRADE.



## 2.4. Evaluación Nacional 2004<sup>8</sup>

Dentro del mismo modelo de criterios, en la evaluación nacional 2004 se definió el objetivo de la evaluación en general y luego, con gran detalle, las especificaciones de prueba y formas rotadas. Lo mismo se hizo con las encuestas. Los resultados globales no han sido publicados, aunque sí los correspondientes a las áreas de Matemática y Comunicación, de primaria y secundaria, en documentos de análisis pedagógico. Los análisis presentados son sumamente interesantes, pues presentan y analizan los resultados de los ítems de Matemática y Comprensión de Lectura. Si bien la prueba incluyó una evaluación de Formación Ciudadana, a la fecha de escribir el presente informe (julio del 2006) los resultados no se encontraban disponibles.

Una novedad de la EN 2004 fue la presentación de los resultados de la evaluación a docentes, que si bien se aplicó de manera voluntaria en las mismas aulas donde se administraron las pruebas a estudiantes de sexto de primaria y quinto de secundaria, alrededor de 94% accedieron a responderla. En el área de Comunicación se presentaron a los docentes textos equivalentes a un grado de dificultad para estudiantes de quinto de secundaria; en Matemática, ítems equivalentes a un grado de dificultad de segundo de secundaria. Los ítems administrados a los docentes fueron similares, pero no exactamente iguales a los de los estudiantes. El informe mostrado a la prensa dice que “Los docentes de los alumnos evaluados muestran dificultades en el dominio de algunas habilidades lectoras y matemáticas. La mayoría de ellos tiene capacidad de resolver solo las tareas más sencillas”.<sup>9</sup> También muestran una clara asociación entre el resultado de los profesores y el de sus estudiantes. Este último resultado es, sin embargo, correlacional; por lo tanto, no se pueden hacer atribuciones causales sobre su base: puede ser que los docentes con rendimiento bajo sean asignados con estudiantes de rendimiento bajo y los docentes altos con estudiantes similares. Por otro lado, tampoco se puede descartar una relación causal: los docentes con rendimiento bajo enseñan menos o peor que los docentes de rendimiento alto. Independientemente de lo anterior, los resultados muestran que muchos profesores tienen un pobre dominio de los contenidos curriculares que deben enseñar. Esta evidencia, que seguramente será explorada y presentada con mayor detalle por la UMC, es en principio más fuerte para reflejar las carencias docentes que los datos de las evaluaciones realizadas para nombramiento de docentes en la década pasada y la presente. Decimos esto porque los contenidos de las

8 Durante este periodo la UMC estuvo dirigida por Giuliana Espinoza (enero del 2003 a marzo del 2005) y Liliana Miranda (abril del 2005 a la fecha).

9 Ver “Evaluación Nacional 2004”, documento en Power Point UMC (2005), disponible en <<http://www.minedu.gob.pe/umc/presentaciones.php>>.

pruebas de docentes mencionadas han sido sumamente variados y no contaban con el sustento de especificaciones de contenido ni datos de pruebas piloto u otra forma de sustentar su validez predictiva (para la evaluación nacional de docentes del 2002, ver Piscocoya 2005). Los ítems administrados por la UMC a los docentes en la EN 2004 tuvieron un solo y claro objetivo: aproximarse al nivel de dominio de los docentes de los contenidos curriculares vigentes para estudiantes en Comunicación y Matemática.

### 3. RESULTADOS DE LAS PRUEBAS NACIONALES

Como se ha explicado antes, las evaluaciones de 1996 y 1998 fueron normativas; por tanto, no tiene mayor sentido hablar de porcentajes de ítems correctamente respondidos para la prueba en conjunto, dado que la dificultad fue prevista para obtener una tasa de acierto de aproximadamente 50%. Sí vale la pena, en cambio, comparar la posición relativa de las regiones del Perú en cada evaluación. Los resultados de este ejercicio se deben interpretar con cautela, pues, como se especificó en el cuadro 1, la representatividad de las muestras es limitada en las dos primeras evaluaciones. De todas formas, el sesgo de cada muestra es similar entre regiones (aunque el peso de lo rural podría hacer variar los resultados). Los datos de la EN 2001 y la EN 2004 son los únicos representativos de toda la población nacional estudiantil. En la última columna del cuadro 2 se presenta la posición relativa de la región en cuanto a niveles de riqueza.<sup>10</sup> En el apéndice 4 se presentan los mismos análisis, por separado, para Comunicación y Matemática. No se presentan datos para todos los grados evaluados, pues solo se han incluido datos representativos a escala regional (verbigracia, con márgenes de error aceptables).

Los resultados de las regiones en las cuatro evaluaciones nacionales son bastante consistentes, con algunas excepciones. En general, las regiones de la costa (incluidas Arequipa, Lima, Tacna y Moquegua) muestran rendimientos más altos, mientras que las regiones de la sierra y la selva (incluidas Apurímac, Huanavelica, Huánuco, Loreto y Ucayali) muestran los rendimientos más bajos.

La correlación entre el promedio de ranking de las evaluaciones nacionales a nivel regional para primaria y secundaria con el ranking de riqueza es de 0,62 y 0,68 respectivamente.<sup>11</sup> En ambos casos, la correlación es estadísticamente sig-

10 Para ordenar las regiones por niveles de riqueza se utilizó el documento *La pobreza en el Perú 2003*, elaborado por Javier Herrera y el equipo del Instituto Nacional de Estadística e Informática (INEI). El rango en el cual se sitúan las regiones va de menos pobre (1) a más pobre (24).

11 Para la correlación de las variables se usó el coeficiente de Spearman.

Cuadro 2  
Posición relativa de las regiones del Perú en las cuatro evaluaciones nacionales

	Primaria						Secundaria						Ránking de riqueza <sup>d</sup>			
	2.º		4.º		6.º		Ránking Promedio EN		3.º		4.º			5.º		Ránking Promedio EN
	EN 2004 <sup>a</sup>	EN 1996 <sup>b</sup>	EN 1998 <sup>c</sup>	EN 1998	EN 2004	EN 2004	EN 2004	EN 1998	EN 2001	EN 1998	EN 2004	EN 2004		EN 2004		
Huancavelica	17	16	14	15	18	16	22	23	24	19	22	22	24			
Huánuco	21	17	15	21	17	19	18	21	14	15	18	17	23			
Puno	11	14	19	20	19	16	24	22	16	20	24	22	22			
Cajamarca	16	11	11	11	14	13	19	13	9	12	15	15	21			
Amazonas	12	4	10	14	15	11	20	7	12	10	16	14	20			
Ayacucho	20	23	22	17	21	21	13	20	6	17	14	19	19			
Loreto	24	21	23	18	23	22	21	16	20	16	20	19	18			
Apurímac	23	24	24	24	24	24	23	18	22	24	21	22	17			
Ucayali	22	20	20	23	20	21	17	24	23	22	23	21	16			
Piura	14	10	12	9	12	12	9	14	10	9	10	10	15			
San Martín	18	18	21	22	16	19	15	17	13	21	17	16	14			
Pasco	13	6	17	13	13	13	10	8	11	14	9	10	13			
Junín	10	8	6	4	8	8	6	4	3	2	5	4	12			
Ancash	15	12	8	6	11	11	16	11	17	7	15	14	11			
Cusco	19	15	16	12	22	17	14	15	19	13	13	15	10			
La Libertad	8	1	4	7	6	6	5	6	5	6	6	6	9			
Lambayeque	6	9	7	8	7	7	8	9	8	11	8	9	8			
Arequipa	4	2	1	1	3	3	2	1	1	3	2	2	7			
Lima	3	3	2	2	2	3	4	2	4	1	4	3	6			
Moquegua	2	5	3	3	4	3	3	5	7	5	3	4	5			
Ica	5	13	9	10	5	8	12	10	15	8	12	12	4			
Madre de Dios	9	22	18	19	9	14	7	19	18	18	7	13	3			
Tacna	1	7	5	5	1	3	1	3	2	4	1	2	2			
Tumbes	7	19	13	16	10	12	11	12	21	23	11	15	1			

Notas

- a. Datos tomados de la UMC. De acuerdo con sus autores, los datos para regiones de la EN 2004 deben ser interpretados con cautela ya que el error muestral es excesivamente alto para algunas regiones.
- b. Para el cálculo del ordenamiento de cada departamento en la evaluación GRECEP 1996 se utilizaron las bases de datos disponibles en el portal UMC ([www.minedu.gob.pe/umc](http://www.minedu.gob.pe/umc)). En el caso de Lambayeque, los resultados deben ser interpretados con cautela ya que hubo problemas para cubrir toda la muestra seleccionada.
- c. Datos tomados del informe de la UMC. En la EN 1998 la muestra fue representativa a nivel regional pero solo de escuelas urbanas.
- d. El ránking de riqueza va de la región más pobre, Huancavelica, a la menos pobre, Tumbes.

nificativa al 5%. En otras palabras, a menor pobreza de la región mejor resultado en las EN, siendo esta asociación más fuerte en la educación secundaria. Sin embargo, hay regiones que siendo relativamente pobres tienen un desempeño inesperadamente alto en las pruebas (como Amazonas o Cajamarca), y otras que no teniendo altos indicadores de pobreza tienen un rendimiento relativamente bajo (como Tumbes). ¿Qué factores explican estas discrepancias? Esta pregunta es especialmente relevante en el contexto actual de descentralización educativa. Posibles explicaciones incluyen analizar las características del cuerpo docente y de los institutos formadores de docentes en estas zonas.

Otro aspecto importante es saber si el gasto público en educación por parte del gobierno se ejecuta de manera que discrimine positivamente a las regiones que han mostrado un bajo rendimiento educativo. Para promover una mayor equidad, el gasto público en educación<sup>12</sup> por habitante debería ser mayor en aquellas regiones con menor rendimiento en las EN. Los gráficos 1 y 2 muestran la dispersión entre el gasto en educación por alumno en el sector estatal para el 2004<sup>13</sup> y el ranking promedio en las evaluaciones nacionales para primaria y secundaria<sup>14</sup> (se debe recordar que el ranking promedio de las EN se ha establecido de modo que el que figura con el puesto 1 tiene el mejor rendimiento).

En los gráficos 1 y 2 se observa una asociación negativa entre el gasto por alumno y el promedio del ranking a nivel regional en primaria y secundaria. La correlación<sup>15</sup> entre el gasto público por alumno en educación primaria y el ranking de la región en primaria es de  $-0,37$  (significativa al 10%), mientras que la correlación entre el gasto público por alumno en educación secundaria y el ranking de la región en secundaria es de  $-0,40$  (significativa al 5%). En otras palabras, las regiones donde los alumnos obtienen mejores resultados educativos son aquellas en las cuales el Estado tiende a gastar más por alumno.

---

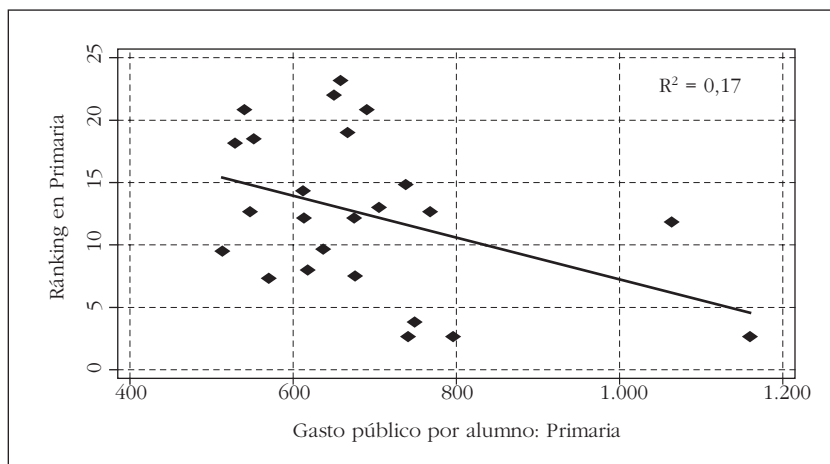
12 Para el cálculo del gasto público en educación por alumno se utilizó la base de datos del Sistema Integrado de Administración Financiera del Ministerio de Economía y Finanzas, disponible en <http://ofi.mef.gob.pe/transparencia/default.aspx>, así como la información de estadísticas básicas del 2004 de la Oficina de Estadística del Ministerio de Educación. En el apéndice 5 se muestran las partidas consideradas para el cálculo del gasto público en educación.

13 Para el cálculo del gasto por alumno en educación por parte del Estado se consideró dentro de la función Educación y Cultura las partidas relacionadas con la educación primaria y secundaria. No se consideraron los rubros destinados al pago de personal administrativo. Los promedios para los niveles primaria y secundaria del gasto público en educación por alumno en el 2004 fueron US\$ 199,7 y US\$ 274,4 (el tipo de cambio usado es el promedio de diciembre del 2004: 3,28 soles por dólar). El estudio de Jaramillo y Arteaga (2003) cita datos del 2001 respecto del gasto por alumno en los niveles primaria y secundaria, situándolo para el Perú en US\$ 122, mientras que en el documento de Saavedra y Suárez (2002) los montos ascienden a US\$ 135 para primaria y US\$ 191 para secundaria.

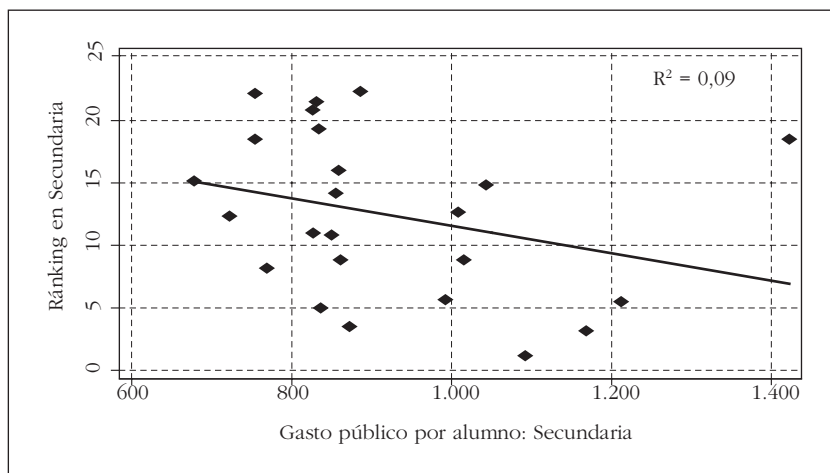
14 Para este análisis se considera solo a las escuelas públicas.

15 El coeficiente de correlación calculado es el de Spearman en ambos casos.

**Gráfico 1**  
**Relación entre el promedio regional de evaluaciones nacionales a nivel primaria (escuelas públicas solamente) y el gasto público en educación primaria por alumno (nuevos soles)\***



**Gráfico 2**  
**Relación entre el promedio regional de evaluaciones nacionales a nivel secundaria (escuelas públicas solamente) y el gasto público en educación por alumno (nuevos soles)\***



\* En ambas regresiones solo se identificó como punto influyente a la región Moquegua en el caso de primaria y la región Tumbes en el caso de secundaria. Sin embargo, la omisión de dichas observaciones en la regresión no modifica el signo de la relación entre el ránking (primaria o secundaria) y el gasto público por alumno, aunque reduce el nivel de ajuste a la recta en el gráfico de primaria ( $R^2=0,07$ ) y en el caso de secundaria se incrementa el nivel de ajuste ( $R^2=0,29$ ).

La evaluación del 2001 y la del 2004 fueron criteriosales y por tanto permiten estimar el porcentaje de estudiantes con un rendimiento satisfactorio. Los cuadros 3 y 4 presentan los resultados en Lenguaje y Matemática respectivamente. En las evaluaciones anteriores, el único nivel equivalente a “aprobado” es el denominado “suficiente”. Como dice el informe de la EN 2004, “Los estudiantes de este nivel demuestran un dominio suficiente y necesario en los desempeños evaluados para el grado”.

**Cuadro 3**  
**Niveles de desempeño en Comunicación en las Evaluaciones Nacionales**  
**2001 y 2004**

Grados	EN 2001			EN 2004			
	Suficiente	Básico	Debajo de básico	Suficiente	Básico	Previo	Debajo del previo
2.º de primaria	...	...	...	15,1	23,7	15,1	46,1
4.º de primaria	34,8	13,7	51,4	...	...	...	...
6.º de primaria	9,1	17,7	73,2	12,1	28,1	35,7	24,1
3.º de secundaria	...	...	...	15,1	23,8	26,9	34,3
4.º de secundaria	21,3	20,5	58,2	...	...	...	...
5.º de secundaria	...	...	...	9,8	45,3	14,8	30,1

Como se puede ver en los cuadros 3 y 4, el porcentaje con nivel suficiente es menor en Matemática que en Comunicación; y, por lo menos en Matemática, hay menos con nivel suficiente en secundaria que en primaria. Es importante considerar si los rendimientos han evolucionado positivamente. Si bien los resultados de la EN 2001 y la EN 2004 parecen fáciles de comparar, la falta de un número considerable de ítems comunes entre ambas evaluaciones lo dificulta; en el futuro, la UMC debería preocuparse por poner suficientes ítems comunes entre evaluaciones para poder ser rigurosos y comparar resultados. En cambio en la presentación de resultados de la EN 2004 se menciona que en la compa-

**Cuadro 4**  
**Niveles de desempeño en Matemática en las Evaluaciones Nacionales**  
**2001 y 2004**

Grados	EN 2001			EN 2004			
	Suficiente	Básico	Debajo de básico	Suficiente	Básico	Previo	Debajo del previo
2.º de primaria	...	...	...	9,6	27,4	18,8	44,2
4.º de primaria	26,6	44,3	29,1	...	...	...	...
6.º de primaria	7,1	40,6	52,3	7,9	34,7	12,7	44,7
3.º de secundaria	...	...	...	6,0	19,9	19,0	55,1
4.º de secundaria	5,2	8,6	86,2	...	...	...	...
5.º de secundaria	...	...	...	2,9	11,0	17,7	68,5

ración de puntajes en la prueba de Matemática en sexto grado de primaria entre 1998 y el 2004 no hay diferencias estadísticamente significativas. En quinto de secundaria se encuentra un rendimiento similar en Comunicación y ligeramente inferior (significativo a nivel estadístico) en Matemática el 2004.<sup>16</sup> La práctica de comparar puntajes históricamente es importante, aunque resulta compleja en términos técnicos pues obliga a usar ítems similares (o, mejor aún, cuadernillos iguales) y muestras comparables en los dos momentos que se analizan. La UMC puso un gran cuidado al hacerlo, y es el único caso en el Perú —que conozcamos— en el cual es posible establecer este tipo de comparaciones con la certeza de que se cumplieron tales requisitos.

16 Se debe recordar que la EN 1998 incluyó solamente estudiantes en contextos urbanos; y por tanto, la comparación es con sus pares en la EN 2004. No existen comparaciones de estudiantes rurales en el tiempo.

#### 4. EVALUACIONES INTERNACIONALES DEL RENDIMIENTO ESCOLAR

El Perú ha participado en dos evaluaciones internacionales del rendimiento escolar: la del Laboratorio Latinoamericano de Evaluación de la Calidad Escolar (LLECE)<sup>17</sup> organizada por la Oficina Regional de Educación para América Latina y el Caribe de la UNESCO, y PISA (*Programme for International Student Assessment*), organizada por la Organización para la Cooperación y el Desarrollo Económico (OCDE).<sup>18</sup> La primera prueba se administró en 1997 y la segunda en el 2001.

La participación en LLECE involucró evaluar el rendimiento de estudiantes en Lenguaje y Matemática en tercer y cuarto grado de primaria a escala nacional. Perú participó junto a otros doce países, pero inicialmente se rehusó a autorizar la publicación de los resultados. Si bien se dieron explicaciones técnicas respecto a tal decisión, todo indica que el secuestro de información obedeció a razones políticas (se debe recordar que en la misma época no se publicaban los datos de la EN 1996). El modelo de la evaluación fue de normas, como lo indica la pobre especificación de los contenidos de las pruebas. Los resultados se orientaron a comparar países y estratos dentro de los países (megaciudad, urbano y rural, por un lado, y público y estatal, por otro). Los resultados fueron publicados en el Perú por el gobierno de transición a fines del año 2000.<sup>19</sup> El cuadro 5 presenta los promedios del Perú y los de los otros países.

Como se puede ver en el cuadro 5, lo más notable es la superioridad del rendimiento de los estudiantes cubanos. El promedio fue fijado en 250 y la desviación estándar en 50, lo cual indica que los estudiantes cubanos están por encima de una desviación estándar en todas las pruebas sobre el segundo y varias por encima de los últimos. Luego de Cuba viene un segundo grupo de países que incluye a Argentina, Brasil y Chile, y luego el grupo de la cola, en el que se encuentra el Perú. Un análisis de Torreblanca (2001) muestra que el rendimiento de los estudiantes peruanos fue estadísticamente similar al de sus pares de Paraguay, República Dominicana y Honduras en todas las pruebas, y en algunas de ellas similar al de México, Colombia, Bolivia y Venezuela. Sobre los resultados de esta y otras pruebas es importante destacar que a menudo se han cometido gruesas exageraciones. Así, Piscoya (2005: 102) reporta que “De acuerdo con las pruebas de la UNESCO (1998-2000) nos disputamos con Haití el último lugar del ránking de la región de América Latina y el Caribe”. Esta afirmación, repetida por otros especialistas, es falsa, pues Haití no participó en la evaluación citada, realizada solamente en 1997.

17 Ver <<http://llece.unesco.cl/>> para objetivos y publicaciones.

18 Ver <<http://www.pisa.oecd.org/>> para los informes internacionales.

19 Ver al respecto la información que aparece en el *Boletín UMC 9* (UMC y GRADE 2001).



**Cuadro 5**  
**Resultados en la evaluación LLECE (1997) por grados y países (medianas)**

Países	Lenguaje		Matemática	
	3.º grado	4.º grado	3.º grado	4.º grado
Cuba	343	349	351	353
Argentina	263	282	251	269
Chile	259	286	242	265
Brasil	256	277	247	269
Venezuela	242	249	220	226
Colombia	238	265	240	258
Bolivia	232	233	240	245
Paraguay	229	251	232	248
México	224	252	236	256
Perú	222	240	215	229
R. Dominicana	220	232	225	234
Honduras	216	238	218	231

Fuente: Tomado de *Boletín UMC* 9.

Si se analiza el rendimiento por subgrupos, se nota que es principalmente el sector público —y en este, los estudiantes de zonas rurales— el que explica la baja posición global del Perú. Sin embargo, el rendimiento de los estudiantes peruanos del sector privado también fue relativamente bajo: en tercer grado, el sector no estatal del Perú se ubica en el puesto 6 en Lenguaje (empate con Paraguay) y 8 de Matemática (de once países, considerando que Cuba no tiene sector privado). En cuarto grado, los estudiantes del sector privado del Perú ocupan el puesto 7 en Lenguaje y 6 en Matemática (empate con Bolivia).<sup>20</sup> Es importante destacar que LLECE se encuentra organizando el segundo estudio de rendimiento escolar en América Latina, en el cual el Perú está participando. La evaluación debía de ocurrir en el segundo semestre del 2006 y los resultados se publicarían a inicios del 2008.

La segunda evaluación internacional en la que participó el Perú es PISA (OECD y UNESCO 2003). El propósito de la evaluación es determinar el nivel de alfabetización en tres áreas: Comprensión de Lectura, Matemática y Ciencia. La evaluación se realiza con estudiantes de 15 años matriculados en el sistema educativo. En el caso del Perú solo se incluyó a estudiantes que estaban en secundaria (se excluyó a estudiantes de 15 años en educación primaria). A diferencia de LLECE y las EN, que se centraron en el currículo para determinar los

20 Ibidem, ob cit.

objetos de evaluación, PISA se centró en evaluar “habilidades para la vida”. En otras palabras, se buscaba determinar en qué medida los estudiantes próximos a entrar en la adultez tenían habilidades para manejar situaciones cotidianas problemáticas.

El Perú quedó último entre los países evaluados en las tres áreas. Esto podría ser esperable, sin embargo, en la medida en que los países de la OCDE en general muestran altos niveles de desarrollo, muy por encima del Perú. Es notable, empero, que en la comparación con otros países de América Latina —Argentina, Brasil, Chile, México y, por equiparaciones posteriores, Uruguay— los estudiantes del Perú resultaron bastante atrasados. El cuadro 6 presenta los resultados de los estudiantes en las cinco categorías de rendimiento; la quinta es la categoría más alta y la tercera indica un rendimiento aceptable.

**Cuadro 6**  
**Porcentaje de estudiantes latinoamericanos en**  
**prueba de Comprensión de Lectura (PISA)**

	Argentina	Brasil	Chile	México	Perú	Uruguay
Nivel 5	1,7	0,6	0,5	0,9	0,1	5,0
Nivel 4	8,6	3,1	4,8	6,0	1,0	11,0
Nivel 3	20,3	12,9	16,6	18,8	4,9	20,0
Nivel 2	25,5	27,7	30,0	30,3	14,5	24,0
Nivel 1	21,3	32,5	28,3	28,1	25,5	20,0
Nivel 0	22,6	23,3	19,9	16,1	54,1	20,0

Nota: Los resultados corresponden a la evaluación PISA 2000-2001, excepto los de Uruguay, que corresponden a PISA 2003. Elaborado sobre la base de OECD y UNESCO (2003) y ANEP (2004).

Como se puede notar en el cuadro 6, el Perú tiene un bajo rendimiento que se explica tanto por la concentración de estudiantes peruanos en la categoría de rendimiento más baja como por la casi total ausencia de estudiantes en la categoría más alta. El nivel 0 identifica a estudiantes que no pueden localizar la idea principal en un texto, identificar información explícita en un texto sencillo y realizar conexiones entre información del texto y hechos de la vida cotidiana; lamentablemente, 54% de los estudiantes peruanos cayeron en este nivel, muy por encima de lo encontrado en los otros países latinoamericanos. Como sería de esperar, los resultados son peores entre estudiantes con menores niveles de escolaridad (es decir estudiantes de 15 años en primero o segundo de secundaria, en comparación con cuarto o quinto de secundaria).

Posteriores análisis describen el rendimiento de los estudiantes del sector privado. Así, en Comprensión de Lectura los estudiantes peruanos del sector

privado tuvieron el más bajo desempeño en comparación con sus pares en Argentina, Brasil, Chile y México; sin embargo, la diferencia es estadísticamente significativa (al 1%) solo con los estudiantes de México, que estuvieron por encima de una desviación estándar sobre los peruanos. De hecho, estadísticamente, el rendimiento del sector estatal mexicano es similar al privado peruano.<sup>21</sup> En Matemática, el sector privado peruano nuevamente tiene el rendimiento más bajo, esta vez en comparación con Argentina, Brasil y Chile, siendo estadísticamente inferior al primero. En Ciencia se repite el patrón del sector privado peruano con el más bajo rendimiento, con una diferencia estadísticamente significativa a favor de los estudiantes chilenos.<sup>22</sup> Estos resultados confirman de alguna manera lo reportado antes para el sector privado en LLECE y sugieren que cuando un sistema público es de mala calidad (como el peruano), el privado no puede ser de excelencia.

PISA es una evaluación planificada para realizarse cada tres años, con énfasis, cada vez, en un área de evaluación (en la primera, Comprensión de Lectura; en la segunda, Matemática; y la planificada para el 2006, Ciencia). Consideramos una oportunidad perdida que el Perú se haya retirado de la segunda y la tercera evaluación. Las razones para el retiro del Perú tuvieron que ver con una supuesta falta de relevancia de resultados que nuevamente mostrarían al Perú a la cola del rendimiento. Frente a este argumento, pensamos que es importante destacar que el énfasis de la segunda evaluación PISA fue Matemática, y hubiera sido útil conocer y comparar a los estudiantes peruanos con estándares internacionales en esta área para sugerir reformas orientadas al currículo, la preparación de textos escolares, la formación de docentes y la capacitación de docentes en ejercicio. Es cierto, sin embargo, que para que esto ocurra es preciso tener un sistema educativo con especialistas que aprenden de los resultados de las evaluaciones; este no parece haber sido el caso en la primera evaluación PISA y tal vez en ninguna EN o LLECE. En general, los resultados de las evaluaciones solo parecen haber resultado en titulares catastróficos de los diarios y una declaración de emergencia que —en cuanto se puede conocer de los medios públicos— no se ha basado en un análisis cuidadoso de qué pueden y no pueden hacer los estudiantes peruanos. Sin embargo, instituciones civiles, como el Consejo Nacional de Educación, han recogido los resultados de PISA y las EN para fundamentar la necesidad de aplicar profundas reformas en el sistema educativo.

Es importante notar que recientemente se realizó la Primera Reunión Iberoamericana de PISA, que congregó a representantes de Argentina, Brasil, Colombia, Chile, España, México, Portugal y Uruguay. Es sin duda una iniciativa interesante,

21 Ver al respecto el Documento de Trabajo UMC 6 (Caro 2004).

22 Ver al respecto el Documento de Trabajo UMC 10 (Asmad y otros 2004).

orientada a sacar el mayor provecho posible de los resultados de las pruebas. En todo caso, tal vez con tiempo se pueda incluir al Perú en este grupo de avanzada en cuanto a evaluación de rendimiento en América Latina y participar en PISA 2009, que nuevamente pondrá énfasis en Comprensión de Lectura. Sería esperable que luego de ocho años de administrada la primera prueba PISA se consiguieran mejores resultados, no necesariamente comparándonos con el resto de países sino principalmente con nuestros propios resultados al iniciar esta década.

## 5. ANÁLISIS DE FACTORES ASOCIADOS

Un estudio puede arrojar correlaciones engañosas entre variables aun cuando esté bien diseñado. Esto se debe al carácter probabilístico de las estadísticas utilizadas, que pueden resultar en falsos positivos o falsos negativos,<sup>23</sup> lo que se puede superar en gran medida con la técnica de meta-análisis, consistente en incluir en un análisis estadístico los resultados de varios estudios sobre una misma variable. El supuesto es que si bien un estudio puede resultar engañoso, la acumulación de estudios bien diseñados debería brindar evidencia sobre la “verdadera” relación entre variables. En este caso se realizará el meta-análisis por separado para Lenguaje y Matemática.

Se han incluido todos los estudios que usaron las bases de datos de la UMC para ver los factores asociados al rendimiento de los estudiantes en Matemática y Lenguaje en las EN 1996, 1998 y 2001. Una vez recolectados los diferentes estudios, se procedió a revisar las variables consideradas en cada uno. Se encontraron trece estudios en el área de Matemática y nueve en el área de Comunicación.

Para cada variable independiente (explicativa) se contabiliza el número de estudios en los cuales el análisis muestra una asociación estadísticamente significativa, o ausencia de la misma, con el rendimiento. En caso de haber relación significativa, se indica el signo de la asociación. Cabe señalar que para el presente estudio se hubiera querido realizar un meta-análisis con mayor detalle en cuanto al tamaño del efecto de cada variable sobre el rendimiento, pero no todos los estudios cuentan con información relacionada con los estadísticos descriptivos (media, desviación y error estándar) necesarios para poder hacer este tipo de análisis. Los cuadros 7 y 8 presentan un resumen de algunas variables incluidas en el mayor número de estudios. En la lista de referencias bibliográficas y en el apéndice 1 están señalados con un asterisco (\*) los estudios utilizados para el meta-análisis.

---

23 Es decir, concluir, en un estudio, que existe una relación significativa cuando en realidad no la hay; o, por el contrario, que no existe una relación significativa cuando en realidad la hay.

Cuadro 7  
EN 1996, 1998 y 2001. Resumen de variables asociadas al rendimiento en Lenguaje y Matemática a nivel del estudiante

	Variables del estudiante					
	Matemática (n = 13)			Comunicación (n = 9)		
	N.º de estudios en los que está presente	Asociaciones significativas	Asociaciones no significativas	N.º de estudios en los que está presente	Asociaciones significativas	Asociaciones no significativas
<i>Características demográficas</i>						
Género (niños)	13	13	0	9	1	0
Extraedad / Repetición de grado	4	0	4	2	0	2
El estudiante trabaja	4	0	4	4	0	4
<i>Idioma</i>						
Habla el castellano	4	3	0	3	1	0
Lengua materna (castellano)	4	4	0	5	2	0
<i>Background académico</i>						
Cuenta con educación inicial	3	0	0	2	0	0
<i>Actitudes y hábitos del estudiante</i>						
Gusto por la matemática	7	7	0	...	...	...
Asistencia a la escuela	4	3	0	...	...	...
<i>Expectativas de los padres</i>						
Nivel educativo del hijo	5	3	0	3	2	0
<i>Características de la familia</i>						
Vive con ambos padres	5	2	0	3	1	0
Nivel socioeconómico	10	6	0	6	4	0
<i>Capital cultural de la familia del estudiante</i>						
Número de libros en casa	7	4	3	6	4	2
Nivel educativo de la madre	8	4	0	8	5	3
Nivel educativo del padre	7	3	0	7	4	0

Cuadro 8  
EN 1996, 1998 y 2001. Resumen de variables asociadas al rendimiento en Lenguaje y Matemática a nivel del aula o la escuela

	Variables del aula de clase o la escuela					
	Matemática (n = 13)			Comunicación (n = 9)		
	N.º de estudios donde está presente	Asociaciones significativas	Asociaciones no significativas	N.º de estudios donde está presente	Asociaciones significativas	Asociaciones no significativas
<i>Características geográficas y tipo de escuela</i>						
Rural	7	1	0	6	5	2
Sierra	6	1	2	3	4	0
Estatal	9	0	5	4	6	0
Multigrado	5	1	1	3	4	0
<i>Infraestructura educativa y nivel socioeconómico de la escuela</i>						
Índice de recursos educativos	6	1	0	5	5	4
Nivel socioeconómico (promedio de estudiantes)	7	7	0	0	4	4
Número de libros por cada 10 alumnos	2	2	0	0	2	0
Computadoras operativas por cada 1000 alumnos	2	2	0	0	2	0
<i>Características demográficas de los docentes</i>						
Lengua materna (nativa)	2	0	1	1	2	0
Años de experiencia	6	2	1	3	4	2
Docente cuenta con título (universitario, instituto superior pedagógico o instituto superior tecnológico)	4	1	0	3	3	0
Docente es nombrado	3	0	0	3	1	0
<i>Acercá de la labor docente</i>						
Número de horas pedagógicas a la semana que enseña Matemática / Comunicación	4	2	0	2	2	0
Tamaño de la clase	5	1	1	3	4	1
Satisfacción con aspectos de la escuela	3	0	0	3	2	0
<i>Procesos educativos</i>						
% de competencias desarrolladas	4	4	0	0	1	0
La escuela elaboró el Proyecto Curricular de Centro	3	0	0	3	...	...
Manejo del docente de los contenidos curriculares	3	2	0	1	2	0
Fecha de inicio de clases	2	2	0	0	2	1
<i>Rol de los padres en la escuela</i>						
Revisar las tareas	3	2	0	1	1	0
Satisfacción de los padres con la escuela	2	2	0	0	2	0

Es importante entender las limitaciones del anterior análisis. Como se ha mencionado en varios casos, los diseños de todas las evaluaciones han sido transversales y ninguno longitudinal. Así, los resultados deben ser interpretados principalmente como asociaciones entre variables y no como estrictas relaciones de causa y efecto. Entendemos que la UMC se encuentra planificando un estudio longitudinal, que constituye un mejor diseño para establecer relaciones causales; esto, porque permite aislar el incremento de puntajes de un punto de tiempo a otro en función del puntaje inicial y las características de las escuelas, aulas y procesos educativos involucrados. Por lo demás, el diseño de investigación preferido para establecer este tipo de relaciones es el experimental, con asignación aleatoria a las condiciones de tratamiento (experimental) y control,<sup>24</sup> pero este tipo de diseños es prácticamente inexistente en el sector educación en el Perú (a diferencia de campos como salud o investigación agrícola, en los que son más comunes).

Se debe recordar que los análisis anteriores corresponden a modelos que incluyen y controlan por el efecto de varias variables explicativas simultáneamente. Así, los efectos señalados como significativos son relativamente robustos dependiendo del modelo estadístico empleado y del conjunto de variables de control incluidas en cada modelo.

Los resultados se comentan a continuación, por grupos de variables. El primer grupo se refiere a las características relacionadas directamente con los estudiantes:

- *Características de los estudiantes.* Destaca la variable referida al género, incluida en todos los estudios analizados. Mientras que es significativa a favor de los hombres en Matemática en todos los estudios, solo lo es a favor de los hombres en uno de nueve estudios en Comunicación. El resultado favorable a los hombres en Matemática seguramente responde a estereotipos y expectativas sociales respecto del rol de hombres y mujeres, y es consistente con resultados hallados en estudios internacionales (por ejemplo PISA, ver OECD y UNESCO 2003). En segundo lugar, en todos los estudios se encontró que los estudiantes con extraedad y/o que hubieran repetido tuvieron menor rendimiento que sus pares en la edad para el grado. Finalmente, en todos los estudios en los que se incluyó la variable “El estudiante trabaja” se encontró un relativo menor rendimiento respecto de los que no trabajan. En cuanto a la lengua materna del estudiante, en casi todos los estudios en Matemática se aprecia un efecto favorable en el rendimiento del dominio del castellano *versus* una lengua indígena, mientras que en Comunicación

---

24 Ver por ejemplo Mosteller y Boruch (2002).

solo dos de los cinco estudios que la incluyeron muestran una asociación significativa. Este resultado, que merece un estudio más amplio, podría estar asociado a la dificultad de entender códigos matemáticos en castellano cuando la lengua es indígena.

- *Educación inicial.* La educación inicial no muestra un patrón de asociaciones significativas con rendimiento. Este resultado, que parece ilógico y de hecho refuta estudios internacionales, se puede deber a que no importaría tanto el haber asistido a educación inicial sino durante cuánto tiempo y en qué condiciones funcionaba el centro. También se debe considerar que el dato de educación inicial es provisto por lo general por los padres, que pueden denominar así a una amplia variedad de centros, incluidas las guarderías, que no necesariamente buscan impactar en el rendimiento escolar posterior. Otra posibilidad es que el efecto de haber asistido a educación inicial sea notable solo en los primeros grados de primaria, equiparándose luego en rendimiento los que han asistido y los que no. Finalmente, se debe recordar que las variables de control socioeconómico podrían estar capturando (o enmascarando) el efecto de la educación inicial, pues las familias con mayor poder adquisitivo envían con mayor probabilidad a sus hijos a educación inicial que las familias más pobres.
- *Actitudes y hábitos del estudiante.* En este grupo de variables se puede apreciar que la denominada “Gusto por las matemáticas” fue utilizada en siete estudios; en todos ellos, la relación resultó significativa y positiva. Esto podría indicar una relación endógena, pues seguramente los que tienen un mayor gusto por la matemática tienen un mejor rendimiento en Matemática. La asistencia a la escuela (durante todo el año o las últimas dos semanas) estuvo asociada a un mejor rendimiento en Matemática en tres de los cuatro estudios que la incluyeron.
- *Expectativas de los padres.* Las expectativas de los padres en cuanto a la educación de sus hijos muestra un patrón de asociación positiva y significativa con el rendimiento en la mayoría de estudios, pero este resultado también puede ser endógeno: los padres de hijos con mejor rendimiento tienen mayores expectativas educativas.
- *Características de las familias.* En este caso se usaron dos variables: la primera es la presencia de la madre y el padre en la casa, que muestra una asociación con el rendimiento en tres de los ocho estudios (en los restantes cinco no fue significativa); y la segunda, el nivel socioeconómico (NSE) del estudiante. Esta variable muestra una fuerte asociación con el rendimien-



to, pero se debe notar que en los diferentes estudios se usan definiciones similares pero no iguales de NSE.

- *Capital cultural.* En este aspecto las variables presentan una mayor ambigüedad en cuanto al sentido de su relación con el rendimiento de los niños. Así, el número de libros en casa tiene asociación positiva y negativa con rendimiento tanto en Comunicación como en Matemática, y la relación es significativa en todos los estudios en los que se utilizó esta variable. En cuanto al nivel educativo de los padres, parecería que, sin ser un gran predictor, la educación del padre es menos ambigua que la educación de la madre.

La anterior lista de variables sugiere algunas posibilidades para compensar desigualdades entre grupos, como focalizar los esfuerzos en elevar el rendimiento de las niñas en Matemática, y mejorar las oportunidades de los estudiantes que trabajan, tienen una lengua materna indígena y/o tienen padres con escaso nivel educativo (o que, en general, viven en condiciones de pobreza).

El segundo grupo de variables corresponde al ambiente escolar o del aula:

- *Características geográficas y tipo de escuela.* Las escuelas rurales, de por sí, no parecen tener menores resultados que sus pares urbanas (esto es, luego de controlar estadísticamente por otras variables vinculadas a las características de los estudiantes). Esto no significa que las rurales no tengan carencias; múltiples estudios demuestran que sí las tienen. Lo que el resultado sugiere es que la diferencia principal entre escuelas urbanas y rurales reside en la relativa mayor pobreza de los estudiantes que atienden escuelas rurales. Por otro lado, a pesar de los controles estadísticos, en la mitad de los estudios el rendimiento en escuelas estatales es menor que en escuelas privadas, lo que sugiere que hay una diferencia en cuanto a la calidad del servicio educativo a favor de estas últimas.
- *Infraestructura educativa y nivel socioeconómico de la escuela.* Los recursos educativos de la escuela (materiales educativos) parecen ser más importantes en Comunicación (seguramente en relación con disponibilidad de libros) que en Matemática. Esto se confirma en la variable “Número de libros por cada 10 estudiantes”, que también muestra una asociación positiva con el rendimiento. El nivel socioeconómico promedio de los estudiantes en un aula se asocia positivamente con el rendimiento en todos los estudios, lo que sugiere que la composición del aula es importante más allá de las características individuales del estudiante. Finalmente, el número de computadoras operativas en la escuela por estudiante también muestra una asociación

positiva con el rendimiento, aunque este resultado puede ser endógeno (las escuelas con estudiantes de más recursos tienen más computadoras y estudiantes con mejor rendimiento).

- *Características demográficas de los docentes.* Ninguna de las características docentes incluidas en el análisis parecen tener mayor asociación con el rendimiento de los estudiantes. Esto puede estar explicado por la selección de variables o porque el resultado de los estudiantes tiene que ver con las características de varios profesores a lo largo de su carrera educativa y no solo del profesor de aula al momento de la evaluación.
- *Labor docente.* Tampoco parece haber mayor relación entre las tres variables incluidas en el presente meta-análisis y el rendimiento en Matemática o Comunicación. De nuevo, estas prácticas se refieren a características solo durante un año y por tanto pueden no reflejar la historia educativa de los estudiantes.
- *Procesos educativos.* El porcentaje de competencias curriculares desarrolladas parece un buen predictor del rendimiento en Matemática, mas no en Comunicación (aunque solo ha sido incluida en un estudio en esta área). Contar con un proyecto curricular de centro no parece tener mayor relación con el rendimiento en Matemática. El manejo docente de los contenidos curriculares ha sido incluido en pocos estudios, pero sí parece tener asociación con el rendimiento. Igual sucede con la fecha de inicio de clases: mientras más pronto sea resulta mejor, posiblemente porque aumenta el tiempo de aprendizaje durante el año escolar.
- *Rol de los padres.* La ayuda que los padres puedan brindar a sus hijos al revisar las tareas ha sido incluida en cinco estudios; sin embargo, solo dos de ellos muestran una asociación positiva con el rendimiento, en Matemática. La satisfacción de los padres con la escuela tiene una asociación positiva con el rendimiento en Matemática, mas no en Comunicación.

Tal vez la variable de mayor importancia en cuanto al ambiente escolar para explicar el rendimiento estudiantil sea una que apenas se ha esbozado en la EN 2004: el dominio que tengan los docentes de las competencias que deben enseñar. Como se vio antes, la EN 2004 mostró un bajo rendimiento de muchos docentes (aun cuando los informes disponibles al momento son limitados). Más allá de cualquier análisis estadístico, es lógico pensar que el dominio de los contenidos por parte del docente es un prerrequisito necesario para el aprendizaje estudiantil: si un docente no domina una competencia, no la puede enseñar

(podría ser, sin embargo, que un docente que la domine enseñe mal, pero si no la domina no hay forma de que pueda enseñar lo que ignora). Pensamos que una línea de investigación en esta dirección sería importante en cuanto a la selección de futuros docentes y capacitación de los actuales. En el Congreso de la República existe un proyecto de Carrera Pública Magisterial pendiente de discusión. Por otro lado, el Congreso aprobó en mayo del 2006 la legislación referida al Sistema Nacional de Evaluación, Acreditación y Certificación Educativa (SINEACE); sin embargo, a julio del mismo año no ha sido implementado. Los dos anteriores tienen importantes elementos de evaluación de los docente que deberían incluir su dominio de las materias que van a enseñar.

## 6. EVALUACIÓN ESTANDARIZADA DEL RENDIMIENTO ESCOLAR EN EL PERÚ: LOS SIGUIENTES PASOS

Como se ha sugerido a lo largo del presente capítulo, la UMC ha logrado importantes avances en solo diez años de actividad. En la presente sección se discuten algunas perspectivas de posibles desarrollos en cuanto al futuro de las actividades de evaluación del rendimiento escolar y el estudio de factores asociados vinculados a la mejora de la calidad y equidad educativa.

a) *Uso de la información generada.* Si bien pensamos que la información generada por la UMC es potencialmente útil y relevante, es indudable que el conocimiento que muchos actores principales tienen de ella es bastante limitado. Es fundamental que antes de realizar una evaluación se tenga claro *qué* se quiere saber y *cómo* se utilizarán los resultados para mejorar el sistema educativo. Así, los datos pueden seguir luego este plan (Rodríguez y Cueto 2001). Si no hay claridad en estos puntos, no tiene sentido evaluar.<sup>25</sup>

Experiencias internacionales sugieren la importancia de trabajar con los actores en la comprensión y el análisis de la información producida. La experiencia de Brasil puede resultar relevante en este sentido. En ese país existe el Comité de Productores de Información Educativa (COMPED), creado y presidido por el Instituto Nacional de Estudios y Pesquisas (INEP). El COMPED reúne a quince instituciones que buscan desarrollar un “sistema articulado de diseminación de informaciones educacionales”. Por ejemplo,

---

25 Es notable, en este contexto, que los resultados de las pruebas no se usen como uno de los insumos para definir el gasto público, que de hecho es mayor en las regiones con mejor rendimiento (es decir, lo contrario de lo que se esperaría de un Estado que promueve la equidad educativa).

el COMPED apoya la publicación de títulos que contribuyan a la formación inicial y continuada de profesores.

- b) *Continuar diseñando evaluaciones criteriosales.* Es indudable que el cambio de un modelo de evaluación normativo a uno criterial entre las dos primeras y dos últimas EN ha sido conveniente. La primera evaluación de LLECE fue también normativa, pero la evaluación PISA fue criterial. Este modelo criterial, que permite establecer puntos de rendimiento aceptables y descripciones detalladas de lo que los estudiantes han aprendido o no pueden hacer, es el que hay que mantener y reforzar para que sea realmente informativo de lo que pueden hacer los estudiantes y lo que necesitan aprender (en este sentido, participar en PISA fue sin duda una experiencia singularmente formativa para el personal especializado de la UMC).
- c) *Uso de diversas modalidades y diseños para evaluar el rendimiento escolar y factores asociados.* En esta línea se han iniciado interesantes experiencias, como el uso de formatos diversos de ítems (opción múltiple, respuestas abiertas, expresión escrita, expresión oral, solución de problemas con materiales concretos y escalas de actitudes), con diversas poblaciones (evaluación de Comunicación en lengua materna y en castellano como segunda lengua en poblaciones quechua y aimara) y con diversos diseños (además de las evaluaciones transversales usadas, en la UMC se tiene previsto usar diseños cuantitativos longitudinales, es decir, el seguimiento de los mismos sujetos en dos momentos diferentes; y cualitativos vinculados a eficacia escolar). Este tipo de trabajos e iniciativas ha recibido poca difusión, pero es relevante para entender la complejidad del rendimiento escolar y los factores asociados, y debería seguir siendo apoyado.
- d) *Importancia de comparar resultados en el tiempo.* La EN 2004 es la primera que incluyó una nota sobre la posibilidad de comparar rendimientos (1998 y 2004).<sup>26</sup> Este es un requerimiento fundamental para entender la evolución del rendimiento escolar. Sugerimos que, con fines de comparación, se considere a partir de ahora que siempre se deben evaluar los rendimientos de los estudiantes en el grado final de cada nivel (esto es, sexto de primaria y quinto de secundaria), en Comunicación y Matemática. Por lo demás, cada evaluación podría variar en cuanto a los otros grados y áreas curriculares incluidos en la evaluación. Por ejemplo, las evaluaciones que se han hecho

---

26 Las EN 1996, 1998 y 2001 también incluyeron procedimientos para comparar resultados, pero estas comparaciones nunca han sido publicadas.

hasta ahora han sido todas transversales. Las evaluaciones longitudinales permitirían una estimación de cuánto han aprendido los estudiantes en el tiempo transcurrido. Este diferencial, o valor agregado, podría ser luego relacionado con las características de los centros educativos relevantes. De este modo se podrían establecer relaciones de causalidad más fuertes que con evaluaciones transversales.

- e) *Evaluación de altas o bajas implicancias.* Esta es una de las grandes cuestiones planteadas desde un inicio para la labor de la UMC. Existen modelos como el chileno, que utilizan las evaluaciones para informar a los padres de familia sobre los promedios de los centros educativos; como el mexicano, que las usan con fines de promoción docente; o como el costarricense, que las usan para definir la graduación de sus estudiantes. Cualquiera de estas alternativas es sumamente difícil de implementar de manera “limpia” técnicamente. De hecho, actualmente hay grandes críticas a los sistemas de altas implicancias por cuestiones de validez de las inferencias que se pueden realizar. Las objeciones giran alrededor de temas como hasta qué punto son justas las comparaciones entre centros educativos que atienden a poblaciones diferentes, en contextos diferentes, con recursos humanos y materiales diferentes. Los ajustes estadísticos que habría que hacer para que las comparaciones fueran válidas parecerían simplemente imposibles de establecer de manera empírica. En segundo lugar, ¿cómo se puede juzgar la calidad de un docente, y por tanto definir su promoción en la carrera, a partir del rendimiento de sus estudiantes, dadas las preguntas formuladas en el punto anterior?<sup>27</sup> Finalmente, si se establecen puntajes mínimos en pruebas estandarizadas nacionales para otorgar diplomas educativos (de secundaria, por ejemplo), es claro que los estudiantes de familias más pobres tendrán menos probabilidades de alcanzarlos. Esto seguramente resultaría, para este grupo, en menos oportunidades de acceder a trabajos y estudios superiores. Las anteriores discusiones serían de tan compleja y costosa solución (verbigracia, seguramente requeriría mediciones repetidas de todos los estudiantes de al menos algunos grados, además de datos de su contexto familiar y comunal) que pensamos que, por lo menos por el momento, el sistema peruano se debería mantener en un sistema de bajas implicancias, en el cual es suficiente tener evaluaciones muestrales realizadas cada tres o cuatro años. La información debería estar orientada a obtener datos sobre la evolución de los rendimientos escolares de acuerdo con las expectativas

---

27 Esto es diferente de evaluar a los docentes para conocer si dominan los contenidos que van a enseñar, que sí debería ser parte de la evaluación docente.

establecidas en el currículo escolar. Esta información debería ser publicada, de modo que fuera relevante para los actores principales del sistema educativo peruano, como funcionarios del Ministerio y oficinas descentralizadas; directores, profesores y padres de familia en centros educativos; y profesores y estudiantes en centros de formación y capacitación docente.

- f) *Participación en evaluaciones internacionales.* Es importante participar en evaluaciones internacionales siempre y cuando estas provean de resultados relevantes para identificar el rendimiento de los estudiantes en contenidos específicos (así, la simple ubicación de estudiantes peruanos en un ránking internacional de países es interesante, pero insuficiente para justificar la participación del Perú). En segundo lugar, la participación del Perú en estas evaluaciones permite mejorar la capacidad técnica del equipo de la UMC. En esta línea, la participación del Perú en la segunda evaluación de LLECE es importante, pero podría ser complementada con la participación en un estudio de Educación Cívica y Ciudadanía el 2008 (organizado por la IEA), PISA (evaluación de Comprensión de Lectura, Ciencia, Matemática y Resolución de Problemas) el año 2009 o las de TIMSS (evaluación en Ciencia y Matemática) o PIRLS (evaluación de Comprensión de Lectura), que tienen procesos previstos para el 2011. No se está sugiriendo, por cierto, participar en todas, sino considerar opciones complementarias que permitan dar una visión comprehensiva del rendimiento de los estudiantes peruanos en diferentes grados y áreas, que complemente la información de las EN y permitan planificar acciones para mejorar el rendimiento.
- g) *Posible rol del Sistema Nacional de Evaluación, Acreditación y Certificación de la Calidad Educativa (SINEACE).* En la Ley General de Educación aprobada por el Congreso de la República el 2003 se creó el Sistema Nacional de Evaluación, Acreditación y Certificación de la Calidad Educativa. Este sistema debería cumplir funciones de evaluación del rendimiento escolar, certificación de competencias profesionales y acreditación de instituciones educativas. En mayo del 2006 el Congreso aprobó una ley para el establecimiento del SINEACE, que debía ser implementado en los siguientes meses. Una alternativa interesante a considerar al respecto es que las evaluaciones que realiza la UMC actualmente pasen al nuevo instituto responsable del SINEACE, de modo que sus actividades tengan un carácter público pero autónomo. Esta experiencia parece haber sido positiva en México y Brasil.
- h) *Evaluación y descentralización.* En el contexto actual de descentralización política y educativa existen ya algunas experiencias regionales de evaluación educativa (en Piura y Arequipa, por ejemplo). Sería importante que la

UMC y estas oficinas compartieran los conocimientos adquiridos, así como con otras oficinas regionales; no con vistas a imponer objetivos sino, por el contrario, para fomentar el desarrollo de experiencias diversas, pero técnicamente sólidas, que permitan conformar un movimiento de evaluación del rendimiento diverso de acuerdo con las necesidades regionales y nacionales de conocimiento.

Las actividades de la UMC se han dado en un contexto internacional de importancia creciente de la evaluación estandarizada del rendimiento escolar y, en general, de desarrollo de políticas educativas basadas en evidencias. Ejemplo de esto último son diversas iniciativas internacionales, tales como:

- *What Works Clearinghouse* (WWC,<sup>28</sup> “Políticas que funcionan”), en Estados Unidos. Es una iniciativa que ha acumulado información sobre programas educativos evaluados rigurosamente y que han demostrado tener impactos positivos en el rendimiento. En el portal de WWC se encuentran resúmenes del estado de conocimiento en áreas específicas como matemática, lectura, educación en valores, deserción escolar, educación en la primera infancia, inglés como segunda lengua, conducta delictiva, educación de adultos y aprendizaje asistido por pares.
- *Evidence for Policy and Practice Information* (EPPI,<sup>29</sup> “Evidencias para información de políticas y prácticas”), en Gran Bretaña. De manera similar a WWC, acumula información de estudios rigurosos sobre intervenciones eficaces en educación. EPPI incluye revisiones en salud y educación. En este último campo cuenta con revisiones acerca del uso de la evaluación sumativa, educación ciudadana, desarrollo profesional en servicio, desarrollo infantil, enseñanza del inglés, estereotipos de género, educación inclusiva, educación matemática, motivación, educación en ciencias e influencia de los directores en el aprendizaje de sus estudiantes.

No queremos decir que las evaluaciones y los informes publicados por la UMC sean equivalentes a WWC o EPPI. De hecho, en estos últimos las evaluaciones de programas demuestran su eficacia de preferencia con diseños experimentales (es decir, de asignación aleatoria a las condiciones experimental y control). Las evaluaciones transversales de rendimiento y el estudio de factores asociados están lejos de alcanzar este estándar. Sin embargo, el punto es que

28 Ver <<http://www.whatworks.ed.gov/>>.

29 Ver <<http://eppi.ioe.ac.uk/EPPIWeb/home.aspx>>.

actualmente hay una fuerte corriente internacional que promueve que las decisiones de política educativa se tomen sobre la base de información empírica y no solamente de posturas filosóficas o corrientes teóricas. Esta tendencia no parece haber decrecido recientemente sino todo lo contrario. En tal sentido, pensamos que la continua inversión en el fortalecimiento del trabajo especializado que realiza la UMC, sea desde el Ministerio o el Sineace, y la difusión y discusión de sus publicaciones, tienen un gran potencial como insumo para el desarrollo de políticas educativas orientadas hacia la calidad y equidad.



## REFERENCIAS BIBLIOGRÁFICAS

- AERA, APA y NCTM-AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, AMERICAN PSYCHOLOGICAL ASSOCIATION y NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- ANEP-ADMINISTRACIÓN NACIONAL DE EDUCACIÓN PÚBLICA (2004). *Primer Informe Nacional PISA 2003 Uruguay*. Montevideo: ANEP.
- ASMAD, Úrsula; David PALOMINO, Mary TAM y Gloria ZAMBRANO (2004). *Una aproximación a la alfabetización matemática y científica de los estudiantes peruanos de 15 años. Resultados del Perú en la evaluación internacional PISA*. Documento de Trabajo 10. Lima: UMC-Ministerio de Educación.
- CARO, Daniel (coordinador) (2004). *Una aproximación a la alfabetización lectora de los estudiantes peruanos de 15 años. Resultado del Perú en la evaluación internacional PISA*. Documento de Trabajo 6. Lima: UMC-Ministerio de Educación.
- CIZEK, Gregory J. (2001) (editor). *Setting Performance Standards. Concepts, Methods and Perspectives*. Nueva Jersey: Lawrence Erlbaum Associates, Publishers.
- FERRER, Guillermo (2005). *Estado de situación de los sistemas nacionales de evaluación de logros de aprendizaje en América Latina*. PREAL, disponible en <<http://www.preal.cl/GTEE/docr/Estado%20de%20Situacion.pdf>>.
- FERRER, Guillermo y Patricia ARREGUI (2003). *Las pruebas internacionales de aprendizaje en América Latina y su impacto en la calidad de la educación: criterios para guiar futuras aplicaciones*. Documento de Trabajo 26, PREAL: Santiago. Disponible en <[http://www.preal.org/docs-trabajo/Arregui\\_et\\_Al\\_26.pdf](http://www.preal.org/docs-trabajo/Arregui_et_Al_26.pdf)>.
- JARAMILLO, Miguel e Irma ARTEAGA (2003). *La inversión pública en educación: proceso de asignación y determinantes de la distribución del gasto por alumno*. Informe de investigación entregado al Consorcio de Investigación Económica y Social. Disponible en <[www.consortio.org](http://www.consortio.org)>.
- MINISTERIO DE EDUCACIÓN DEL PERÚ (1998). *Evaluación del Rendimiento Estudiantil "CRECER"*. Informe de Resultados de la Primera Prueba Nacional. Lima: Ministerio de Educación.
- DIRECCIÓN DE EDUCACIÓN BÁSICA REGULAR (1981). *Diagnóstico del rendimiento académico de los educandos del sexto grado de E. B. R.* Lima: Ministerio de Educación (mimeo).

- MOSTELLER, Frederick y Robert BORUCH (editores) (2002). *Evidence matters. Randomized Trials in education Research*. Washington, DC: Brooking Institution Press.
- OECD y UNESCO (2003). *Literacy Skills for the World of Tomorrow. Further Results from PISA 2000*. Disponible en <<http://www.pisa.oecd.org/dataoecd/43/9/33690591.pdf>>.
- PISCOYA, Luis (2005). *Cuánto saben nuestros maestros. Una entrada a los diez problemas cardinales de la educación peruana*. Lima: Corporación Financiera de Desarrollo y Fondo Editorial de la Universidad Nacional Mayor de San Marcos.
- POPHAM, James (1980). *Problemas y técnicas de la evaluación educativa*. Madrid: Anaya.
- RAVELA, Pedro (2002). *¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina?* Documento de Trabajo 22, PREAL: Santiago. Disponible en <<http://www.preal.org/GTEE/pdf/Resultados2.pdf>>.
- (2003). *¿Cómo aparecen los resultados de las evaluaciones educativas en la prensa?* PREAL. Disponible en <<http://www.preal.org/GTEE/pdf/prensa.pdf>>.
- RAVELA, Pedro; Richard WOLFE, Gilbert VALVERDE y Juan Manuel ESQUIVEL (2001). *Los próximos pasos: ¿cómo avanzar en la evaluación de aprendizajes de América Latina?* Documento de Trabajo 20, PREAL, Santiago. Disponible en <<http://www.preal.org/doctra20.pdf>>.
- RODRÍGUEZ, José y Santiago CUETO (julio del 2001). *¿Cuánto aprenden nuestros estudiantes? Presentación de la Evaluación Nacional del 2001*. En *Creceer* 2, 20-24. Lima: Ministerio de Educación y GRADE.
- SAAVEDRA, Jaime y Pablo SUÁREZ (2002). *El financiamiento de la educación pública en el Perú: el rol de las familias*. Documento de Trabajo 38. Lima: Grade.
- TORREBLANCA, Alberto (2001). *¿Cuán diferentes son los resultados de las pruebas de matemática y lenguaje en Latinoamérica? Una mirada a través del Laboratorio Latinoamericano de Medición de la Calidad Educativa*, *Revista Creceer* 2, julio 2001. Lima: UMC-Ministerio de Educación.
- UMC y GRADE-UNIDAD DE MEDICIÓN EDUCATIVA Y GRUPO DE ANÁLISIS PARA EL DESARROLLO (2001). *Boletín UMC* 9, “El Perú en el primer estudio internacional comparativo de la UNESCO sobre lenguaje, matemática y factores asociados en el tercer y cuarto grado”. Lima: Ministerio de Educación.

## APÉNDICE 1

### PUBLICACIONES DE LA UNIDAD DE MEDICIÓN DE LA CALIDAD EDUCATIVA

El presente apéndice incluye todas las publicaciones, editadas por la Unidad de Medición de la Calidad Educativa (UMC) o por otros autores, que han usado las bases de datos de las evaluaciones nacionales o internacionales comprendidas en el presente estudio. Todas las publicaciones de la UMC se encuentran disponibles en el portal <[www.minedu.gob.pe/umc](http://www.minedu.gob.pe/umc)>.

#### BOLETINES

*Boletín UMC 27.* Análisis de las respuestas de la evaluación de producción de textos en lenguas nativas (quechua y aimara): Evaluación Nacional 2001. Cuarto grado de primaria. UMC (2004). Lima: Ministerio de Educación.

*Boletín UMC 26.* Las oportunidades de aprendizaje en Matemática: un estudio para 4.º de secundaria. Zambrano, G. y UMC (2004). Lima: Ministerio de Educación. (\*)

*Boletín UMC 25.* Gestión de la dirección en centros educativos polidocentes completos. Espinosa, G.; D. Caro y UMC (2003). Lima: Ministerio de Educación.

*Boletín UMC 24.* Acuerdo entre el profesorado y la construcción del proyecto educativo en centros educativos polidocentes completos. Espinosa G., D. Caro y UMC (2003). Lima: Ministerio de Educación.

*Boletín UMC 23.* Trabajo en equipos dentro del aula. Espinoza G. y UMC (2003). Lima: Ministerio de Educación.

*Boletín UMC 22.* Las oportunidades de aprendizaje en Lógico-Matemática: un estudio para cuarto grado de primaria. Zambrano, G. y UMC (2003). Lima: Ministerio de Educación.

*Boletín UMC 21.* El programa internacional para la evaluación de estudiantes de la OCDE (PISA) y la participación del Perú. UMC (2002). Lima: Ministerio de Educación.

*Boletín UMC 20.* Análisis de ítems de las pruebas CRECER 1998. Producción de textos en quinto grado de secundaria. UMC y GRADE (2002). Lima: Ministerio de Educación.

---

(\*) Estudio utilizado en el meta-análisis.

- Boletín UMC 19.* Análisis de ítems de las pruebas CRECER 1998. Resultados de Comunicación en quinto grado de secundaria. UMC y GRADE (2002). Lima: Ministerio de Educación.
- Boletín UMC 18.* Análisis de ítems de las pruebas CRECER 1998. Resultados de Matemática en quinto grado de secundaria. UMC y GRADE (2002). Lima: Ministerio de Educación.
- Boletín UMC 17.* Análisis de ítems de las pruebas CRECER 1998. Resultados de Comunicación en cuarto grado de secundaria. UMC y GRADE (2002). Lima: Ministerio de Educación.
- Boletín UMC 16.* Análisis de ítems de las pruebas CRECER 1998. Resultados de Matemática en quinto grado de secundaria. UMC y GRADE (2002). Lima: Ministerio de Educación.
- Boletín UMC 15.* Análisis de ítems de las pruebas CRECER 1998. Producción de textos en sexto grado de primaria. UMC y GRADE (2001). Lima: Ministerio de Educación.
- Boletín UMC 14.* Análisis de ítems de las pruebas CRECER 1998. Resultados de Comunicación Integral en sexto grado de primaria. UMC y GRADE (2001). Lima: Ministerio de Educación.
- Boletín UMC 13.* Análisis de ítems de las pruebas CRECER 1998. Resultados de Lógico-Matemática en sexto grado de primaria. UMC y GRADE (2001). Lima: Ministerio de Educación.
- Boletín UMC 12.* Análisis de ítems de las pruebas CRECER 1998. Producción de textos en cuarto grado de primaria. UMC y GRADE (2001). Lima: Ministerio de Educación.
- Boletín UMC 11.* Análisis de ítems de las pruebas CRECER 1998. Resultados de Comunicación Integral en cuarto grado de primaria. UMC y GRADE (2001). Lima: Ministerio de Educación.
- Boletín UMC 10.* Análisis de ítems de las pruebas CRECER 1998. Resultados de Lógico-Matemática en cuarto grado de primaria. UMC y GRADE (2001). Lima: Ministerio de Educación.
- Boletín UMC 9.* El Perú en el primer estudio internacional comparativo de la UNESCO sobre lenguaje, matemática y factores asociados en el tercer y cuarto grado. UMC y GRADE (2001). Lima: Ministerio de Educación.
- Boletín UMC 8.* Efecto de la escuela en el rendimiento en Lógico-Matemática en cuarto grado de primaria. UMC y GRADE (2001). Lima: Ministerio de Educación.

- Boletín CRECER 7.* Resultados de las pruebas de Ciencias Sociales y Ciencias Naturales. Evaluación Nacional de 1998. UMC (2001). Lima: Ministerio de Educación.
- Boletín CRECER 5/6.* Resultados de las pruebas de Matemática y Lenguaje. ¿Qué aprendimos a partir de la evaluación CRECER 1998? UMC y GRADE (2000). Lima: Ministerio de Educación.
- Boletín CRECER 4.* Las escuelas y las expectativas de las madres y los padres. UMC y GRADE (2000). Lima: Ministerio de Educación.
- Boletín CRECER 3.* Las tareas escolares. UMC y GRADE (2000). Lima: Ministerio de Educación.
- Boletín CRECER 2.* ¿Te gustan las clases de Matemáticas? ¿Y las clases de Lenguaje? UMC y GRADE (2000). Lima: Ministerio de Educación.
- Boletín CRECER 1.* Algunos aspectos de la formación docente en el Perú. UMC y GRADE (1999). Lima: Ministerio de Educación.

## REVISTAS

### ***Revista Crecer* n.º 2, julio del 2001. Lima: UMC-Ministerio de Educación**

- BELLO, Manuel. "Hacia una educación basada en evidencias".
- RODRÍGUEZ, José y Santiago CUETO. "¿Cuánto aprenden nuestros estudiantes? Presentación de la Evaluación Nacional del 2001".
- TORREBLANCA, Alberto. "¿Cuán diferentes son los resultados de las pruebas de matemática y lenguaje en Latinoamérica? Una mirada a través del Laboratorio Latinoamericano de Medición de la Calidad Educativa".
- TRAHTEMBERG, León. "La polémica tarea de evaluar la calidad".
- VALVERDE, Gilbert A. "La política educativa peruana y las evaluaciones educacionales internacionales".
- VARGAS, Silvana. "El programa Internacional para la Evaluación de Estudiantes (PISA-OCDE): ¿se está formando capital humano joven en el Perú?".
- VEGA, Juan Fernando. "Comentarios respecto a la participación del Perú en el Laboratorio Latinoamericano de Medición de la Calidad Educativa".

### ***Revista Crecer*, n.º 1, abril del 2000. Lima: UMC-Ministerio de Educación**

- BECERRA, Hernán y David SIFUENTES. "Modelos para entender la comprensión de textos".

GARCÍA, Felipe. “Trabajando para compartir una información confiable, oportuna y pertinente”.

CUADROS, María Esther (entrevista con). “Cómo y para qué evaluamos en CRECER”.

CUETO, Santiago. “Retos para la enseñanza y el aprendizaje en el Perú”.

GALARZA, Néstor. “Innovar en la enseñanza de las Ciencias Naturales”.

PALACIOS, María Luisa. “Historia: conocer el pasado para comprender el presente”, *Revista Crecer* 1, abril 2000. Lima: UMC-Ministerio de Educación.

SAAVEDRA, Hólger. “Romper la tradición para acercar la matemática a cada estudiante”.

TRAHTEMBERG, León. Reflexiones en torno a la “medición de la calidad educativa”, *Revista Crecer* 1, abril 2000. Lima: UMC-Ministerio de Educación.

UMC. “Sistema Nacional de Evaluación del Rendimiento Escolar en América Latina”.

## DOCUMENTOS DE TRABAJO

*Año 2004*

Documento de Trabajo 11

MONTANÉ, Angélica; Fernando LLANOS y Jéssica TAPIA. *La evaluación de la alfabetización lectora de PISA y el rendimiento de los estudiantes peruanos.*

Documento de Trabajo 10 (\*)

ASMAD, Úrsula; David PALOMINO, Mary TAM y Gloria ZAMBRANO. *Una aproximación a la alfabetización matemática y científica de los estudiantes peruanos de 15 años. Resultados del Perú en la evaluación internacional PISA.*

Documento de Trabajo 9 (\*)

EQUIPO DE ANÁLISIS DE LA UNIDAD DE MEDICIÓN DE LA CALIDAD EDUCATIVA. *Factores asociados al rendimiento estudiantil. Resultados de Evaluación Nacional 2001.*

Documento de Trabajo 8

DIBÓS, Alessandra; Susana FRISANCHO y Yolanda ROJO. *Propuesta de evaluación de Formación Ciudadana.*

Documento de Trabajo 7

PALOMINO, David. *El aprendizaje de la medición. Análisis de las pruebas de material concreto aplicadas en la Evaluación Nacional 2001 a alumnos peruanos de cuarto y sexto grados de primaria.*

---

(\*) Estudios utilizados en el meta-análisis.

## Documento de Trabajo 6 (\*)

CARO, Daniel (coordinador). *Una aproximación a la alfabetización lectora de los estudiantes peruanos de 15 años. Resultado del Perú en la evaluación internacional PISA.*

*Año 2003*

## Documento de Trabajo 5

MONTANÉ, Angélica (coordinadora). *Cómo rinden los estudiantes peruanos en Comunicación y Matemática: resultados de la Evaluación Nacional 2001. Cuarto y sexto grados de primaria. Informe pedagógico.*

## Documento de Trabajo 4

MONTANÉ, Angélica (coordinadora). *Cómo rinden los estudiantes peruanos en Comunicación y Matemática: Resultados de la Evaluación Nacional 2001. Cuarto grado de secundaria. Informe pedagógico.*

## Documento de Trabajo 3

MONTANÉ, Angélica (coordinadora). *Cómo rinden los estudiantes peruanos en Comunicación y Matemática: resultados de la Evaluación Nacional 2001. Sexto grado de primaria. Informe pedagógico.*

## Documento de Trabajo 2

MONTANÉ, Angélica (coordinadora). *Cómo rinden los estudiantes peruanos en Comunicación y Matemática: resultados de la Evaluación Nacional 2001. Cuarto grado de primaria. Informe pedagógico.*

## Documento de Trabajo 1

ESPINOSA, Giuliana y Alberto TORREBLANCA. *Cómo rinden los estudiantes peruanos en Comunicación y Matemática: resultados de la Evaluación Nacional 2001. Informe descriptivo.*

**ARTÍCULOS**

BURGA, Andrés (2005). "Evaluación del rendimiento académico. Introducción a la teoría de respuesta al ítem. Lima: UMC-Ministerio de Educación". Ver en <<http://www.minedu.gob.pe/umc/articulosUMC.php>>.

---

(\*) Estudio utilizado en el meta-análisis.

BURGA, Andrés (2005). “La unidimensionalidad de un instrumento de medición. Perspectiva factorial”. Ver en <<http://www.minedu.gob.pe/umc/publicaciones/artiumc/2.pdf>>.

## ENCARTES UMC

UMC (2004). *Evaluar para aprender. Todo lo que usted necesita saber sobre las evaluaciones nacionales del rendimiento estudiantil*. Lima: Ministerio de Educación.

UMC (2005). *¿Cuánto aprenden sus hijos en la escuela? Resultados de la Evaluación Nacional 2004*. Lima: Ministerio de Educación.

## SOBRE LAS EVALUACIONES

### Evaluación Nacional 2004

#### *Marcos de trabajo*

UMC (2004). Marco de trabajo de la evaluación nacional 2004. Lima: Ministerio de Educación.

UMC (2005). Evaluación Nacional 2004. Marco de trabajo de las pruebas de rendimiento. Lima: Ministerio de Educación.

UMC (2005). Evaluación Nacional 2004. Marco de trabajo de los instrumentos de factores asociados al rendimiento estudiantil. Lima: Ministerio de Educación.

UMC (2005). Evaluación Nacional 2004. Marco de trabajo de las pruebas de rendimiento de Comunicación Integral en lenguas nativas. Lima: Ministerio de Educación.

#### *Informes pedagógicos*

UMC (2005). Evaluación Nacional del Rendimiento Estudiantil 2004. Informe pedagógico de resultados. Formación Matemática. Segundo grado de primaria. Sexto grado de primaria. Lima: Ministerio de Educación.

UMC (2005). Evaluación Nacional del Rendimiento Estudiantil 2004. Informe pedagógico de resultados. Comprensión de textos escritos. Segundo grado de primaria. Sexto grado de primaria. Lima: Ministerio de Educación.

UMC (2005). Evaluación Nacional del Rendimiento Estudiantil 2004. Informe pedagógico de resultados. Formación Matemática. Tercer grado de secundaria. Quinto grado de secundaria. Lima: Ministerio de Educación.



UMC (2005). Evaluación Nacional del Rendimiento Estudiantil 2004. Informe pedagógico de resultados. Comprensión de textos escritos. Tercer grado de secundaria. Quinto grado de secundaria. Lima: Ministerio de Educación.

### **Evaluación Nacional 2001**

#### *Documentos técnicos*

CARO, Daniel (2002). Estimación del nivel socioeconómico de las familias: Propuesta metodológica para la Evaluación Nacional de rendimiento del 2001. Lima: Ministerio de Educación.

CRUZ, Gustavo, Giuliana ESPINOSA, Angélica MONTANÉ y Carlos RODRÍGUEZ (2002). Informe técnico de la consulta sobre puntos de corte para la Evaluación Nacional 2001. Lima: UMC-Ministerio de Educación.

TORREBLANCA, Alberto y Daniela ZACHARÍAS (2002). Ficha técnica: diseño muestral Evaluación Nacional 2001. Primaria. Lima: UMC-Ministerio de Educación.

TORREBLANCA, Alberto y Daniela ZACHARÍAS (2002). Ficha técnica: diseño muestral Evaluación Nacional 2001. Secundaria. Lima: UMC-Ministerio de Educación.

UMC y GRADE (2001). Evaluación Nacional del rendimiento escolar 2001. Especificaciones de las pruebas de Lógico-Matemática. Lima: Ministerio de Educación.

UMC y GRADE (2001). Evaluación Nacional del rendimiento escolar 2001. Fundamentación de la elaboración de las pruebas de Lógico Matemática y Matemática. Lima: Ministerio de Educación.

UMC y GRADE (2001). Evaluación Nacional del rendimiento escolar 2001. Fundamentación de la elaboración de las pruebas de Comunicación Integral y Comunicación. Lima: Ministerio de Educación.

UMC y GRADE (2001). Evaluación Nacional del rendimiento escolar 2001. Fundamentación de los instrumentos de factores asociados. Lima: Ministerio de Educación.

UMC (2001). Fundamentación de la evaluación de actitudes en la Evaluación Nacional del 2001. Lima: Ministerio de Educación.

*Informes*

BENAVIDES, Martín y otros (2002). Cuarto de primaria y secundaria: factores asociados al rendimiento en Matemática. Versión preliminar. Lima: UMC-Ministerio de Educación. (\*)

UMC (2002). Resultados de la Evaluación Nacional del rendimiento de los estudiantes 2001. Lima: UMC-Ministerio de Educación.

**Evaluación Nacional 1998**

MINISTERIO DE EDUCACIÓN (1998). Evaluación del rendimiento estudiantil "CRECER". Informe de resultados de la primera prueba nacional. Cuarto grado de educación primaria. Centros educativos polidocentes. Lima: Ministerio de Educación.

**Evaluación Nacional 1996**

MINISTERIO DE EDUCACIÓN (1998). Evaluación del Rendimiento Estudiantil "CRECER". Informe de Resultados de la Primera Prueba Nacional. Lima: MED.

*Documentos técnicos*

TORREBLANCA, Alberto; Arturo CALDERÓN y Cholly FARRO (2000). Diseño muestral en la aplicación nacional CRECER 1996. Lima: UMC-Ministerio de Educación.

BAZÁN, Jorge (2000). Evaluación psicométrica de las preguntas y pruebas CRECER 96. Lima: UMC-Ministerio de Educación.

*Documento en Power Point*

UMC (2005). Resultados de la IV Evaluación Nacional del rendimiento estudiantil 2004. Lima: UMC-Ministerio de Educación.

---

(\*) Estudio utilizado en el meta-análisis.

## APÉNDICE 2

### DOCUMENTOS DEL PROGRAMA DE MEJORAMIENTO DE LA CALIDAD DE LA EDUCACIÓN PRIMARIA (MECEP)

- ARREGUI, Patricia (2001). *¿Y cómo estamos en escritura? Resultados de la prueba de expresión escrita Crecer 1996*. Documento de Trabajo 6. Lima: MECEP-Ministerio de Educación.
- MECEP (1997). *Recomendaciones técnico-pedagógicas para la enseñanza de lenguaje en la educación primaria n.º 1. Comunicación Oral*. “CRECER”, Primera prueba nacional 1996. Lima: MECEP-Ministerio de Educación.
- MECEP (1997). *Recomendaciones técnico-pedagógicas para la enseñanza de lenguaje en la educación primaria n.º 2. Comprensión Lectora*. “CRECER”, Primera prueba nacional 1996. Lima: MECEP-Ministerio de Educación.
- MECEP (1997). *Recomendaciones técnico-pedagógicas para la enseñanza de lenguaje en la educación primaria n.º 3. Nociones gramaticales*. “CRECER”, Primera prueba nacional 1996. Lima: MECEP-Ministerio de Educación.
- MECEP (1997). *Recomendaciones técnico-pedagógicas para la enseñanza de lenguaje en la educación primaria n.º 4. Vocabulario*. “CRECER”, Primera prueba nacional 1996. Lima: MECEP-Ministerio de Educación.
- MECEP (1997). *Recomendaciones técnico-pedagógicas para la enseñanza de lenguaje en la educación primaria n.º 5. Producción de textos*. “CRECER”, Primera prueba nacional 1996. Lima: MECEP-Ministerio de Educación.
- MECEP (1997). *Recomendaciones técnico-pedagógicas para la enseñanza de la matemática en educación primaria n.º 1*. “CRECER”, Primera prueba nacional 1996. Lima: MECEP-Ministerio de Educación.
- MECEP (1997). *Recomendaciones técnico-pedagógicas para la enseñanza de la matemática en educación primaria n.º 2*. “CRECER”, Primera prueba nacional 1996. Lima: MECEP-Ministerio de Educación.
- RODRÍGUEZ, Jorge y Silvana VARGAS (editores) (2002). *Análisis de los resultados y metodología de las pruebas CRECER 1998*. Documento de Trabajo 13. Lima: MECEP-Ministerio de Educación. (\*)

---

(\*) Estudio utilizado en el meta-análisis.

### APÉNDICE 3

#### ANÁLISIS REALIZADOS POR AUTORES FUERA DEL MINISTERIO DE EDUCACIÓN

VALDIVIA, Martín (2004). “La escuela y el rendimiento escolar de los niños en el Perú: sobre cómo la heterogeneidad condiciona el proceso de descentralización”. Draft. Lima: GRADE. (\*)

AGÜERO, Jorge y Santiago. CUETO (2004). “Dime con quién estudias y te diré cómo rindes: *Peer-effects* como determinantes del rendimiento escolar”. Informe de investigación entregado al Consorcio de Investigación Económica y Social. Disponible en <www.consorcio.org>. (\*)

CUETO, Santiago; Fernando ANDRADE y Juan LEÓN (2003). *Las actitudes de los estudiantes peruanos hacia la lectura, la escritura, la matemática y las lenguas indígenas*. Documento de Trabajo 44. Lima: GRADE.

CUETO, Santiago.; Cecilia RAMÍREZ, Juan LEÓN y Oscar PAIN (2003). *Oportunidades de aprendizaje y rendimiento en matemática en una muestra de estudiantes de sexto grado de primaria de Lima*. Documento de Trabajo 43. Lima: GRADE. (\*)

BANCO MUNDIAL (2001). “Peruvian Education at a Cross-Road: Challenges and Opportunities for the 21st Century”. World Bank Country Study. (\*)

---

(\*) Estudios utilizados en el meta-análisis.

APÉNDICE 4  
**RESULTADOS DE LAS EVALUACIONES NACIONALES POR ÁREAS EVALUADAS**

**I. Comunicación (lenguaje)**

**Cuadro 1**  
**Posición relativa de las regiones del Perú en las cuatro evaluaciones nacionales en el área de Comunicación**

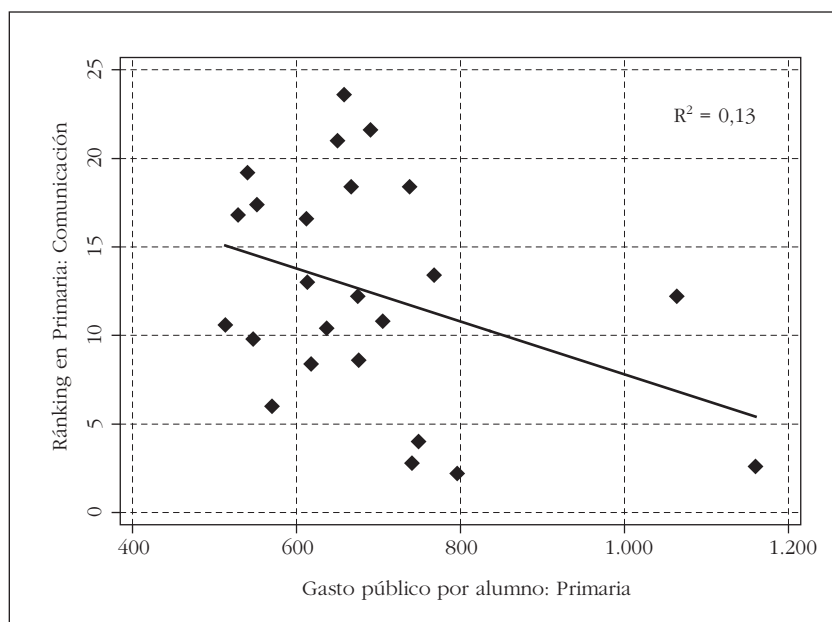
	Primaria						Secundaria						Ránking promedio EN	Ránking de riqueza
	2.º		4.º		6.º		3.º		4.º		5.º			
	EN 2004	EN 1996	EN 1998	EN 1998	EN 1998	EN 2004	EN 2004	EN 1998	EN 2001	EN 1998	EN 2004			
Huancavelica	17	17	18	17	17	20	18	23	24	23	23	23	24	
Huánuco	19	15	14	21	21	15	17	20	15	14	18	17	23	
Puno	11	16	21	22	18	18	18	24	20	19	22	22	22	
Cajamarca	15	13	13	11	17	17	14	21	14	7	12	15	21	
Amazonas	12	5	11	13	16	16	11	19	7	14	15	14	20	
Ayacucho	22	23	22	19	21	21	21	16	22	6	18	15	19	
Loreto	24	22	23	15	23	23	21	18	19	20	11	19	18	
Apurímac	23	24	24	24	24	24	24	22	21	23	24	22	17	
Ucayali	21	21	19	20	19	19	20	14	24	21	19	20	16	
Piura	13	8	9	8	10	10	10	10	13	12	7	12	15	
San Martín	18	19	20	23	14	14	19	13	18	13	21	16	14	
Pasco	14	10	15	14	13	13	13	12	12	11	17	13	13	
Junín	10	9	7	5	12	12	9	7	5	4	5	5	12	
Ancash	16	11	8	7	9	9	10	17	10	16	8	13	11	
Cusco	20	14	16	12	22	22	17	15	15	18	16	16	10	
La Libertad	8	1	4	6	6	6	5	5	6	3	6	5	9	
Lambayeque	6	7	6	9	7	7	7	8	8	8	9	8	8	
Arequipa	4	2	1	1	3	3	2	3	1	1	4	2	7	
Lima	3	3	3	2	2	4	3	4	2	5	1	3	6	
Moquegua	2	6	2	3	3	4	3	2	4	10	2	4	5	
Ica	5	12	10	10	5	5	8	11	11	9	10	10	4	
Madre de Dios	9	20	17	18	8	8	14	6	17	17	13	12	3	
Tacna	1	4	5	4	4	4	3	1	3	2	3	1	2	
Tumbes	7	18	12	16	11	11	13	9	9	22	20	14	1	

**Cuadro 2**  
**Correlaciones entre el ránking regional en pruebas de Comunicación, nivel de riqueza y gasto público en educación**

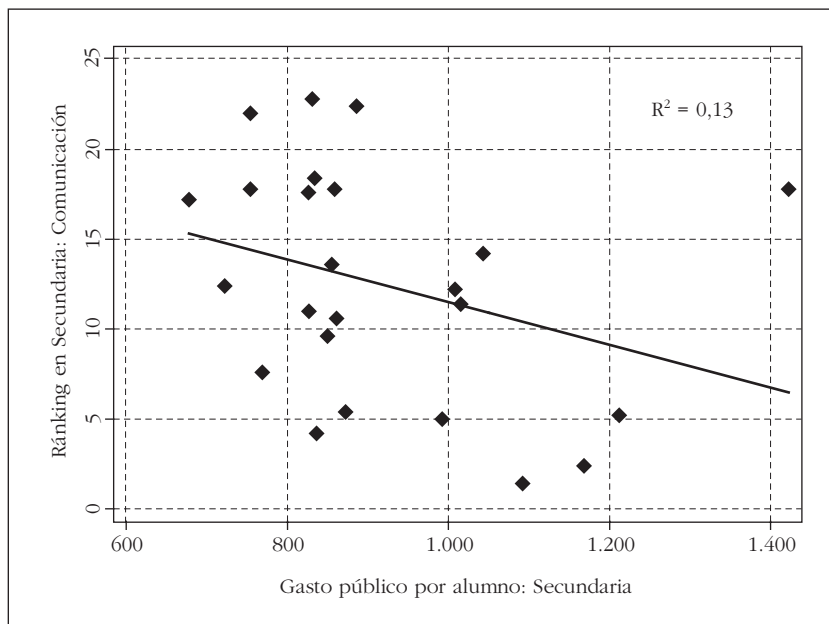
	Riqueza	Primaria	Gasto Secundaria
Ránking en Comunicación - Primaria	0,60 *** (0,00)	-0,30 (0,16)	
Ránking en Comunicación - Secundaria	0,61 *** (0,00)		-0,37 * (0,07)

\*\*\*  $p < 0,01$ ; \*\*  $p < 0,05$ ; \*  $p < 0,10$

**Gráfico 1**  
**Relación entre el promedio regional de evaluaciones nacionales en Comunicación a nivel primario (escuelas públicas solamente) y el gasto público por alumno en educación**



**Gráfico 2**  
**Relación entre el promedio regional de evaluaciones nacionales en Comunicación a nivel secundario (escuelas públicas solamente) y el gasto público por alumno en educación**



## II. Matemática

Cuadro 3  
Posición relativa de las regiones del Perú en las cuatro evaluaciones nacionales en el área de Matemática

	Primaria						Secundaria						Ránking de riqueza		
	2.°		4.°		6.°		Ránking promedio EN	3.°		4.°		5.°		Ránking promedio EN	
	EN 2004	EN 1996	EN 1998	EN 1998	EN 1998	EN 2004		EN 2004	EN 1998	EN 2001	EN 1998	EN 2004			
Huancavelica	17	16	12	15	15	15	15	19	20	22	17	19	19	19	24
Huánuco	21	17	16	19	18	18	18	16	23	13	15	17	17	17	23
Puno	11	11	18	17	17	17	15	24	22	14	18	24	20	20	22
Cajamarca	16	10	11	10	14	14	12	18	9	10	14	16	13	13	21
Amazonas	12	5	10	14	16	16	11	20	8	12	8	18	13	13	20
Ayacucho	18	22	22	18	20	20	20	12	16	7	13	12	12	12	19
Loreto	24	21	23	22	24	24	23	23	18	21	19	22	21	21	18
Apurímac	23	24	24	24	23	23	24	22	17	20	22	20	20	20	17
Ucayali	22	20	21	23	22	22	22	21	24	24	24	23	23	23	16
Piura	15	12	15	9	11	11	12	10	13	6	12	11	10	10	15
San Martín	20	18	20	20	19	19	19	17	19	15	20	21	18	18	14
Pasco	13	2	17	12	9	9	11	8	6	11	11	6	8	8	13
Junín	10	4	6	3	8	8	6	5	3	2	1	4	3	3	12
Áncash	14	14	8	6	12	12	11	15	12	18	7	13	13	13	11
Cusco	19	15	14	13	21	21	16	14	15	16	10	9	10	10	10
La Libertad	7	1	4	8	6	6	5	6	7	9	6	7	7	7	9
Lambayeque	6	9	7	7	7	7	7	9	11	8	16	10	11	11	8
Arequipa	4	3	1	1	3	3	2	2	2	1	2	2	2	2	7
Lima	3	7	2	2	2	2	3	4	1	5	3	5	4	4	6
Moquegua	2	6	3	5	4	4	4	3	5	3	5	3	4	4	5
Ica	5	13	9	11	5	5	9	11	10	23	9	15	14	14	4
Madre de Dios	9	23	19	21	13	13	17	7	21	17	21	8	15	15	3
Tacna	1	8	5	4	1	1	4	1	4	4	4	1	3	3	2
Tumbes	8	19	13	16	10	10	13	13	14	19	23	14	17	17	1



**Cuadro 4**  
**Correlaciones entre el ránking para cada nivel en Matemática,**  
**nivel de riqueza y gasto público en educación**

	Riqueza	Gasto Primaria	Gasto Secundaria
Ránking en Matemática - Primaria	0,54 *** (0,00)	-0,32 (0,12)	
Ránking en Matemática - Secundaria	0,41 ** (0,04)		-0,33 (0,11)

\*\*\*  $p < 0,01$ ; \*\*  $p < 0,05$ , \*  $p < 0,10$ .

**Gráfico 3**  
**Relación entre el ránking regional en pruebas de Matemática a nivel primario**  
**(escuelas públicas solamente) y el gasto público por alumno en educación**

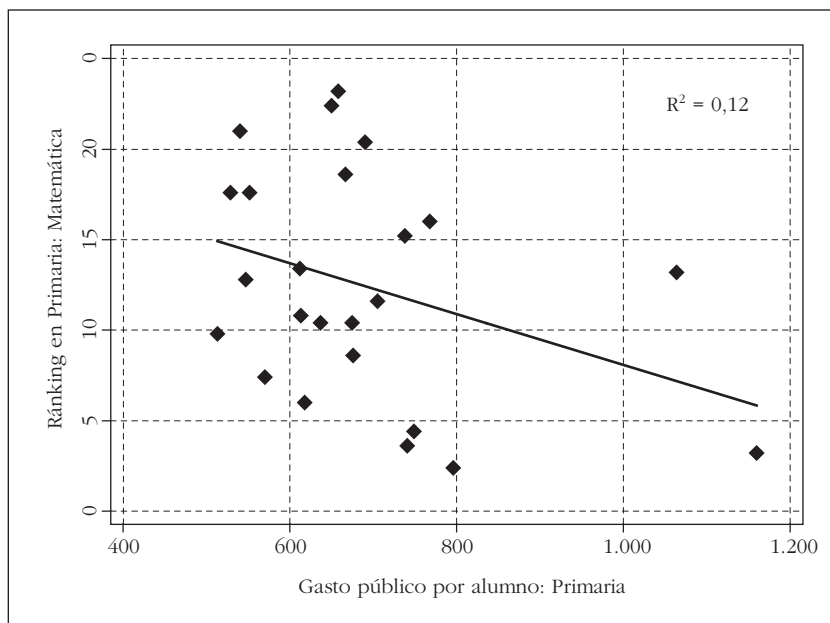
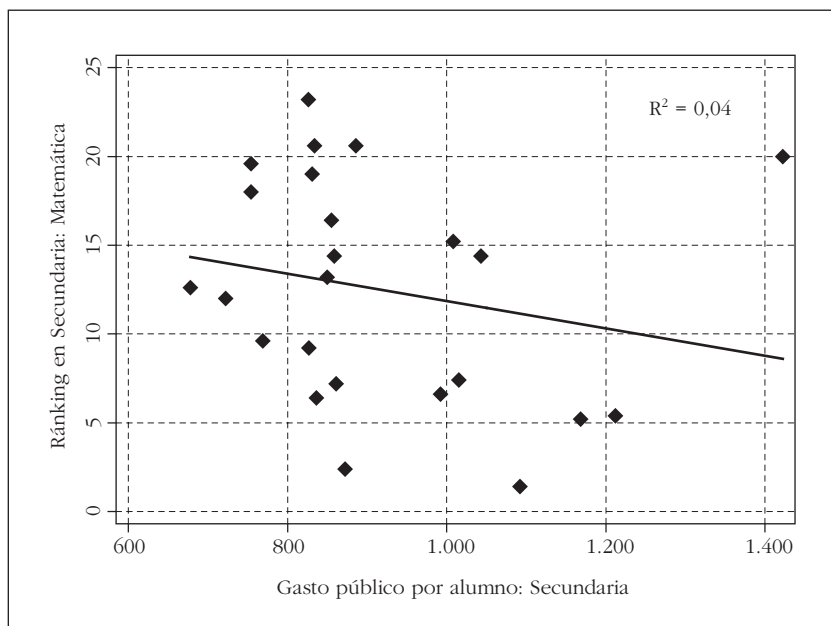


Gráfico 4  
Relación entre el promedio regional de evaluaciones nacionales  
en Matemática a nivel secundario (escuelas públicas solamente)  
y el gasto público por alumno en educación



**APÉNDICE 5**  
**RUBROS CONSIDERADOS PARA LA CONSTRUCCIÓN**  
**DEL INDICADOR DEL GASTO PÚBLICO EN EDUCACIÓN**  
**POR FUNCIÓN, PROGRAMA Y SUBPROGRAMA**

<b>Función</b>	<b>Programa</b>	<b>Subprograma</b>
Educación y cultura	Educación primaria	Enseñanza primaria
		Erradicación del analfabetismo
	Educación secundaria	Formación general
		Formación ocupacional
	Edificaciones escolares	Edificaciones escolares (primaria)
		Edificaciones escolares (secundaria)