

Realizado en 1993, el Diagnóstico General de la Educación Peruana identificó las áreas de mayor necesidad y urgencia para una intervención educativa. Dado este contexto, en 1996 el Ministerio de Educación inicia el Programa de Mejoramiento de la Calidad de la Educación Peruana, MECEP. Cofinanciado por el Banco Mundial y dirigido a los alumnos y profesores de los centros educativos hispanohablantes de nivel primario, el Programa cuenta, entre sus principales líneas de acción: la modernización curricular, la dotación de material educativo y mobiliario escolar, la capacitación docente, la modernización de la gestión educativa, la medición de la calidad educativa, y la construcción y rehabilitación de infraestructura escolar. Al mismo tiempo, el Ministerio de Educación da inicio al Plan Nacional de Educación Bilingüe Intercultural, orientado hacia la población escolar vernaculohablante.

En el año 1997, con la suscripción de un contrato de préstamo con el BID, el Programa MECEP se extiende a la educación inicial para los niños de cinco años. Comienza a intervenir, asimismo, en la educación secundaria y la formación profesional técnica. En el primer caso, aborda los aspectos de modernización curricular, capacitación docente y promoción de innovaciones pedagógicas; en el segundo, los temas relacionados con el diseño del sistema de formación para el trabajo y el catálogo de títulos profesionales.

Desde el año 2001, gracias a un nuevo aporte financiero del BID, se viene trabajando en la Primera Fase del Programa de Mejoramiento de la Calidad de la Educación Secundaria. Entretanto, con el apoyo técnico del Banco Mundial han venido siendo diseñadas las líneas de intervención para mejorar la calidad de la educación rural, revalorizar la carrera magisterial e incorporar las nuevas tecnologías de la información y la comunicación en la educación pública, todo ello a través de nuevos y alentadores proyectos.

ISBN 9972-881-12-1

PROGRAMA ESPECIAL MEJORAMIENTO DE LA CALIDAD DE LA EDUCACION PERUANA

MECEP-MED

Análisis de los resultados y metodología de las pruebas CRECER 1998

PROGRAMA ESPECIAL MEJORAMIENTO DE LA CALIDAD DE LA EDUCACION PERUANA



## Análisis de los resultados y metodología de las pruebas CRECER 1998

José Rodríguez / Silvana Vargas  
Editores

13

DOCUMENTO DE TRABAJO

## ANÁLISIS DE LOS RESULTADOS Y METODOLOGÍA DE LAS PRUEBAS CRECER 1998





# Análisis de los resultados y metodología de las pruebas CRECER 1998

José Rodríguez / Silvana Vargas

Editores

13

DOCUMENTO DE TRABAJO

Esta investigación fue realizada en el marco del Programa Especial Mejoramiento de la Calidad de la Educación Peruana, MECEP. Su contenido es responsabilidad exclusiva de los autores y no compromete al Programa MECEP, al Ministerio de Educación del Perú ni a las entidades cofinanciadoras.

© Ministerio de Educación del Perú, 2002  
Calle Van de Velde N° 160, Lima 41 - Perú  
Teléfono 435 3900  
[www.minedu.gob.pe](http://www.minedu.gob.pe)

ISBN 9972-881-12-1  
Depósito legal N° 1501212002-0483

DISEÑO Y CUIDADO DE EDICIÓN: Rosario Rey de Castro  
COORDINACIÓN POR EL MINISTERIO DE EDUCACIÓN: Martín Oyata  
COMPOSICIÓN DE TEXTOS: Nova Print Digital S.R.L. Teléfono: 423 5117  
IMPRESIÓN: FIMART S.A.C. Editores e Impresores

Reservados todos los derechos.  
Se autoriza a citar o reproducir en todo o en parte el presente documento, siempre y cuando se cite la fuente.

Impreso en Lima, Perú  
Tiraje 1000 ejemplares  
Febrero del 2002

# CONTENIDO

Presentación	11
EL CURRÍCULO IMPLEMENTADO COMO INDICADOR DEL PROCESO EDUCATIVO / <i>Claudia Galindo</i>	13
1. El modelo de oportunidades de aprendizaje	14
2. Una aproximación al cambio curricular peruano	16
3. Currículo intencional e implementado utilizado por los docentes peruanos	18
4. Relación entre el currículo implementado y el currículo aprendido	26
5. Resumen y conclusiones	30
Anexo 1. Muestra y procedimientos	32
Anexo 2. Prueba de Lógico-Matemática y encuestas para profesores	33
Anexo 3. Tamaño de las muestras, número de docentes por grupos	38
ELABORACIÓN Y APLICACIÓN DEL PROYECTO DE DESARROLLO INSTITUCIONAL EN LOS CENTROS EDUCATIVOS DE PRIMARIA Y SECUNDARIA / <i>Claudia Galindo, Francisco Indacochea</i>	39
1. El PDI como herramienta de descentralización y gestión	39
2. Porcentaje de centros educativos que ha elaborado y aplicado el PDI	40
3. Personas que han participado en la elaboración del PDI	43
4. Apoyo de las instancias administrativas superiores para elaborar, desarrollar o implementar el PDI	49
5. Resumen y conclusiones	49
Anexo 1. Preguntas analizadas	51
Anexo 2. Muestra y procedimientos	53
Anexo 3. Número de casos del cuadro 3 y gráficos 5 y 6	54
RENDIMIENTO Y ACTITUDES HACIA LA MATEMÁTICA EN EL SISTEMA ESCOLAR PERUANO / <i>Jorge Bazán, Giuliana Espinosa, Cholly Farro</i>	55
1. Marco teórico: Aproximación al concepto de actitudes	55
2. Método	58
3. Resultados	60
4. Discusión	65
Anexo. Construcción del indicador de actitud	68

¿QUÉ PIENSAN LOS DOCENTES DE SUS ALUMNOS? / <i>Fernando Andrade</i>	71
1. Padres y estudiantes de los CE estatales y no estatales	72
2. Percepción de los docentes sobre sus alumnos	75
3. A modo de cierre	80
PARA EXPLICAR LAS DIFERENCIAS EN EL RENDIMIENTO EN MATEMÁTICA DE CUARTO GRADO EN EL PERÚ URBANO: ANÁLISIS DE RESULTADOS A PARTIR DE UN MODELO BÁSICO / <i>Martín Benavides</i>	83
1. Las tradiciones de investigación, la explicación del rendimiento y los problemas fundamentales	83
2. Preguntas de investigación, muestra, variables del estudio y métodos	87
3. Algunos resultados descriptivos	88
4. Resultados multivariados y multinivel: El caso de cuarto grado en Matemática	99
5. Conclusiones	103
DISEÑO MUESTRAL EN LA APLICACIÓN NACIONAL CRECER 1998 / <i>Arturo Calderón, Cholly Farro, Jorge Bazán</i>	109
1. Antecedentes	109
2. Objetivo	110
3. Diseño muestral	110
4. Estructura general de la muestra	111
Anexo. Fórmula de ponderación	129
ESTIMACIÓN DEL ERROR ESTÁNDAR EN LAS PRUEBAS CRECER 1998 / <i>Cholly Farro, Jorge Bazán, Alberto Torreblanca, Óscar Millones</i>	131
1. Metodología basada en el diseño de la muestra	131
2. Metodología basada en remuestreo	135
3. Conclusiones	136
Anexo 1. Estimación del error estándar de la media de rendimiento en Arequipa y Amazonas basado en el diseño de la muestra	138
Anexo 2. Algoritmo del programa que calcula EE utilizando la técnica de remuestreo Jackknife	140
EVALUACIÓN PSICOMÉTRICA DE LAS PREGUNTAS DE LAS PRUEBAS CRECER 1998 / <i>Jorge Bazán, Óscar Millones</i>	141
1. Metodología para el análisis de las preguntas	141
2. Resultados	145
3. Conclusiones	151
Anexo 1. Análisis de las preguntas observadas en las pruebas CRECER 1998	152
Anexo 2. Glosario de términos psicométricos	164
EVALUACIÓN PSICOMÉTRICA DE LAS PRUEBAS CRECER 1998 / <i>Jorge Bazán, Óscar Millones</i>	171
1. El marco para el análisis de las pruebas	171
2. Constructos y contenidos de las pruebas	180
3. Resultados de los indicadores en la evaluación de las pruebas	182
4. Conclusiones y sugerencias	185
Anexo 1. Tablas de especificaciones de las pruebas CRECER 1998	187
Anexo 2. Desarrollo y construcción de las pruebas	195

*La educación es un derecho humano con un inmenso poder de transformación. En sus cimientos descansan las piedras angulares de la libertad, la democracia y el desarrollo humano sostenible.*

*Cuando se garantiza el derecho a la educación el mundo entero gana.*

*No existe una solución milagrosa para la contravención de ese derecho.*

*La solución empieza por un simple aserto: que en vísperas del siglo XXI no haya mayor prioridad ni misión más importante que la educación para todos.*

KOFI A. ANNAN  
Secretario General de  
Naciones Unidas  
1999



## PRESENTACION

*El compromiso de medir la calidad del sistema educativo de un país involucra un desarrollo de políticas y acciones sustentadas en información confiable y oportuna. Sin embargo, la medición no sólo pasa por la generación e interpretación de información relevante, confiable y apropiada, sino también por su difusión entre todos los actores sociales. Así, devolver la información a elementos clave de la educación es una tarea indispensable para construir una cultura de evaluación. Este trabajo responde al compromiso adquirido para tal efecto por la Unidad de Medición de la Calidad (UMC) del Ministerio de Educación.*

*La UMC es una de las tres unidades que forman la Oficina de Planificación Estratégica y Medición de la Calidad Educativa (PLANMED). Su tarea primordial consiste en producir información útil para la planificación eficiente de la gestión educativa. El objetivo de la UMC es recolectar, analizar y difundir información acerca del sistema educativo nacional (i.e., rendimiento escolar), con el fin de contribuir en el planeamiento y formulación de políticas en el sector. Las actividades prioritarias de la UMC son llevadas a cabo a través de mediciones periódicas nacionales del rendimiento estudiantil en las áreas de aprendizaje de mayor importancia del currículo. Hasta el momento se han llevado a cabo tres operativos: en 1996, 1998 y el 2001.*

*Las evaluaciones nacionales constan, en general, de pruebas que miden el aprendizaje de los alumnos en ciertas áreas. Hasta la fecha, estas áreas han incluido Matemática, Lenguaje, Ciencia y Ambiente, y Personal Social. El proceso de elaboración de las pruebas implica la selección, la validación en campo y en gabinete, y la consulta a expertos acerca del conjunto de ítems. Asimismo, las evaluaciones comprenden una serie de instrumentos complementarios que recopilan información acerca de los alumnos, sus centros educativos, docentes y familias.*

*Esta publicación recoge sobre todo la información que se ha generado a partir de la evaluación nacional CRECER\* 1998. Incluimos nueve informes producidos por la UMC, que pueden ser clasificados en dos grupos: 1) aquellos reportes que se basan en el análisis de resultados y orientados a la discusión teórica en el área de la investigación educativa; y, 2) aquellos orientados a la discusión de precisiones técnicas del diseño de CRECER 1998.*

*El conjunto de informes de corte más pedagógico busca identificar una serie de elementos cruciales que ayuden a explicar el rendimiento escolar en el Perú. Este ejercicio se realiza desde diversos ángulos y privilegia una serie de actores y temas. Así, se presenta información trabajada en las áreas del currículo, el Proyecto de Desarrollo Institucional (PDI), las actitudes de los*

---

\* Crecer con Calidad y Equidad en el Rendimiento.

alumnos hacia el aprendizaje y las percepciones de los docentes sobre sus alumnos. También se presenta un estudio pionero en el país que identifica los factores asociados al rendimiento a través de la aplicación del modelo jerárquico lineal.

Así, el primer grupo incluye “El currículo implementado como indicador del proceso educativo”, de Claudia Galindo; “Elaboración y aplicación del Proyecto de Desarrollo Institucional en los centros educativos de primaria y secundaria”, de Claudia Galindo y Francisco Indacochea; “Rendimiento y actitudes hacia la Matemática en el sistema escolar peruano”, de Jorge Bazán, Giuliana Espinosa y Cholly Farro; “¿Qué piensan los docentes de sus alumnos?”, de Fernando Andrade; y “Para explicar las diferencias en el rendimiento en Matemática de cuarto grado en el Perú urbano: Análisis de resultados a partir de un modelo básico”, de Martín Benavides.

De otro lado, el conjunto de informes de carácter técnico-metodológico contiene discusiones acerca de temas que combinan criterios estadísticos y psicométricos. En esa línea se presentan los trabajos “Diseño muestral en la aplicación nacional CRECER 1998”, de Arturo Calderón, Cholly Farro y Jorge Bazán; “Estimación del error estándar en las pruebas CRECER 1998”, de Cholly Farro, Jorge Bazán, Alberto Torreblanca y Óscar Millones; “Evaluación psicométrica de las preguntas de las pruebas CRECER 1998”, y “Evaluación psicométrica de las pruebas CRECER 1998”, ambos de Jorge Bazán y Óscar Millones.

La UMC espera que el conjunto de reportes aquí presentado estimule la discusión y genere interés en los asuntos relacionados con la medición de sistemas educativos, tanto en términos teóricos como técnicos y de política educativa. Creemos que no podremos estar satisfechos con nuestra labor —y que nuestros objetivos no se cumplirán plenamente— hasta que no se asuma que la medición del desempeño del sistema educativo de un país no depende exclusivamente de una instancia ubicada en la sede central del Ministerio de Educación. Ésta es una tarea que compromete no sólo a los docentes y autoridades del sector, sino también a los alumnos, los padres de familia, los centros de formación docente, los centros de investigación y la sociedad civil. La UMC recibe y agradece cualquier comentario, crítica y sugerencia sobre el trabajo que realiza; éstos pueden hacerse llegar a: [medicion@minedu.gob.pe](mailto:medicion@minedu.gob.pe).

# EL CURRÍCULO IMPLEMENTADO COMO INDICADOR DEL PROCESO EDUCATIVO

CLAUDIA GALINDO

Con frecuencia se afirma que uno de los principales problemas de la educación primaria peruana está relacionado con la calidad de la enseñanza, es decir, con lo que aprenden los estudiantes en la escuela y con la manera como lo hacen. Esta sección del documento de trabajo<sup>1</sup> estudia lo que los docentes enseñan y lo que los estudiantes aprenden. Tomando como base el currículo de cuarto grado de Primaria en Lógico-Matemática, se analizan específicamente el nivel de generalización de la nueva estructura curricular (currículo intencional), las capacidades que los docentes de escuelas polidocentes completas urbanas del país reportan haber tratado en el aula (currículo implementado), y la relación del currículo implementado con los logros de los estudiantes en pruebas de rendimiento de Lógico-Matemática.

Se utiliza como marco de referencia el modelo de oportunidades de aprendizaje (ODA) desarrollado por el Estudio de Oportunidades en Matemática y Ciencia (SMSO por sus siglas en inglés: Survey of Mathematics and Science Opportunities)<sup>2</sup>. En otros documentos se intenta analizar las prácticas pedagógicas de los docentes de aula y su relación con el rendimiento escolar de los alumnos.

Este trabajo es descriptivo correlacional, y busca acercarse al currículo intencional e implementado de Lógico-Matemática. Para el análisis se utiliza información recogida por el Proyecto de Medición de Rendimiento Estudiantil, CRECER 1998<sup>3</sup>. Las pruebas fueron

- 1 Elaborado de manera conjunta con GRADE. Arturo Calderón (de la UMC) participó en el análisis estadístico de la información. La autora desea expresar su agradecimiento a Gilbert Valverde, José Rodríguez y Santiago Cueto por los comentarios y sugerencias; y a los especialistas en Matemática Tania Pacheco, Freddy Raymundo, Holger Saavedra (UMC) y Martha Villavicencio (GRADE) por su participación en la interpretación de los resultados. También se reconocen los aportes de los distintos especialistas de la UMC, quienes en una reunión de trabajo presentaron sus recomendaciones y comentarios.
- 2 En Schmidt, William; Doris Jorde; Leland Cogan; Ignacio Gonzalo; Emilie Barrier; Urs Moser; Katsuhiko Shimizu; Toshio Sawada; Gilbert Valverde; Curtis McKnight; Richard Prawat; David Wiley; Senta Raizen; Edward Britton y Richard Wolfe: *Characterizing Pedagogical Flow. An Investigation of Mathematics and Science Teaching in Six Countries*. Boston: Kluwer Academic Publishers, 1996. En 1991 se formó un grupo de trabajo cuyos objetivos principales fueron: 1) desarrollar un modelo teórico que permitiera entender los principales elementos de las experiencias educativas de los alumnos en clase; y, 2) elaborar un conjunto de instrumentos para obtener información relacionada con los factores asociados al rendimiento de los alumnos de seis países participantes en el TIMSS (Third International Mathematics and Science Study).
- 3 CRECER es el proyecto de evaluación del rendimiento escolar liderado por la UMC. Consiste en aplicar periódicamente pruebas en algunas áreas de desarrollo (Comunicación Integral, Lógico-Matemática, entre otras) y encuestas (para directores, profesores, padres de familia y alumnos) con el fin de retroalimentar los programas y proyectos desarrollados por el Ministerio de Educación (MED) y contribuir con la formulación de las políticas educativas. Los instrumentos fueron aplicados a fines

aplicadas en centros educativos (CE) polidocentes completos<sup>4</sup> de zonas urbanas de diversas partes del país<sup>5</sup> a una muestra de 16 827 alumnos de cuarto grado de primaria, y las encuestas a 572 profesores de Matemática de los alumnos evaluados.

Con el fin de medir el grado de generalización del currículo intencional, se preguntó a los profesores qué estructura curricular usaron durante el año para enseñar Lógico-Matemática<sup>6</sup>. Por otro lado, con el objeto de evaluar el currículo implementado se seleccionaron aquellas capacidades de la estructura curricular básica incluidas en la prueba de rendimiento<sup>7</sup>. Para cada una de las capacidades seleccionadas se preguntó a los profesores si habían preparado a sus alumnos para resolver problemas y ejercicios como los planteados en las pruebas, en qué momento del año trabajaron el tema y, si fuera el caso, por qué no lo habían hecho<sup>8</sup>.

En el primer acápite de este documento se presenta una aproximación teórica al modelo de ODA trabajado por el SMSO como indicador de proceso educativo, de manera general, y a la concepción del currículo en sus tres dimensiones (intencional, implementado y aprendido), de manera específica. En el segundo acápite se hace referencia al cambio curricular en el Perú a raíz del diagnósti-

co educativo realizado en 1993. En el tercero se hace un análisis descriptivo de los resultados en términos del currículo intencional e implementado por los docentes en el aula. En el cuarto se presenta la relación entre la cobertura de las capacidades y el rendimiento de los alumnos. Por último, en el quinto acápite se resumen los hallazgos y se señalan las conclusiones del análisis.

Los resultados se presentan para la muestra total, así como para CE estatales y no estatales. Esta distinción resulta pertinente, ya que en diversos estudios nacionales se han encontrado diferencias significativas entre estas escuelas. Tales diferencias no sólo atañen al rendimiento de los alumnos, sino también a las prácticas pedagógicas y a la formación de los docentes, estilos de gestión, recursos de infraestructura, equipos y materiales, entre otras. Sin embargo, aquí se desarrolla con mayor exhaustividad la información relacionada con las escuelas de gestión estatal, en las que el MED tiene un amplio campo de acción.

Este trabajo adolece de algunas limitaciones, principalmente relacionadas con la muestra<sup>9</sup> y con el hecho de que los profesores constituyeron la única fuente de información para saber si las capacidades fueron enseñadas. No se utilizaron métodos alternativos —por ejemplo, observaciones de aula—. Sin embargo, es una primera aproximación al tema y sugiere aspectos importantes que es necesario profundizar.

## **1 EL MODELO DE OPORTUNIDADES DE APRENDIZAJE**

Los sistemas educativos, sobre todo en períodos de reforma, buscan monitorear el cumplimiento de las metas. Para ello se han utilizado comúnmente indicadores de insumos y de resultados que permiten evaluar los objetivos definidos en ese ámbito. Entre otras cosas, tales indicadores brindan información acerca del tamaño de la brecha entre lo deseado y lo efectivamente logrado; permiten identificar problemas, analizar la efectividad de las políticas, programas y proyectos<sup>10</sup>, y obtener información relaciona-

---

de noviembre de 1998 (término del año escolar). Por ello, se espera que todas las capacidades evaluadas ya hayan sido trabajadas por los profesores.

4 Esto es, en escuelas donde existe por lo menos un profesor para cada grado.

5 Para mayor detalle de la muestra y procedimientos, véase el anexo 1.

6 A pesar de que el nuevo currículo fue establecido en 1997, su adopción ha sido paulatina.

7 No se incluyeron todas las capacidades de la estructura curricular básica.

8 Para mayor detalle de las capacidades evaluadas en las pruebas y de las preguntas de la encuesta para profesores, se recomienda analizar el anexo 2.

9 En algunos casos los tamaños han sido dispares; en otros, demasiado reducidos. Para mayor detalle de los tamaños de muestra analizados, véase el anexo 3.

10 Rodríguez, José y Juan Pablo Silva: "Flujos y stocks en el sistema educativo y los indicadores de eficiencia interna y de escolarización". Lima: MED, 1998 (mimeo).

da con el mejoramiento de las condiciones escolares<sup>11</sup>.

Sin embargo, la visión desde los insumos empleados y los resultados es insuficiente para entender de manera integral lo que sucede en el sistema educativo. Por ejemplo, una escuela puede tener recursos de infraestructura y materiales educativos adecuados (insumos), pero sus alumnos obtienen malos resultados en pruebas de aprendizaje porque los docentes no han utilizado plenamente estos insumos. Surge entonces la necesidad de definir indicadores que brinden información acerca de los procesos que tienen lugar en la escuela. Así se identificarán las causas y posibles soluciones a los problemas.

En este contexto, el concepto de ODA empieza a ser utilizado como herramienta de política, como un indicador importante para entender el proceso de aprendizaje de los alumnos, como el vínculo definido entre las metas y los resultados de los estudiantes<sup>12</sup>, y como mecanismo para generar responsabilidad en la escuela. De esta manera se busca tener una mejor aproximación al complejo proceso educativo que se imparte en ella.

Mc Donnell<sup>13</sup> sostiene que el concepto de ODA se desarrolló inicialmente hace 30 años como un concepto técnico que buscaba medir la validez de las evaluaciones internacionales realizadas por la International Association for the Evaluation of Educational Achievement (IEA), tomando en consideración si los contenidos evaluados habían sido enseñados por los profesores. Luego se incorporó a este concepto una nueva dimensión relacionada con la manera como estos contenidos eran presentados y con las características de los profesores. Por último, se incluyó una dimensión adicional relacionada con los procesos organizacionales y los recursos de la escuela<sup>14</sup>.

El concepto de ODA, que surgió como respuesta a un objetivo específico, se fue haciendo más complejo hasta generar un modelo integral de las oportunidades de aprendizaje de los alumnos como forma de analizar sus experiencias educativas y así entender su rendimiento en la escuela<sup>15</sup>. En este modelo se reconoce la coexistencia de diversos factores que influyen en el rendimiento desde diferentes niveles.

El gráfico de la página siguiente presenta el modelo desarrollado por el SMSO.

Como se observa en el gráfico, en el SMSO se identificaron factores que influyen en el proceso educativo en cuatro niveles diferentes (los sistemas, las escuelas, las clases y los alumnos) para responder a las siguientes preguntas: ¿qué se espera que aprendan los alumnos?; ¿quién imparte la instrucción?; ¿cómo se organiza la instrucción?; y, ¿qué han aprendido los alumnos? Estas preguntas se organizan en torno de dos ejes principales: *el profesor* como articulador y coordinador de la relación entre el alumno y lo que se enseña, y *el currículo*.

Para entender las experiencias educativas es importante analizar lo que se enseña en el aula, es decir, aproximarse al currículo<sup>16</sup> desde sus tres dimensiones<sup>17</sup>: 1) el currículo intencional; 2) el currículo implementado; y, 3) el currículo aprendido.

En el currículo intencional se reflejan los lineamientos, políticas y decisiones oficiales que buscan guiar y dirigir el proceso educativo en el aula. Es considerado como el enlace entre planificadores, profesores y alumnos. En él se definen los contenidos que deben ser enseñados y se determinan las metas de aprendizaje. La estructura curricular es el documento que mejor refleja este currículo.

El currículo implementado es definido como el conjunto de capacidades puestas a disposición de los alumnos a través de las

11 Mc Donnell, Lorraine: "Opportunity to Learn a Research Concept and a Policy Instrument". *Educational Evaluation and Policy Analysis* 17 (3), 1995, pp. 305-322.

12 Valverde, Gilbert: "Evaluación y estándares de currículo", en Benjamín Álvarez y Mónica Ruiz Casares: *Evaluación y reforma educativa: Opciones de política*. Santiago de Chile: PREAL, 1997.

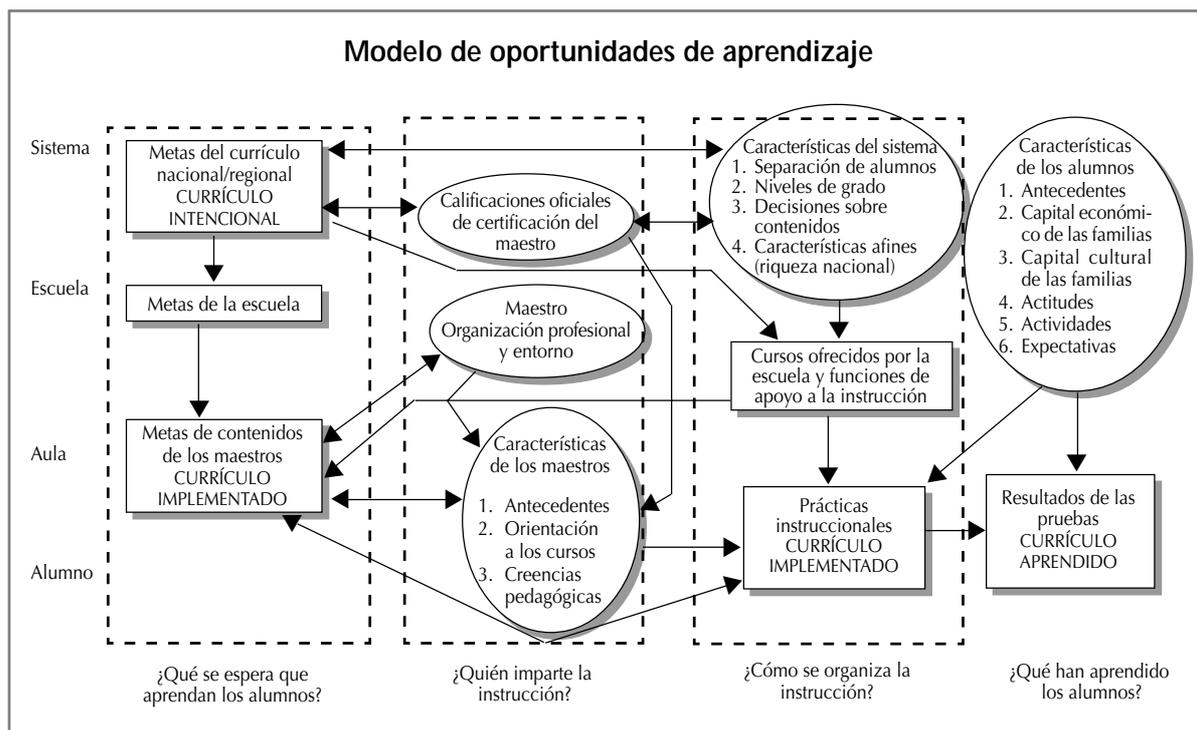
13 "Opportunity to Learn...", *op. cit.*, 1995.

14 En algunas investigaciones como el National Assessment of Education Progress (NAEP) y el Estudio Longitudinal de Estudiantes realizado por el NCES (National Center for Educational Statistics).

15 Este modelo ha sido desarrollado por el SMSO (1996).

16 Esto es, objetivos, metas y oportunidades educativas.

17 El currículo y las dimensiones planteadas han sido desarrolladas por el SMSO en 1996.



prácticas y actividades realizadas por el profesor en el aula. El currículo que se aplica en el aula tiene como base el currículo intencional, pero es el profesor quien usa la estructura formal de manera selectiva y quien define cómo se desarrollan las capacidades, la secuencia de enseñanza, el tiempo dedicado a desarrollar cada tema, la forma de incentivar a sus alumnos para el aprendizaje, entre otros aspectos.

Las decisiones tomadas por el profesor están mediadas por un complejo conjunto de interacciones entre él y sus estudiantes, por ritmos individuales de aprendizaje, por comportamientos y actitudes de los alumnos y por sus percepciones y expectativas frente a un tema. No podemos dejar de mencionar que este escenario está inmerso en uno mayor en el que existen, en el propio ámbito de la escuela, aspectos que se deben tomar en cuenta, como por ejemplo la flexibilidad que le es permitida al profesor para dejar de desarrollar algunos temas o capacidades, o la cantidad de horas disponibles para enseñar el curso. También es importante considerar la influencia que en este proceso ejerce el contexto familiar y cultural externo a la escuela.

Por último, el currículo aprendido representa el conjunto de capacidades efectiva-

mente adquiridas por los alumnos y que pueden ser medidas a través de pruebas de rendimiento.

## 2. UNA APROXIMACIÓN AL CAMBIO CURRICULAR PERUANO

Como consecuencia del Diagnóstico de la Educación Peruana realizado en 1993, el MED está poniendo en práctica una serie de estrategias destinadas a mejorar la calidad de la educación en el país. Una de las acciones identificadas como necesarias por la reforma educativa fue la redefinición de la estructura curricular.

En el currículo vigente en 1993 se encontraron una serie de limitaciones y problemas. En primer lugar, se llegó a la conclusión de que éste no respondía a las necesidades democráticas de una sociedad, al logro de la identidad nacional, al desarrollo histórico y a las demandas productivas del país. En segundo lugar, se reconocieron dificultades tanto en su concepción como en sus posibilidades de concreción, dada la falta de coherencia interna y externa entre sus componentes. En tercer lugar, se pudo observar una desconexión con la realidad y un exce-

sivo énfasis en conductas mecánicas y prácticas memorísticas para la adquisición de conocimientos. Por último, se halló que el currículo no respondía a las exigencias de modernización que debía enfrentar el sistema educativo peruano, a la flexibilidad requerida para hacer frente con éxito a los desafíos del país, a los intereses y necesidades de los alumnos y a las aspiraciones de los padres de familia<sup>18</sup>.

Por ello, desde 1996 se está formulando y difundiendo una nueva estructura curricular para la educación primaria que busca resolver las limitaciones y problemas encontrados, y que propone un cambio radical del proceso de enseñanza y aprendizaje para brindar a los individuos herramientas que les permitan contribuir con el desarrollo social, político y económico del país. La nueva estructura curricular es definida como la expresión explícita de las orientaciones para el proceso educativo, y se le reconocen dos funciones principales: 1) asegurar la unidad del sistema educativo; y, 2) servir de base para la construcción de los proyectos curriculares de los CE, fomentando la participación activa de directores, profesores, padres de familia y especialistas locales<sup>19</sup>.

En la estructura curricular se reflejan las innovaciones y nuevas tendencias en términos de competencias y prácticas pedagógicas que deberían ser desarrolladas en el aula por el profesor. Los alumnos son considerados como el centro del aprendizaje; así, dejan de ser vistos y tratados como estancos de información y, en cambio, son considerados participantes activos en la construcción de su propio conocimiento. Los profesores, en su calidad de facilitadores y responsables de la planificación y organización coherente de las actividades, son piezas clave en este proceso de intercambio. Por último, es necesario que las escuelas sean agentes de cambio y ambientes de aprendizaje integral. Así, no sólo deben ser responsables de la adquisición de conocimientos, sino también de la transmisión de habilidades y de la motivación de los alumnos para involucrarse con el aprendizaje. En tanto organizaciones activas, las escuelas podrán tomar decisiones relacionadas con los contenidos en los que es preciso poner énfasis,

con la secuencia por seguir, con la adecuación de las competencias a su realidad local, entre otras.

La nueva estructura curricular se organiza en tres ciclos de dos grados cada uno y cinco áreas de desarrollo (o cursos) comunes a todos los ciclos. Así se responde a la necesidad de contar con mayor tiempo para que los procesos de aprendizaje de los alumnos se consoliden respetando los ritmos particulares.

El programa curricular básico para el ciclo II<sup>20</sup> (tercer y cuarto grados) de educación primaria se difundió<sup>21</sup> en 1988 en todo el país. En este ciclo se espera que los alumnos afirmen las capacidades adquiridas en el primero y que consoliden la formación de estructuras de conocimiento y conceptos que sirvan de base para nuevos aprendizajes.

En particular, como la Matemática es una disciplina de estructura de pensamiento formal, se considera que el aprendizaje en esta área contribuye de manera significativa con el desarrollo de estructuras cognitivas, del pensamiento creativo, de la capacidad de análisis y de crítica, y de actitudes como la confianza en las habilidades propias, la perseverancia y el gusto por aprender. Para construir conceptos y desarrollar la capacidad de abstracción, se pone énfasis en la importancia de establecer relaciones entre situaciones y objetos, ya que así el proceso inductivo parte de experiencias concretas (manipulación de materiales; elaboración de dibujos, gráficos y esquemas; juegos espaciales, etcétera)<sup>22</sup>. Más allá de estimular una serie de habilidades y actitudes relevantes para el

---

18 Sánchez, Rodolfo; Rosa María Saco y Carmen Lopera: *Diagnóstico del currículo de la educación primaria. Diagnóstico general de la educación*. Lima: Ministerio de Educación, Programa de las Naciones Unidas para el Desarrollo y Agencia Alemana de Cooperación, 1993.

19 *Estructura curricular básica de educación primaria. Segundo ciclo*. Lima: MED-Dirección Nacional de Educación Primaria e Inicial, 1997.

20 El programa curricular básico para el primer ciclo se generalizó en 1997.

21 Es decir, adquirió valor oficial.

22 *Estructura curricular básica de educación primaria. Segundo ciclo, op. cit.*

**Cuadro 1**  
**Programa curricular utilizado para enseñar Lógico-Matemática.**  
**Resultados de la muestra total, y de escuelas estatales y no estatales**  
**(Porcentajes)**

Programa curricular que usa el profesor	Total	Estatal	No estatal
PC básico de ciclo II de primaria (nuevo)	86,8	91,4	61,1
PC de educación primaria (anterior)	10,9	7,2	31,7
Otros <sup>23</sup>	2,4	1,4	6,9
Total	100,0	100,0	100,0
Número de casos	562	477	85

PC: programa curricular.

óptimo desarrollo del individuo, la Matemática es considerada como una materia instrumental debido a que proporciona un lenguaje que permite acceder a otros campos del conocimiento.

En las orientaciones metodológicas para trabajar el área de Lógico-Matemática se recomienda que el profesor defina la secuencia de trabajo considerando la estructura interna del área, pero también incorporando aspectos como la construcción de los conceptos, la influencia del medio sociocultural, entre otros. La secuencia de trabajo debe ser independiente de la secuencia de presentación de las capacidades en el documento.

Se recomienda, asimismo, trabajar esta área en forma de espiral; es decir, los conceptos centrales deben ser trabajados reiteradas veces y, dependiendo del progreso de los alumnos, se irán introduciendo temas nuevos y ampliando los niveles de elaboración y dificultad de los temas que fueron desarrollados previamente. Sin embargo, se hace hincapié en que el nivel de dificultad debe ser graduado y en que las experiencias de aprendizaje deben ser agradables para los estudiantes.

A continuación se presentan los resultados encontrados en el estudio del currículo intencional e implementado utilizado por los profesores de Lógico-Matemática de cuarto de primaria y de la relación entre la enseñanza

de las capacidades y el rendimiento de los alumnos. Estos resultados son representativos de los docentes y del estudiante promedio de este tipo de escuelas en todo el país.

### **3. CURRÍCULO INTENCIONAL E IMPLEMENTADO UTILIZADO POR LOS DOCENTES PERUANOS**

De lo señalado hasta aquí se puede concluir que el currículo es uno de los ejes del modelo de oportunidades de aprendizaje desarrollado por el SMSO. Aun cuando se reconoce también la importancia de las otras dimensiones (las prácticas pedagógicas y el profesor como guía del proceso, por ejemplo), aquí nos aproximamos al currículo nacional a través del currículo intencional e implementado.

En este acápite se analiza la difusión de la estructura curricular del ciclo II de primaria y se observa cómo está siendo implementado el currículo.

#### **¿CUÁN GENERALIZADO ESTÁ EL USO DEL CURRÍCULO INTENCIONAL ENTRE LOS DOCENTES DE LÓGICO-MATEMÁTICA?**

Sobre la base de los recientes esfuerzos por difundir el nuevo programa curricular, se consideró pertinente conocer qué estructura curricular usan los profesores en el aula. El cuadro 1 presenta los resultados de la muestra total, y de las escuelas estatales y no estatales.

23 Otros incluye: PC de escuelas unitarias o de educación bilingüe intercultural, otro programa curricular, y el desarrollo del curso sin programa curricular sobre la base de su experiencia.

Allí se puede observar que la mayoría de profesores de la muestra total usa el nuevo programa curricular<sup>24</sup>. Además, que el porcentaje de los que lo usan es mayor en las escuelas estatales que en las no estatales. Estos resultados pueden ser explicados porque en 1998 el uso del nuevo programa curricular era obligatorio para las escuelas estatales pero no para las no estatales. Entonces, se puede concluir que el nuevo programa curricular es ampliamente utilizado en los CE estatales.

### **¿CÓMO ESTÁ SIENDO IMPLEMENTADO EL CURRÍCULO INTENCIONAL POR LOS PROFESORES DE LÓGICO-MATEMÁTICA?**

Para aproximarnos al currículo no basta conocer cuál es el programa curricular que usan los profesores de aula; también es necesario analizar el currículo implementado, es decir, observar cuáles de las capacidades definidas en él han sido cubiertas en el aula. En este acápite se analiza cuán extendida está la enseñanza del conjunto de capacidades evaluadas y de cada una de ellas. A la vez, se observa si son desarrolladas en momentos específicos o a lo largo del año, y, también, si no han sido enseñadas porque se desarrollan en el currículo de un grado posterior.

Como una muestra de las capacidades del programa curricular, se han utilizado aquellas evaluadas en la prueba de Lógico-Matemática de 1998. Dado que el nivel de difusión y uso de la estructura curricular es amplio y que las pruebas fueron aplicadas al final del año escolar, se esperaban porcentajes altos de profesores que declararan haber trabajado todas las capacidades evaluadas.

En el cuadro 2 se presentan las capacidades del currículo evaluadas en la prueba de Lógico-Matemática de cuarto de primaria.

### **¿Cuán extendida está la enseñanza del conjunto de capacidades evaluadas entre los profesores de Lógico-Matemática?<sup>25</sup>**

Como se puede observar en el cuadro 3, el porcentaje de profesores de la muestra total

que reporta no haber preparado a sus alumnos en todas las capacidades evaluadas es alto. Dos hipótesis podrían explicar este resultado. Por un lado, los programas curriculares son demasiado amplios y el tiempo establecido para cubrir el curso no es suficiente. Por otro, los profesores todavía se estaban adaptando al uso del nuevo programa curricular. De ser cierta esta segunda hipótesis, se esperaría encontrar diferencias en la proporción de capacidades enseñadas según el programa curricular que utilizan los profesores.

Adicionalmente, se puede observar que el porcentaje de profesores que enseña todas las capacidades es mayor en las escuelas no estatales que en las estatales, a pesar de que sólo 61,1% de los primeros declaró haber adoptado el nuevo programa curricular, frente a 91,4% de los segundos que señaló haberlo hecho<sup>26</sup>. Esto sugiere que las capacidades evaluadas, que provienen del nuevo programa, también son cubiertas en otros programas curriculares. Por otro lado, estos resultados podrían estar relacionados con el tiempo disponible para la enseñanza, que suele ser mayor en centros educativos no estatales (el año escolar tiene más semanas y el día escolar suele tener más horas).

Al analizar sólo los resultados de las escuelas estatales considerando el programa curricular que utilizan los docentes (véase el cuadro 4), se observa que el porcentaje de docentes que cubren todas las capacidades al utilizar el nuevo programa es mayor que el de docentes que aplican el programa tradicional. Sin embargo, la diferencia entre estos dos grupos no es estadísticamente significativa. Una

24 En Lógico-Matemática, 13,3% de la muestra total usa otro programa.

25 Para mayor detalle del procedimiento para generar la proporción de capacidades cubiertas, véase el anexo 2.

26 Se aplicó la prueba de diferencias Mann-Whitney, ya que la proporción de capacidades enseñadas no tiene una distribución normal. La diferencia entre los resultados de escuelas estatales y no estatales es estadísticamente significativa con un nivel de confianza del 95% ( $p < 0,05$ );  $p < 0,05$  significa que la probabilidad de encontrar resultados como los reportados por azar es menor de 5%; por lo tanto, se asume que la diferencia es verdadera.

**Cuadro 2**  
**Capacidades evaluadas en la prueba de Lógico-Matemática**

Conocimiento de números y numeración	Habilidades operativas de cálculo	Medición	Organización del espacio: Iniciación de la Geometría
Compara y ordena fracciones homogéneas usando la terminología pertinente y su respectiva gráfica.	Resuelve situaciones problemáticas usando operaciones fundamentales (suma, resta, multiplicación y división) de números naturales menores de 99 000.	Resuelve problemas de la vida diaria de compra y venta relacionados con unidades usuales de masa.	Reconoce rectas perpendiculares, paralelas y ángulos. Identifica y clasifica triángulos y cuadriláteros.
Registra e interpreta información cuantitativa correspondiente a su entorno, estableciendo relación de orden en los números naturales menores de 99 000.	Resuelve ejercicios usando operaciones fundamentales con números naturales menores de 99 000.	Resuelve problemas de la vida diaria relacionados con unidades usuales de tiempo.	Identifica ciertas características de cuerpos geométricos.
Compara y ordena números decimales hasta el orden de las centésimas.	<p>Convierte fracciones a números decimales hasta las centésimas.</p> <p>Resuelve problemas de la vida cotidiana relacionados con adición y sustracción de fracciones homogéneas.</p> <p>Resuelve problemas de la vida cotidiana usando operaciones de adición y sustracción de números decimales hasta el orden de las centésimas.</p>	<p>Estima longitudes y masa en unidades convencionales y no convencionales.</p> <p>Resuelve problemas de la vida diaria de compra y venta relacionados con unidades usuales de longitud.</p>	

**Cuadro 3**  
**Porcentaje de profesores que ha cubierto todas las capacidades evaluadas. Resultados de la muestra total, y de escuelas estatales y no estatales**

	Muestra total	Escuelas estatales	Escuelas no estatales
Todas las capacidades evaluadas	51,4	49,6	61,6
Número de casos	294	241	53
Número de docentes de la muestra	572	486	86

**Cuadro 4**  
**Porcentaje de profesores que ha cubierto todas las capacidades evaluadas según programa curricular.**  
**Resultados de escuelas estatales**

	PC nuevo	PC tradicional	Promedio
Todas las capacidades evaluadas	49,8	38,2	49,6
Número de casos	217	13	241
Número de docentes de la muestra total	436	50	486

posible explicación para ello es que los tamaños de los grupos son muy dispares, y en uno de ellos hay una cantidad reducida de casos.

**¿Cuán extendida está la enseñanza de cada una de las capacidades evaluadas entre los profesores de Lógico-Matemática?**

Como se puede observar en el cuadro 5, las capacidades trabajadas por un mayor porcentaje de docentes están relacionadas con los números naturales menores de 99 000 (establecer re-

laciones de orden, resolver ejercicios y problemas). Esto era esperable, pues los números naturales constituyen ejes fundamentales para enseñar Matemática y son trabajados por los profesores de aula desde grados anteriores<sup>27</sup>.

Las capacidades relacionadas con las fracciones homogéneas y números decimales (comparar y ordenar, resolver problemas de adición y sustracción) han sido trabajadas en segundo lugar. Es importante resaltar, sin embargo, que en la resolución de ejercicios en los que se combinan capacidades vinculadas al uso de fracciones y números decimales<sup>28</sup>, un alto porcentaje de profesores (26,86%) re-

**Cuadro 5**  
**Capacidad. Resultados de la muestra total**

Capacidad	Porcentaje
Establece relaciones de orden en series y secuencias con números naturales	97,2
Resuelve ejercicios usando operaciones básicas con números naturales	95,5
Resuelve problemas usando operaciones básicas con números naturales <sup>29</sup>	95,1
Compara y ordena números decimales hasta las centésimas	91,4
Resuelve problemas relacionados con la adición y sustracción de fracciones homogéneas	90,0
Resuelve problemas relacionados con la adición de números decimales	89,6
Resuelve problemas de compra y venta con unidades de masa	88,7
Compara y ordena fracciones homogéneas	88,3
Resuelve problemas de compra y venta con unidades de longitud	84,4
Resuelve problemas con unidades de tiempo	75,6
Estima longitudes y peso en unidades convencionales	75,1
Resuelve ejercicios convirtiendo fracciones a números decimales	73,1
Reconoce rectas perpendiculares, paralelas y ángulos; triángulos y cuadriláteros	61,6
Reconoce cubos, prismas y pirámides	52,6

27 Los alumnos son introducidos en el conocimiento de los números naturales desde el primer grado de primaria.

28 La capacidad evaluada es la conversión de fracciones a números decimales.

29 Con números menores de 99 000.

porta no haber preparado a sus alumnos. Este porcentaje es relativamente alto si se lo compara con los observados para las capacidades donde se trabajan estos temas de manera independiente. Una posible explicación para estos resultados es que en el nuevo programa curricular las fracciones y decimales son introducidos sólo en cuarto grado, de manera que queda para grados más avanzados el uso combinado de estas capacidades, en tanto requieren de un mayor nivel de complejidad cognitiva.

Las capacidades relacionadas con la resolución de problemas son desarrolladas por un menor porcentaje de profesores. Se pueden observar porcentajes menores en los problemas relacionados con unidades de longitud, estimación de longitudes y peso, y con unidades de tiempo. Algunos autores consideran que la resolución de problemas es más difícil que la resolución de ejercicios, ya que implica poner en práctica un mayor número de estrategias: 1) comprender el enunciado; 2) traducir el lenguaje natural al lenguaje simbólico; 3) tomar decisiones; y, 4) seleccionar alguna de las estrategias posibles para resolverlos<sup>30</sup>. Sin embargo, esta dificultad no es independiente del contenido temático de los problemas. Por ejemplo, se han observado porcentajes más altos de profesores que enseñan problemas relacionados con números naturales menores de 99 000, adición y sustracción de fracciones homogéneas, con compra y venta usando unidades de masa y con adición de números decimales.

Dado el alto porcentaje de profesores que no desarrollan las capacidades relacionadas con la resolución de problemas con unidades de tiempo y con estimación de longitudes y peso, se consideró necesario esbozar algunas hipótesis que expliquen esta situación. A pesar de que la resolución de problemas con unidades de tiempo no es una innovación del nuevo programa, el porcentaje de profesores que no la desarrolla es alto. Una posible explicación de los resultados es que para resolver estos problemas se requiere incorporar el

sistema de numeración sexagesimal, sistema distinto del decimal, trabajado a lo largo de todo el grado. Por ello, probablemente algunos profesores podrían preferir trabajarlos en grados posteriores. Por el contrario, las estimaciones son innovaciones recién incorporadas en el nuevo programa; su desarrollo no ha sido muy elaborado en los libros de texto y algunos profesores podrían mostrarse reacios a trabajarlas, ya que es posible que perciban que las estimaciones no pertenecen a la Matemática, tradicionalmente conceptualizada como ciencia exacta.

Por último, las capacidades relacionadas con la iniciación a la Geometría son enseñadas por un menor número de profesores; aquellas vinculadas con el reconocimiento de rectas, ángulos y con la identificación de triángulos o cuadriláteros no han sido trabajadas por 38,41% de los profesores, y las relacionadas con el reconocimiento de sólidos geométricos no han sido trabajadas por 47,40% de ellos. La enseñanza de estas capacidades es una innovación del nuevo programa para cuarto grado de primaria, pues en programas curriculares anteriores se las enseñaba en grados superiores. Al parecer, los profesores no han logrado cumplir del todo el programa curricular, y son las capacidades relacionadas con la Geometría las que han sido dejadas de lado con mayor frecuencia para ser trabajadas en grados posteriores.

En el cuadro 6 se puede observar que el porcentaje de profesores que ha trabajado cada capacidad tiende a ser mayor o igual en las escuelas no estatales que en las estatales. Se han encontrado diferencias mayores de 10 puntos porcentuales en las siguientes capacidades: compara y ordena fracciones homogéneas; resuelve ejercicios convirtiendo fracciones a números decimales; reconoce rectas perpendiculares, paralelas y ángulos, triángulos y cuadriláteros; y reconoce cubos, prismas y pirámides. Estos resultados sugieren que en las escuelas no estatales se enseña un mayor número de capacidades que en las estatales, especialmente las capacidades relacionadas con fracciones y números decimales, y que la iniciación a la geometría es desarrollada por un mayor grupo de profesores de estas escuelas.

30 Pozo, Juan Ignacio *et al.*: *La solución de problemas*. Madrid: Santillana, 1994.

**Cuadro 6**  
**Porcentaje de profesores que ha enseñado cada capacidad.**  
**Resultados de escuelas estatales y no estatales**

Capacidad	Porcentaje CE estatales	Porcentaje CE no estatales
Establece relaciones de orden en series y secuencias con números naturales	97,1	97,9
Resuelve ejercicios usando operaciones básicas con números naturales	95,4	95,8
Resuelve problemas usando operaciones básicas con números naturales menores de 99 000	94,3	100,0
Compara y ordena números decimales hasta las centésimas	90,9	93,8
Resuelve problemas relacionados con adición y sustracción de fracciones homogéneas	88,8	95,8
Resuelve problemas relacionados con la adición de números decimales	90,4	87,5
Resuelve problemas de compra y venta con unidades de masa	87,8	94,7
Compara y ordena fracciones homogéneas	86,9	97,4
Resuelve problemas de compra y venta con unidades de longitud	84,3	85,4
Resuelve problemas con unidades de tiempo	75,1	79,0
Estima longitudes y peso en unidades convencionales	75,9	70,8
Resuelve ejercicios convirtiendo fracciones a números decimales*	70,2	92,1
Reconoce rectas perpendiculares, paralelas y ángulos; triángulos y cuadriláteros	59,8	70,8
Reconoce cubos, prismas y pirámides	50,2	64,6

\* Diferencia estadísticamente significativa entre los grupos.

Adicionalmente, se han encontrado diferencias (entre 5 y 10 puntos porcentuales) en las capacidades relacionadas con la resolución de problemas que impliquen la adición y sustracción de fracciones homogéneas, de compra y venta con unidades de masa, con operaciones básicas de números naturales menores de 99 000, así como en la capacidad relacionada con la estimación de longitudes y peso en unidades convencionales. Sólo en esta última se observa un mayor porcentaje en las escuelas estatales.

Al analizar las diferencias se encontró que sólo en una capacidad, resolución de ejercicios convirtiendo fracciones a decimales, la diferencia era estadísticamente significativa. Este resultado puede ser explicado por los tamaños dispares de las muestras y porque uno de los grupos tiene pocos casos<sup>31</sup>. Por ello estos resultados, antes que como conclusiones, se toman como indicios que deben ser estudiados con mayor profundidad.

Al analizar sólo los resultados de las escuelas estatales considerando el programa curricular que utilizan los docentes (véase el cuadro 7), se pueden observar diferencias mayores de 10 puntos porcentuales en las siguientes capacidades, en el caso de los profesores que afirman usar la nueva estructura curricular: resolución de ejercicios con operaciones básicas; resolución de problemas con adición y sustracción de fracciones homogéneas; de compra y venta con unidades de masa y con unidades de tiempo; reconocimiento de rectas, ángulos, triángulos y cuadriláteros; y reconocimiento de cubos, pirámides y prismas. Es decir, se observa algún impacto del uso del nuevo programa en el porcentaje de profesores que enseñan estas capacidades.

31 Para el tamaño de las muestras en los distintos análisis, véase el anexo 3.

**Cuadro 7**  
**Porcentaje de profesores que ha enseñado cada capacidad según el programa curricular que utilizan. Resultados de escuelas estatales<sup>32</sup>**

Capacidad	Porcentaje nuevo PC	Porcentaje PC tradicional
Establece relaciones de orden en series y secuencias con números naturales	97,6	90,5
Resuelve ejercicios usando operaciones básicas con números naturales	96,2	85,7
Resuelve problemas usando operaciones básicas con números naturales	94,2	92,3
Compara y ordena números decimales hasta las centésimas	91,4	85,7
Resuelve problemas relacionados con la adición y sustracción de fracciones homogéneas	89,5	76,2
Resuelve problemas relacionados con la adición de números decimales	90,0	95,2
Resuelve problemas de compra y venta con unidades de masa	88,0	76,9
Compara y ordena fracciones homogéneas	86,7	84,6
Resuelve problemas de compra y venta con unidades de longitud	84,3	81,0
Resuelve problemas con unidades de tiempo	75,1	61,5
Estima longitudes y peso en unidades convencionales	76,7	71,4
Resuelve ejercicios convirtiendo fracciones a números decimales*	68,9	76,9
Reconoce rectas perpendiculares, paralelas y ángulos; triángulos y cuadriláteros	61,0	47,6
Reconoce cubos, prismas y pirámides	51,9	33,3

\* Diferencias estadísticamente significativas entre los grupos.

Adicionalmente, se han encontrado diferencias (entre 5 y 10 puntos porcentuales) en las siguientes capacidades, en el caso de los profesores que utilizan la nueva estructura curricular: establecimiento de relaciones de orden en series y secuencias; comparación y ordenamiento de números decimales; y estimación de longitudes y peso en unidades convencionales. Sólo en las capacidades relacionadas con resolución de problemas con números decimales y solución de ejercicios convirtiendo fracciones a decimales se observan porcentajes mayores en el grupo de profesores que usan la estructura curricular tradicional.

Al analizar la significación estadística de la diferencia entre las escuelas estatales que utilizan el nuevo programa, por un lado, y las que emplean el tradicional, por otro, se encontraron resultados similares a los hallados en el análisis de la diferencia entre escuelas estatales y no estatales: la diferencia encontrada era estadísticamente significativa sólo en la capa-

cidad relacionada con la resolución de ejercicios convirtiendo fracciones a decimales.

**Entre los profesores de escuelas estatales que utilizan la nueva estructura curricular y enseñan las capacidades, ¿éstas son desarrolladas a lo largo de todo el año?<sup>33</sup>**

El programa curricular hace referencia a que las capacidades deben ser trabajadas durante

32 Dado que el tamaño de las muestras es pequeño (210 ó 225 que usan nuevo currículo, 21 ó 13 que usan el currículo tradicional).

33 A los profesores que mencionaron haber desarrollado las capacidades con los alumnos se les preguntó si fueron trabajadas a lo largo de todo el año, en momentos específicos o si se estaban trabajando en el momento de la aplicación. Para fines del análisis, las dos últimas alternativas se juntaron.

**Cuadro 8**  
**Porcentaje de profesores que desarrolla las capacidades a lo largo del año.**  
**Resultados de escuelas estatales que usan el nuevo programa curricular**

Capacidad	Porcentaje de preparación a lo largo de todo el año
Establece relaciones de orden en series y secuencias con números naturales	69,8
Resuelve ejercicios usando operaciones básicas con números naturales	73,1
Resuelve problemas usando operaciones básicas con números naturales	75,0
Compara y ordena números decimales hasta las centésimas	36,8
Resuelve problemas relacionados con la adición y sustracción de fracciones homogéneas	21,0
Resuelve problemas relacionados con la adición de números decimales	28,5
Resuelve problemas de compra y venta con unidades de masa	53,8
Compara y ordena fracciones homogéneas	33,0
Resuelve problemas de compra y venta con unidades de longitud	30,3
Resuelve problemas con unidades de tiempo	33,3
Estima longitudes y peso en unidades convencionales	19,4
Resuelve ejercicios convirtiendo fracciones a números decimales	27,6
Reconoce rectas perpendiculares, paralelas y ángulos; triángulos y cuadriláteros	14,2
Reconoce cubos, prismas y pirámides	12,0

todo el grado, a manera de espiral; es decir, a lo largo de todo el año. En el cuadro 8 se puede observar que las capacidades relacionadas con números naturales son trabajadas todo el año por un mayor porcentaje de profesores de escuelas estatales que usan el nuevo programa curricular<sup>34</sup>, mientras que al resto se dedican principalmente momentos específicos del año, excepto la capacidad relacionada con la resolución de problemas con unidades de masa<sup>35</sup>. El porcentaje de profesores que enseñan en espiral las capacidades relacionadas con la iniciación a la Geometría es muy reducido. Estos resultados podrían indicar que muchos profesores todavía no utilizan la enseñanza de las capacidades en espiral.

### **Respecto de las escuelas estatales exclusivamente, ¿por qué no se enseñan algunas capacidades?<sup>36</sup>**

En este acápite se pueden observar claramente las limitaciones del tamaño de las muestras para los análisis. Dado que el porcentaje de no enseñanza es reducido en algunas capacidades y que hay altos porcentajes de no res-

puesta (o de “no sabe” como respuesta)<sup>37</sup>, sólo se discutirán las capacidades con un tamaño de muestra mayor de 30 casos.

En el cuadro 9 se puede observar que la mayoría de profesores refiere no enseñar las capacidades relacionadas con la conversión de fracciones a números decimales, incluyendo el reconocimiento de rectas, triángulos, cuadriláteros, cubos, prismas y pirámides, porque se encuentran en el currículo de un grado posterior.

Estos resultados podrían dar ciertos indicios acerca de la poca familiaridad de los profesores con la nueva estructura curricular.

34 Esta tendencia también se observa en escuelas no estatales, independientemente del currículo que utilicen.

35 Donde son similares los porcentajes de profesores que refieren trabajar a lo largo del año o en momentos específicos.

36 Se les preguntó a los profesores que no habían desarrollado las capacidades si lo hicieron en un grado anterior, si serán desarrolladas en un grado posterior y si no saben si fueron o serán desarrolladas o no responden.

37 Para mayor detalle del tamaño de las muestras de los análisis, véase el anexo 3.

**Cuadro 9**  
**Porcentaje de profesores que no han preparado a sus alumnos en cada capacidad porque consideran que están incluidas en el currículo de un grado posterior. Resultados de escuelas estatales que usan el nuevo programa curricular**

Capacidad	La capacidad está incluida en el currículo de un grado posterior (Porcentaje)
Establece relaciones de orden en series y secuencias con números naturales	50,0
Resuelve ejercicios usando operaciones básicas con números naturales	66,7
Resuelve problemas usando operaciones básicas con números naturales	16,7
Compara y ordena números decimales hasta las centésimas	50,0
Resuelve problemas relacionados con la adición y sustracción de fracciones homogéneas	88,0
Resuelve problemas relacionados con la adición de números decimales	70,6
Resuelve problemas de compra y venta con unidades de masa	42,9
Compara y ordena fracciones homogéneas	52,2
Resuelve problemas de compra y venta con unidades de longitud	61,1
Resuelve problemas con unidades de tiempo	48,8
Estima longitudes y peso en unidades convencionales	51,6
Resuelve ejercicios convirtiendo fracciones a números decimales	63,5
Reconoce rectas perpendiculares, paralelas y ángulos; triángulos y cuadriláteros	61,2
Reconoce cubos, prismas y pirámides	58,7

#### **4 RELACION ENTRE EL CURRÍCULO IMPLEMENTADO Y EL CURRÍCULO APRENDIDO**

Luego de tener un panorama del currículo, en este acápite se analiza la relación entre la enseñanza —o ausencia de ella— de las capacidades y la proporción promedio de acierto<sup>38</sup> del aula en la prueba de Lógico-Matemática. Se presentan los resultados de la muestra total y de las escuelas estatales.

Se espera que exista relación entre la enseñanza de las capacidades y los resultados de la prueba aplicada. Sin embargo, es importante considerar que no es la única variable que influye en el rendimiento, y que es posible que la relación sea más evidente en algunas capacidades que en otras. Algunas áreas de desarrollo se aprenden básicamente en la escuela, mientras que otras se aprenden en la escuela y se refuerzan en la vida diaria. En términos generales, las capacidades relacionadas con el área Lógico-Matemática se desarrollan fundamentalmente en la escuela, por lo que es de presumir una mayor “sensibili-

dad” a la enseñanza. Por el contrario, diversos estudios sugieren que el papel que cumple la familia en el desarrollo de las capacidades relacionadas con la comunicación es mayor que el que cumple en el aprendizaje de la Matemática; aunque, como veremos, los resultados no muestran que tal cosa suceda en todas las capacidades evaluadas.

En este acápite se analiza la relación entre la cobertura total de capacidades<sup>39</sup> y el rendimiento de los alumnos, y la relación entre la enseñanza de cada una y el rendimiento en las preguntas que miden la capacidad.

#### **¿EN QUÉ MEDIDA LA COBERTURA TOTAL DE CAPACIDADES ESTÁ RELACIONADA CON EL RENDIMIENTO DE LOS ALUMNOS?**

Para realizar este análisis se correlacionó el número total (por forma) de capacidades en-

<sup>38</sup> En los ítems correspondientes a las capacidades.

<sup>39</sup> De cada forma.

señadas y el puntaje total de los alumnos en las preguntas de la prueba de Matemática que correspondían a las capacidades de cada forma<sup>40</sup>. Como se puede observar en el cuadro 10, hay una asociación significativa ( $p < 0,05$ ) y positiva entre ambas variables (dígito en negrita). Alrededor de 8% de la diferencia en el rendimiento promedio de la clase es explicado por la enseñanza de las capacidades. Es decir, a medida que aumenta el número de capacidades enseñadas, aumenta también el rendimiento de los alumnos.

Además, se puede observar que la enseñanza de las capacidades también influye en el rendimiento al considerar todos los ítems de la prueba. Dado que las capacidades no son necesariamente independientes, la enseñanza de una afecta el rendimiento de las restantes.

Al observar el cuadro 11 se encuentra que la asociación se mantiene tanto en escuelas estatales como en no estatales. La enseñanza de las capacidades explica alrede-

Cuadro 10 Correlaciones entre cobertura de capacidades y rendimiento. Resultados de la muestra total <sup>41</sup>			
Proporción de enseñanza de capacidades	Rendimiento		
	Todas las preguntas	Preguntas de la forma A	Preguntas de forma la B
De la forma A	0,32	<b>0,29</b>	0,32
De la forma B	0,24	0,17	<b>0,27</b>

dor del 7% de la diferencia en el rendimiento promedio de la clase en las escuelas estatales, y el 15% en las no estatales. Sin embargo, la correlación observada para el caso de escuelas no estatales, específicamente de la forma B, es bastante mayor. Alrededor de 30% de la diferencia es explicado por la enseñanza.

Cuadro 11 Correlaciones entre cobertura de capacidades y rendimiento. Resultados de escuelas estatales y no estatales						
Proporción de enseñanza de capacidades	Rendimiento					
	Estatal			No estatal		
	Todas las preguntas	Preguntas de la forma A	Preguntas de la forma B	Todas las preguntas	Preguntas de la forma A	Preguntas de la forma B
De la forma A	0,29	<b>0,27</b>	0,29	0,34	<b>0,29</b>	0,36
De la forma B	0,20	0,15	<b>0,25</b>	0,55	0,50	<b>0,54</b>

40 Las capacidades evaluadas en la prueba de Lógico-Matemática se presentaron en dos formas (A y B). Si se hubieran presentado todas las capacidades en una sola forma, el instrumento se habría extendido demasiado. Así, la mitad de los docentes contestó una forma y la mitad otra.

41 En negritas figuran las correlaciones estadísticamente significativas y de magnitud suficientemente grande como para ser consideradas "de significación práctica". Siguiendo a Ernesto Cohen ("Indicators for the Assessment of Social Policies", en Franco, Rolando; Pascual Gerstenfeld y Ernesto Cohen: *Social Policies and Socioeconomics Indicators for Transitorial Economies*. Santiago de Chile: CEPAL, 1998, pp. 50-60), hemos tomado como punto de corte el valor 0,11 para declarar que una correlación tiene "significación práctica". Este valor puede parecer bajo, pero está dentro del rango de correlaciones halladas en investigaciones en las que se miden características del individuo.

**Cuadro 12**  
**Proporción de acierto en las preguntas de las pruebas correspondientes a las capacidades según enseñanza o no. Resultados de la muestra total**

Capacidad	Número de pregunta	Proporción de acierto si enseña	Proporción de acierto no enseña
Establece relaciones de orden en series y secuencias con números naturales <sup>42</sup>	1	0,71	0,63
	<b>2*</b>	0,87	0,77
Resuelve ejercicios usando operaciones básicas con números naturales <sup>43</sup>	7	0,92	0,90
	8	0,79	0,73
Resuelve problemas usando operaciones básicas con números naturales <sup>44</sup>	9	0,52	0,40
	17	0,86	0,79
	18	0,48	0,41
	19	0,54	0,37
Compara y ordena números decimales hasta las centésimas <sup>45</sup>	5	0,93	0,90
	12	0,19	0,17
Resuelve problemas relacionados con la adición y sustracción de fracciones homogéneas <sup>46</sup>	13	0,40	0,33
	14	0,50	0,31
Resuelve problemas relacionados con la adición de números decimales	23	0,60	0,57
	<b>24*</b>	0,84	0,77
Resuelve problemas de compra y venta con unidades de masa	21	0,20	0,19
	<b>28*</b>	0,67	0,56
Compara y ordena fracciones homogéneas	<b>4*</b>	0,83	0,74
	<b>11*</b>	0,82	0,70
Resuelve problemas de compra y venta con unidades de longitud	30	0,55	0,51
	31	0,40	0,34
Resuelve problemas con unidades de tiempo	25	0,52	0,50
	26	0,20	0,16
Estima longitudes y peso en unidades convencionales	27	0,53	0,50
	29	0,23	0,21
Resuelve ejercicios convirtiendo fracciones a números decimales	<b>15*</b>	0,34	0,24
	22	0,29	0,28
Reconoce rectas perpendiculares, paralelas y ángulos; triángulos y cuadriláteros	<b>3*</b>	0,35	0,26
	<b>10*</b>	0,75	0,68
	<b>16*</b>	0,51	0,43
	<b>20*</b>	0,49	0,40
Reconoce cubos, prismas y pirámides	<b>6*</b>	0,50	0,39
	32	0,25	0,26

\* Diferencias estadísticamente significativas entre los grupos.

42 El tamaño del grupo que no prepara es demasiado pequeño (8 personas) como para realizar inferencias.

43 El tamaño del grupo que no prepara es demasiado pequeño (13 personas) como para realizar inferencias.

44 El tamaño del grupo que no prepara es demasiado pequeño (14 personas) como para realizar inferencias.

45 El tamaño del grupo que no prepara es demasiado pequeño (25 personas) como para realizar inferencias.

46 El tamaño del grupo que no prepara es demasiado pequeño (29 personas) como para realizar inferencias.

**Cuadro 13**  
**Proporción de acierto en las preguntas de las pruebas correspondientes a las capacidades según enseñanza o no. Resultados de escuelas estatales**

Capacidad	Número de pregunta	Proporción de acierto si enseña	Proporción de acierto no enseña
Establece relaciones de orden en series y secuencias con números naturales <sup>47</sup>	1	0,69	0,62
	2	0,84	0,76
Resuelve ejercicios usando operaciones básicas con números naturales <sup>48</sup>	7	0,91	0,89
	8	0,76	0,71
Resuelve problemas usando operaciones básicas con números naturales <sup>49</sup>	9	0,49	0,40
	17	0,85	0,79
	18	0,46	0,41
	19	0,52	0,37
Compara y ordena números decimales hasta las centésimas <sup>50</sup>	5	0,92	0,90
	12	0,16	0,16
Resuelve problemas relacionados con la adición y sustracción de fracciones homogéneas <sup>51</sup>	13	0,37	0,32
	<b>14*</b>	0,46	0,30
Resuelve problemas relacionados con la adición de números decimales <sup>52</sup>	23	0,58	0,56
	24	0,83	0,75
Resuelve problemas de compra y venta con unidades de masa	21	0,19	0,19
	<b>28*</b>	0,65	0,55
Compara y ordena fracciones homogéneas	4	0,82	0,74
	<b>11*</b>	0,80	0,69
Resuelve problemas de compra y venta con unidades de longitud	30	0,51	0,49
	31	0,36	0,33
Resuelve problemas con unidades de tiempo	25	0,52	0,50
	26	0,19	0,16
Estima longitudes y peso en unidades convencionales	27	0,51	0,48
	29	0,22	0,19
Resuelve ejercicios convirtiendo fracciones a números decimales	<b>15*</b>	0,32	0,23
	22	0,28	0,27
Reconoce rectas perpendiculares, paralelas y ángulos; triángulos y cuadriláteros	<b>3*</b>	0,32	0,25
	10	0,71	0,65
	16	0,46	0,41
	<b>20*</b>	0,46	0,37
Reconoce cubos, prismas y pirámides	6*	0,25	0,26
	32	0,48	0,37

\* Diferencias estadísticamente significativas entre los grupos.

47 El tamaño del grupo que no prepara es demasiado pequeño (7 personas) como para realizar inferencias.

48 El tamaño del grupo que no prepara es demasiado pequeño (11 personas) como para realizar inferencias.

49 El tamaño del grupo que no prepara es demasiado pequeño (14 personas) como para realizar inferencias.

50 El tamaño del grupo que no prepara es demasiado pequeño (22 personas) como para realizar inferencias.

51 El tamaño del grupo que no prepara es demasiado pequeño (27 personas) como para realizar inferencias.

52 El tamaño del grupo que no prepara es demasiado pequeño (24 personas) para realizar inferencias.

## ¿EN QUÉ MEDIDA LA COBERTURA DE CADA CAPACIDAD INFLUYE EN EL RENDIMIENTO DE LOS ALUMNOS EN CADA UNA DE LAS PREGUNTAS QUE FORMA LA CAPACIDAD?

En el cuadro 12 se puede observar que, en todos los casos, las proporciones de acierto son mayores o iguales si los profesores refieren haber enseñado la capacidad evaluada. Sin embargo, se han encontrado diferencias estadísticamente significativas sólo en algunas de las preguntas. En relación con las cinco primeras capacidades, no hay casos suficientes para realizar inferencias concluyentes.

Al analizar sólo los resultados de las escuelas estatales (véase el cuadro 13), se puede observar que en casi todos los casos las proporciones de acierto son mayores o iguales si los profesores refieren haber enseñado la capacidad evaluada<sup>53</sup>. Sin embargo, se han encontrado diferencias estadísticamente significativas sólo en algunas de las preguntas. Respecto de las seis primeras capacidades, no hay casos suficientes para realizar inferencias concluyentes.

## 5 RESUMEN Y CONCLUSIONES

En sus tres dimensiones (intencional, implementado y aprendido), el currículo es parte fundamental del modelo de oportunidades de aprendizaje trabajado por el SMSO y constituye un indicador del proceso educativo que nos permite entender qué está pasando en las aulas del Perú en un momento en que se están poniendo en marcha importantes cambios curriculares.

A partir de 1996 el MED está implementando una serie de reformas educativas relacionadas con el currículo de primaria. Específicamente, la nueva estructura curricular fue difundida en 1998. El proceso educativo en primaria ha sido reorganizado en tres ciclos con dos grados y cinco áreas de desarrollo. Las nuevas estructuras curriculares se organizan sobre la base de competencias y capacidades y se sugieren nuevas estrategias pedagógicas. Los alumnos son considerados el centro del proceso educativo, se pone el énfasis

en el aprendizaje más que en la enseñanza y se asigna a los profesores el papel de facilitadores del proceso educativo.

En relación con el área de Lógico-Matemática, en la estructura curricular de cuarto grado de primaria se propone desarrollar la capacidad de establecer relaciones entre situaciones y objetos partiendo de experiencias concretas. Asimismo, se destaca la importancia de trabajar los aprendizajes a manera de espiral, graduando la incorporación de nuevos temas y la ampliación de la dificultad para que las experiencias educativas sean agradables para los alumnos.

En este contexto, el análisis del currículo del área de Lógico-Matemática de cuarto grado de primaria resulta sumamente útil, ya que contribuye con el mejoramiento de la calidad educativa al retroalimentar a las unidades responsables del desarrollo curricular. Sin embargo, cabe mencionar que ninguna de las relaciones analizadas debe ser considerada de causa-efecto, y que al estudiar la información no se ha controlado el impacto de variables "extrañas". Como ya mencionamos, en esta sección del documento se analiza sólo una de las dimensiones planteadas en el modelo del SMSO. Otras variables que también son importantes para entender los resultados encontrados (por ejemplo, las prácticas pedagógicas) serán desarrolladas en futuros trabajos. A continuación se presentan los principales resultados y conclusiones del estudio.

De acuerdo con los resultados de este trabajo, la mayoría de profesores de Lógico-Matemática de cuarto de primaria de escuelas polidocentes completas de zonas urbanas de la muestra usa la nueva estructura curricular. El porcentaje de profesores de escuelas estatales que usa esta nueva estructura curricular es mayor que aquel de escuelas no estatales.

Dado que el programa curricular de Lógico-Matemática presenta las capacidades que deben ser desarrolladas en clase y que la aplicación se realizó a fines del año escolar, todas las capacidades evaluadas debieron haber sido enseñadas por los profesores. Sin embargo, sorprende encontrar un alto porcen-

53 Sólo en algunas preguntas la proporción promedio de acierto es igual.

taje de profesores que afirman no haber desarrollado todas las capacidades evaluadas, porcentaje aún mayor en escuelas estatales que no estatales. Además, se ha encontrado que este porcentaje es mayor si los profesores de escuelas estatales utilizan el programa tradicional y no el nuevo.

El hecho de no enseñar todas las capacidades podría implicar problemas de falta de equidad en la educación. Si el docente no desarrolla algunas capacidades en sus alumnos, evidentemente limita sus posibilidades de aprendizaje. Por otro lado, el que los profesores no hayan enseñado alguna capacidad no necesariamente es inadecuado, ya que esta decisión podría haberse basado en el conocimiento de los procesos de aprendizaje de sus alumnos.

Al analizar la enseñanza de cada capacidad se ha encontrado que las capacidades relacionadas con la iniciación a la Geometría fueron desarrolladas por el menor porcentaje de profesores de escuelas no estatales. La antecede la capacidad relacionada con la estimación de longitudes y peso. Estas capacidades son innovaciones de la nueva estructura curricular.

Aquellas capacidades que han sido cubiertas por el menor porcentaje de profesores de escuelas estatales de la muestra también están relacionadas con la iniciación a la Geometría, a pesar de que son parte de las innovaciones incluidas en el nuevo programa. En segundo lugar de menor cobertura se tiene la capacidad que permite vincular el uso de fracciones y números decimales, y las capacidades relacionadas con la resolución de problemas con unidades de tiempo y con la estimación de longitudes y peso.

Si se analizan estos resultados considerando el programa curricular que utilizan los profesores (nuevo o tradicional), se encuentra que las capacidades relacionadas con la iniciación a la Geometría son cubiertas por un mayor porcentaje de profesores que utilizan el nuevo programa, mientras que la resolución de ejercicios convirtiendo fracciones a decimales es cubierta por un mayor porcentaje de profesores que utilizan el programa tradicional.

La mayoría de profesores de escuelas estatales que utilizan la nueva estructura curricular preparan a lo largo de todo el año las

capacidades relacionadas con los números naturales. Las restantes han sido trabajadas en momentos específicos del año.

Se ha encontrado una asociación positiva y significativa entre la enseñanza de las capacidades y el rendimiento promedio de los alumnos en la prueba aplicada. Sin embargo, al analizar la relación entre la enseñanza y el rendimiento por pregunta, se han hallado asociaciones positivas y significativas sólo en algunas.

Por último, es importante reconocer que la información presentada en este documento debe ser analizada con mayor profundidad en futuras investigaciones utilizando diversas técnicas tales como la observación de clases y el análisis de textos y de cuadernos de los alumnos. Una de las limitaciones principales de este análisis ha sido el tamaño de la muestra de docentes.

A continuación se presentan algunas recomendaciones que se desprenden de los resultados encontrados en este estudio:

- Se podría considerar la reducción del número de capacidades contempladas en el currículo, dado que aparentemente los docentes no llegan a cubrir todas (hay alternativas a esta sugerencia, como, por ejemplo, mantener el número de capacidades y aumentar el tiempo de enseñanza).
- Algunas capacidades son cubiertas por menos docentes. Esto podría deberse a la falta de tiempo, a la escasez de buenos materiales educativos, a la falta de conocimiento de los docentes sobre el tema tratado o a dificultades de los docentes para tratar estos temas pedagógicamente. Lo primero que habría que hacer es averiguar por qué estas capacidades son enseñadas menos frecuentemente, pero es probable que los docentes se beneficiarían de talleres de metodologías pedagógicas activas para la enseñanza de las capacidades que son tratadas en menor proporción.
- Los docentes parecen tener dificultades para seguir la recomendación del currículo de trabajar los temas en espiral, y, por tanto, seguramente se beneficiarían con talleres para el desarrollo de habilidades pedagógicas en este sentido.

# ANEXO 1

## MUESTRA Y PROCEDIMIENTOS

El presente estudio es descriptivo correlacional, busca acercarse al currículo intencional e implementado de Lógico-Matemática y responder a las siguientes preguntas de investigación: ¿cuán generalizado está el uso del currículo intencional?; ¿cuán extendida está la enseñanza del conjunto de capacidades evaluadas entre los profesores?; ¿cuán extendida está la enseñanza de cada una de las capacidades evaluadas?; entre los profesores de escuelas estatales que utilizan el nuevo programa curricular que enseñan las capacidades, ¿éstas son desarrolladas a lo largo de todo el año?; ¿por qué los profesores de escuelas estatales que utilizan la nueva estructura curricular no enseñan las capacidades evaluadas?; y, ¿qué relación existe entre la enseñanza de las capacidades y el rendimiento de los alumnos?

Para el análisis se ha utilizado información recogida por el Proyecto de Medición de Rendimiento Estudiantil 1998\*, en el que se aplicaron pruebas de Lógico-Matemática a una muestra de 16 827 alumnos de cuarto grado de primaria. También se aplicaron encuestas a 572 profesores de los alumnos evaluados. El proyecto CRECER se desarrolló en CE polidocentes completos de zonas urbanas en el ámbito nacional. Por ello, los resultados son representativos de los docentes y del estudiante promedio de este tipo de escuelas en el país.

La aplicación de los instrumentos se realizó la primera semana de diciembre de 1998, último mes del año escolar. Estuvo a cargo de personal externo a los CE evaluados, especialmente preparado para manejar los instrumentos. Se puso énfasis en que las encuestas no iban a ser revisadas por los directores ni por otros colegas del CE.

La encuesta que recogió la información relacionada con el currículo tuvo dos “formas” en las que se presentaban distintas capacidades. En la forma A se preguntó por cinco capacidades, y en la forma B se indagó por nueve. Es decir, un profesor de la muestra respondió una de las formas.

Para medir la cobertura global por forma y hacer compatibles las comparaciones entre ellas, se obtuvo la proporción de la cobertura (sumando aquellas capacidades que habían sido enseñadas por los profesores y dividiendo entre el número total de capacidades), de manera que se obtuvieron índices que toman valores teóricos entre 0 (cuando ninguna capacidad ha sido declarada cubierta) y 1 (cuando todas han sido cubiertas).

---

\* En este proyecto se evaluó el rendimiento de los alumnos de cuarto grado de educación primaria en el área de Matemática. La aplicación de los instrumentos se realizó a fines de noviembre de 1998 (término del año escolar) pues se esperaba que todas las capacidades evaluadas ya hubiesen sido trabajadas por los profesores.

## ANEXO 2

### PRUEBA DE LÓGICO-MATEMÁTICA Y ENCUESTAS PARA PROFESORES

En este anexo se presentan la prueba de Lógico-Matemática, una tabla de especificaciones que registra las capacidades evaluadas y las preguntas utilizadas para medirlas, además de las preguntas de la encuesta para docentes utilizadas para realizar los análisis.

#### PRUEBA DE MATEMÁTICA CUARTO GRADO DE EDUCACIÓN PRIMARIA

- 1) Coloca el signo mayor que ( $>$ ), menor que ( $<$ ) o igual ( $=$ ) en los espacios en blanco y marca la respuesta correcta.

$867 \square 876$

$768 \square 678$

A) = , <

B) > , >

C) > , <

D) < , >

- 2) Escribe el número que sigue y marca la respuesta correcta.

3456

3466

3476

3486

.....

A) 3496

B) 3486

C) 3456

D) 3446

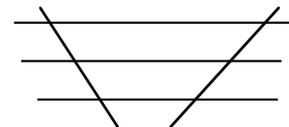
- 3) ¿Cuántas rectas paralelas hay en la siguiente figura?

A) 2

B) 5

C) 4

D) 3



- 4) Sandra, Augusto, Eduardo y Soledad se compraron chocolates para cada uno. Las partes oscuras representan lo que cada uno comió de su chocolate.



Chocolate de Sandra



Chocolate de Augusto



Chocolate de Eduardo



Chocolate de Soledad

¿Quién comió más chocolate?

A) Soledad

B) Augusto

C) Eduardo

D) Sandra

5) María, Julia, Juana y Teresa fueron al mercado y gastaron cada una:

María	S/. 13,75
Teresa	S/. 13,55
Juana	S/. 13,85
Julia	S/. 13,65

¿Quién gastó más?

- A) Julia                      B) Teresa                      C) Juana                      D) María

6) ¿Cuántas caras tiene el siguiente cuerpo geométrico?



- A) 6                      B) 4                      C) 3                      D) 2

7) Realiza la siguiente suma y marca la respuesta correcta.

$$16\ 299 + 21\ 386 =$$

- A) 37 575                      B) 37 585                      C) 37 685                      D) 37 675

8) Resuelve y marca la respuesta correcta.

$$523 + 353 - 658 =$$

- A) 228                      B) 218                      C) 170                      D) 135

9) En una escuela hay 97 estudiantes, de los cuales 49 son **niños**. ¿Cuántas **niñas** hay?

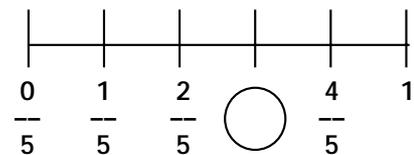
- A) 48                      B) 52                      C) 58                      D) 146

10) La siguiente figura es un:



- A) Rombo                      B) Cuadrado                      C) Rectángulo                      D) Trapecio

11) Marca la fracción que falta en la recta numérica.



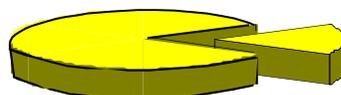
- A) 3/5                      B) 5/3                      C) 3/1                      D) 5/1

12) Pepe tiene 2,40 metros de sogas, Alicia tiene 2,4 metros y María tiene dos metros con 40 centímetros. ¿Cuál de las siguientes respuestas es cierta?

- A) Pepe tiene la sogas más larga                      B) Alicia tiene la sogas más larga  
C) María tiene la sogas más larga                      D) Todas las sogas tienen la misma longitud

13) Lily compró un queso. Su hijo Hernán comió  $\frac{1}{6}$  de queso. ¿Qué parte de queso queda?

- A)  $\frac{1}{6}$  de queso                      B)  $\frac{5}{6}$  de queso  
C)  $\frac{6}{5}$  de queso                      D)  $\frac{6}{1}$  de queso



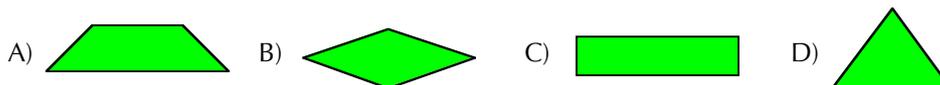
14) Tito estudió Historia durante  $\frac{3}{4}$  de hora, Geografía durante  $\frac{2}{4}$  de hora y dibujó una figura en  $\frac{2}{4}$  de hora. ¿Cuánto tiempo demoró en total?

- A)  $\frac{5}{8}$  de hora      B)  $\frac{5}{4}$  de hora      C)  $\frac{7}{4}$  de hora      D)  $\frac{7}{12}$  de hora

15) Escribe la fracción  $\frac{23}{100}$  en forma decimal.

- A) 0,023      B) 0,23      C) 2,3      D) 23

16) ¿Cuál de las siguientes figuras **no** es un cuadrilátero?



17) En el desfile de un circo había 5 elefantes, 8 leones y 17 caballos. ¿Cuántos animales había en total?

- A) 30      B) 22      C) 20      D) 13

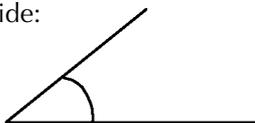
18) Rosa vende diariamente 112 kilogramos de papas. ¿Cuántos kilogramos venderá en 24 días?

- A) 136      B) 672      C) 2688      D) 4704

19) Se divide S/. 4760 entre cinco personas en partes iguales. ¿Cuánto le corresponde a cada una?

- A) S/. 852      B) S/. 952      C) S/. 960      D) S/. 972

20) Este ángulo mide:



- A) Menos de  $90^\circ$       B) Más de  $90^\circ$       C) Más de  $180^\circ$       D)  $90^\circ$

21) Si por la compra de 4 kilogramos de papas se paga S/. 12, ¿cuánto se pagará por la compra de 18 kilogramos?

- A) S/. 34      B) S/. 44      C) S/. 48      D) S/. 54

22) En las siguientes alternativas, marca la igualdad correcta.

- A)  $\frac{321}{100} = 32,1$       B)  $\frac{547}{100} = 5,47$       C)  $\frac{62}{10} = 6,02$       D)  $\frac{83}{100} = 0,083$

23) Inés mide 1,37 metros y Carlos 0,19 metros más que Inés. ¿Cuánto mide Carlos?

- A) 1,18 metros      B) 1,46 metros      C) 1,56 metros      D) 1,57 metros

24) Manuel pagó por un cuaderno S/. 1,40 y por un borrador, 20 céntimos. ¿Cuánto gastó en total?

- A) S/. 21,40      B) S/. 3,40      C) S/. 2,80      D) S/. 1,60

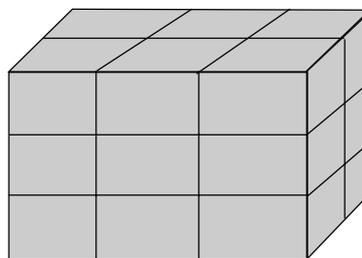
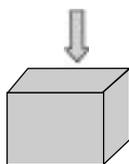
25) Si una hora tiene 60 minutos, ¿cuántas horas y minutos hay en 119 minutos?

- A) 2 horas 1 min      B) 1 hora 59 min      C) 1 hora 49 min      D) 1 hora 19 min

- 26) Roxana se demoró en ayudar a su mamá 35 minutos, en jugar 40 minutos y en hacer su tarea de Matemática 45 minutos. ¿Cuánto tiempo empleó en total?
- A) 2 horas      B) 1 hora 50 minutos      C) 1 hora 20 minutos      D) 1 hora 15 minutos
- 27) ¿Cuánto mide aproximadamente un lápiz nuevo?
- A) 50 centímetros      B) 40 centímetros      C) 20 centímetros      D) 5 centímetros
- 28) Victoria compró en el mercado 3 kilogramos de arroz por S/. 4,80 y 1 kilogramo de frejoles por S/. 3,50. ¿Cuánto gastó Victoria?
- A) S/. 9,10      B) S/. 8,30      C) S/. 7,40      D) S/. 7,30
- 29) ¿Cuánto pesará aproximadamente un huevo de gallina?
- A) 1000 gramos      B) 500 gramos  
C) 70 gramos      D) 5 gramos
- 30) Para cercar un terreno se necesitan 44 metros de alambre. Si el metro de alambre cuesta S/. 2, ¿cuánto se pagará por los 44 metros?
- A) S/. 44      B) S/. 88      C) S/. 90      D) S/. 176
- 31) Por la compra de 10 metros de tela se pagó S/. 40. ¿Cuánto se pagará por la compra de 5 metros?
- A) S/. 10      B) S/. 15      C) S/. 20      D) S/. 55
- 32) ¿Cuántos cubos como el que señala la flecha hay en la figura grande?

- A) 9  
C) 15

- B) 12  
D) 18



### **OPORTUNIDADES DE APRENDIZAJE PARA PROFESORES DE LÓGICO-MATEMÁTICA CUARTO GRADO DE EDUCACIÓN PRIMARIA**

10. ¿Qué currículo está desarrollando este año en la clase de Matemática?
- a) Programa curricular de Escuelas Unitarias.  
b) Programa curricular de Educación Bilingüe Intercultural.  
c) Programa curricular de Educación Primaria (programa tradicional).  
d) Programa curricular Básico de Segundo Ciclo de Educación Primaria (programa nuevo).  
e) Su experiencia le permite desarrollar sus clases sin usar programa curricular.  
f) Otro.

Para cada una de las capacidades evaluadas se mostró a los profesores la definición de éstas y las preguntas de la prueba correspondientes a cada una de ellas. A continuación se presenta un ejemplo y las preguntas analizadas.

<b>I. Habilidad operativa y cálculo</b>
1.2 Resuelve situaciones problemáticas usando operaciones de adición, sustracción, multiplicación y división de números naturales cuyos términos no excedan de 99 000.
<ul style="list-style-type: none"> <li>Hace 18 años, Roberto era tres veces más viejo que su hijo, pero ahora le dobla la edad. ¿Cuántos años tiene Roberto?</li> </ul>

43. Este año, ¿ha preparado usted a sus alumnos para resolver problemas o ejercicios como los anteriores?

Sí No

44. Si ha marcado Sí

a) Trabajó el tema a lo largo de todo el año.

Sí No

b) Trabajó el tema en algún momento del año.

Sí No

c) Está trabajando el tema AHORA (entre el 23 de noviembre y el 11 de diciembre).

Sí No

45. Si ha marcado No

a) El tema se trató en el currículo de un grado ANTERIOR.

Sí No

b) El tema se tratará en el currículo de un grado POSTERIOR.

Sí No

c) No sabe si este tema se ha tratado o se va a tratar en otro grado.

Sí No

# ANEXO 3

## TAMAÑO DE LAS MUESTRAS, N° MERO DE DOCENTES POR GRUPOS

Capacidades	Muestra total		Escuelas estatales			Escuelas estatales sólo nuevo PC					Escuelas no estatales, cualquier PC	
	ND* total	ND que prepara a sus alumnos	ND total	ND que prepara a sus alumnos	ND que no prepara a sus alumnos	ND total	ND que prepara a lo largo del año	ND no respuesta	ND que no prepara porque el currículo es posterior	ND no respuesta	ND total	ND que prepara a sus alumnos
Establece relaciones de orden en series y secuencias con números naturales	288/9	280	240	233	7	205	143	0	1	3	48	47
Resuelve ejercicios usando operaciones básicas con números naturales	288/9	275	240	229	11	202	147	1	4	2	48	46
Resuelve problemas usando operaciones básicas con números naturales	283	269	245	231	14	212	159	0	1	7	38	38
Compara y ordena números decimales hasta las centésimas	288/9	263	240	218	22	192	70	2	6	6	48	45
Resuelve problemas con la adición y sustracción de fracciones homogéneas	288/9	259	240	213	27	188	39	2	16	4	48	46
Resuelve problemas relacionados con la adición de números decimales	288/9	258	240	216	24	189	53	3	12	4	48	42
Resuelve problemas de compra y venta con unidades de masa	283	251	245	215	30	198	106	1	9	6	38	36
Compara y ordena fracciones homogéneas	283	250	245	213	32	195	63	4	12	7	38	37
Resuelve problemas de compra y venta con unidades de longitud	288/9	243	240	202	38	177	53	2	11	15	48	41
Resuelve problemas con unidades de tiempo	283	214	245	184	61	169	56	1	21	13	38	30
Estima longitudes y peso en unidades convencionales	288/9	217	240	183	57	161	31	1	16	18	48	34
Resuelve ejercicios convirtiendo fracciones a números decimales	283	207	245	172	73	155	42	3	33	18	38	35
Reconoce rectas perpendiculares, paralelas y ángulos; triángulos y cuadriláteros	288/9	178	240	144	96	128	18	1	30	33	48	34
Reconoce cubos, prismas y pirámides	288/9	152	240	121	119	109	13	1	37	38	48	31

\* Número de docentes.

# ELABORACIÓN Y APLICACIÓN DEL PROYECTO DE DESARROLLO INSTITUCIONAL EN LOS CENTROS EDUCATIVOS DE PRIMARIA Y SECUNDARIA

CLAUDIA GALINDO  
FRANCISCO INDACOCHEA

Esta sección<sup>1</sup> del documento de trabajo pretende brindar información relacionada con el Proyecto de Desarrollo Institucional (PDI) recogida en CRECER 1998<sup>2</sup>. Como en todo proceso, puede tomar algún tiempo llegar a un nivel en el que la elaboración y la aplicación del PDI se desarrollen en forma eficiente, creando condiciones para la descentralización, la correcta planificación y el mejoramiento de la gestión del centro educativo (CE). Por ello es importante que las oficinas del MED y otras instituciones que trabajan este tema obtengan algunos alcances relacionados con su formulación y aplicación.

Específicamente, se investiga si los CE han elaborado y aplicado el PDI, se especifica quiénes han participado en este proceso, y si han recibido apoyo de las instancias administrativas superiores (Direcciones Regionales, Subregionales, de Lima, de Callao, USE y ADE)<sup>3</sup>. Adicionalmente, interesa saber qué diferencias hay en las variables anteriormente mencionadas al considerar el tipo de gestión del CE y, en escuelas públicas, la participación del director en las capacitaciones de PLANCGED<sup>4</sup>. Para ello se estudia la información recogida en las encuestas aplicadas a los directores, profesores y padres de familia de CE urbanos polidocentes completos a nivel nacional, y de primaria y secundaria<sup>5</sup>. Estas encuestas fueron elaboradas por la UMC en coordinación con la unidad de PLANCGED, cuyo personal participó en la definición de los temas por investigar y en la revisión de las preguntas formuladas.

En el primer acápite del documento se presenta una aproximación al PDI como herramienta de descentralización y de gestión, y al PLANCGED como estrategia de capacitación desarrollada por el MED. En los acápiques segundo, tercero y cuarto se muestran los resultados encontrados. Por último, en el quinto se señalan los hallazgos más importantes del análisis.

## **1 EL PDI COMO HERRAMIENTA DE DESCENTRALIZACIÓN Y GESTIÓN**

La poca autonomía de los CE fue uno de los problemas identificados en el Plan Anual de

- 1 Resultado de las investigaciones llevadas a cabo por la UMC y GRADE. Los autores desean expresar su agradecimiento a José Rodríguez y Santiago Cuento por sus comentarios y sugerencias.
- 2 CRECER es el proyecto de evaluación del rendimiento escolar a cargo de la UMC que aplica periódicamente pruebas en algunas áreas de desarrollo o cursos (*i.e.*, Comunicación Integral, Lógico-Matemática) y encuestas (para directores, profesores, padres de familia y alumnos) para retroalimentar los programas y proyectos desarrollados por el MED, y contribuir con la formulación de las políticas educativas.
- 3 Véase el anexo 1, donde se presentan las preguntas analizadas y alternativas de respuestas.
- 4 Plan Nacional de Capacitación en Gestión Educativa.
- 5 Para mayor detalle de la muestra y procedimientos, véase el anexo 2.

1999 del Ministerio de Educación<sup>6</sup>. En las últimas décadas el MED ha buscado potenciar a los CE como agentes autónomos, fomentando la iniciativa y el compromiso de sus integrantes con el trabajo cotidiano. Para ello se está revalorando, entre otros aspectos, el papel del director como agente planificador y organizador de la gestión, y se está iniciando el traslado de una serie de responsabilidades al CE.

El director, como cabeza que orienta la labor pedagógica del CE y como motor de la gestión institucional, debe coordinar múltiples funciones: i) evaluar los criterios de adaptación del currículo; ii) organizar el trabajo y los lineamientos pedagógicos de los profesores; iii) impulsar proyectos de innovación y estrategias para optimar la gestión pedagógica; iv) planificar; v) definir la organización en términos de su estructura y características; vi) administrar el personal, los recursos y financiamiento, entre otras<sup>7</sup>.

Dado que las responsabilidades del director han sido redefinidas y que algunos no habían sido formados para ejecutar una labor compleja como es la planificación y organización de los CE, el MED inició un plan de capacitación en gestión educativa (PLANCGED) que tiene como objetivo principal brindar herramientas, soporte, capa-

citación y supervisión para que los directores mejoren su gestión institucional y pedagógica. Se viene desarrollando desde 1996<sup>8</sup> y hasta la fecha ha capacitado a 8859 directores de primaria y secundaria de CE urbanos<sup>9</sup>. Adicionalmente, en 1999 se inició la capacitación para directores de CE rurales<sup>10</sup> utilizando la educación a distancia como modalidad de enseñanza.

Una de las principales herramientas para la gerencia y planificación del trabajo de los CE es el PDI, ampliamente tratado en las capacitaciones de PLANCGED. El PDI, como esfuerzo colectivo, debe orientar, organizar y articular el trabajo cotidiano, así como los programas y proyectos de innovación que se ejecuten en el CE. Es un instrumento que permite analizar la situación presente de los CE, definir las metas y objetivos de manera realista, esbozar los planes de acción y clarificar las necesidades de recursos.

En el PDI se deben desarrollar principalmente cuatro aspectos: i) la identidad del CE, es decir, la filosofía y las aspiraciones e ideales que van a guiar el trabajo; ii) el diagnóstico, para identificar las potencialidades y debilidades como punto inicial para definir metas, objetivos, estrategias y planes de acción; iii) la propuesta pedagógica, donde se definen los criterios de acción pedagógica y la propuesta curricular; y, iv) la propuesta de gestión, donde se definen las características de la organización. Para desarrollar estos aspectos se postula incentivar la participación de los profesores y de los padres de familia y trabajar con ellos en forma conjunta.

En este contexto, la UMC decidió incluir en las encuestas de CRECER 1998 preguntas que permitiesen evaluar en alguna medida la elaboración y aplicación del PDI en los CE. A continuación se presentan los resultados.

## **2 PORCENTAJE DE CENTROS EDUCATIVOS QUE HA ELABORADO Y APLICADO EL PDI**

Los procesos de elaboración y aplicación del PDI son complementarios e igualmente im-

6 MED: *Plan Operativo Anual*, 1999, p. 8: "... en las décadas pasadas el Ministerio de Educación se construyó como una organización piramidal y extremadamente centralizada. En consecuencia, las otras instancias de la organización no tuvieron iniciativa ni capacidad para tomar decisiones y desarrollar soluciones para los problemas que surgen día a día".

7 MED-Oficina de Apoyo a la Administración de la Educación y Dirección Nacional de Formación y Capacitación Docentes: *Manual del director*. Lima: MED, 1996.

8 MED: *Plan Nacional de Capacitación para Directores de Educación Primaria y Secundaria*. Lima: MED, 1996.

9 Aproximadamente 62% del total de directores de CE de primaria y secundaria de zonas urbanas. MED-PLANCGED: *Informe de trabajo de PLANCGED*. Lima: MED, 2000.

10 Se ha capacitado aproximadamente a 1440 directores en zonas rurales de un total de 26 500. MED-PLANCGED: *Informe de trabajo de PLANCGED*, op. cit.

**Cuadro 1**  
**Elaboración y aplicación del PDI. Resultados de la muestra total de CE**  
**de primaria y secundaria**  
**(Porcentajes)**

Categorías	Primaria	Secundaria
1. Se ha elaborado y aplicado totalmente	11,6	11,0
2. Se ha elaborado (o se está elaborando) y aplicado parcialmente	46,3	49,2
3. Se ha elaborado (o se está elaborando) y no aplicado	29,2	29,8
4. No elaborado y no aplicado	12,9	9,9
<b>Total</b>	<b>100,0</b>	<b>100,0</b>
<b>Número de casos</b>	<b>568</b>	<b>553</b>

Fuente: Directores de CE.

portantes. Sin embargo, es en la aplicación donde se pueden observar los mayores beneficios para el funcionamiento del CE. No es suficiente formular el PDI: éste debe traducirse en acciones concretas.

Como se puede observar en el cuadro 1, los resultados de los CE de primaria y secundaria con respecto a este tema arrojan porcentajes similares. En la mayoría de CE la elaboración o aplicación del PDI está en proceso (como se percibe al sumar las categorías 2 y 3). Es mayor la cantidad de CE que no lo han aplicado que la de aquellos que no lo han elaborado. Estos resultados son coherentes si se considera que, en primer lugar, no es posible comenzar a aplicar el PDI sin haber iniciado previamente su elaboración y, en segundo lugar, que poner en práctica formulaciones abstractas demanda más tiempo, entrenamiento y despliegue de habilidades diversas.

Cuando se analiza la información tomando en cuenta el tipo de gestión de los CE, como se puede observar en el gráfico 1, es mayor el porcentaje de CE no estatales que estatales que han elaborado y aplicado totalmente el PDI. Estos resultados podrían indicar que los CE estatales enfrentan mayores dificultades, requieren mayor tiempo para aplicar totalmente esta herramienta o no tienen los insumos necesarios para hacerlo. Adicionalmente, no se observan diferencias importantes en la distribución de los porcenta-

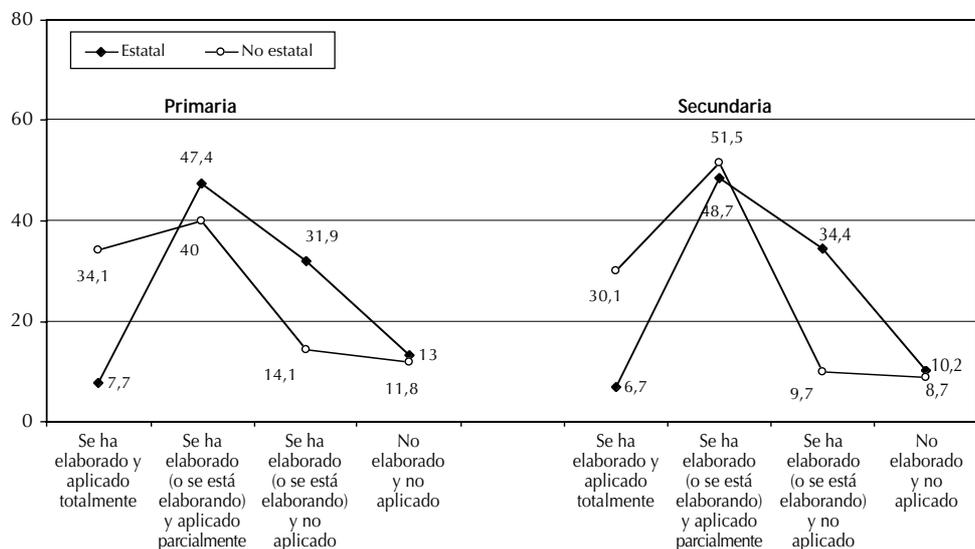
jes entre CE estatales de nivel primario y secundario.

Además de las diferencias observadas en la elaboración y aplicación del PDI según tipo de gestión de los CE, se consideró importante analizar las diferencias en los CE estatales según la asistencia del director a PLANCGED. Esta capacitación tiene como uno de sus objetivos principales brindar las herramientas necesarias tanto para la formulación como para la aplicación del PDI.

Del total de directores que participaron en la evaluación CRECER 1998, 61,4% en primaria y 52,8% en secundaria han participado en alguna de las capacitaciones de PLANCGED. Estos porcentajes suben a 67,1% y a 57,6% respectivamente cuando se analizan solamente los CE del sector estatal de la muestra. Pese a que la capacitación está dirigida principalmente a directores de CE estatales, se ha encontrado que 28% de los directores de CE no estatales de primaria y 31,7% de secundaria han participado en alguna de ellas.

En el gráfico 2 se puede observar que la asistencia de los directores a PLANCGED, en términos generales, está relacionada con la elaboración y aplicación del PDI. Sin embargo, no se aprecian efectos en el grupo que ha elaborado y aplicado totalmente el plan. Al parecer, PLANCGED, como estrategia de capacitación, está brindando herramientas más eficientes para la elaboración que para la aplicación del PDI.

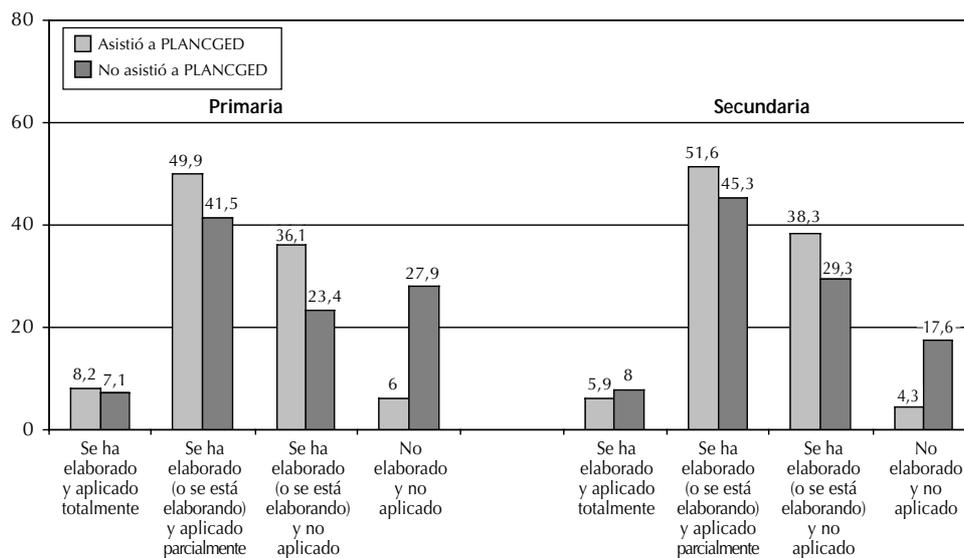
**Gráfico 1**  
**Elaboración y aplicación del PDI según tipo de gestión del CE.**  
**Resultados de CE de primaria y secundaria (Porcentajes)**



*Nota:* Número de casos: CE estatales de primaria: 483; CE no estatales de primaria: 85; CE estatales de secundaria: 450; y CE no estatales de secundaria: 103.

*Fuente:* Directores de CE.

**Gráfico 2**  
**Elaboración y aplicación del PDI según participación en PLANCGED.**  
**Resultados de CE estatales de primaria y secundaria (Porcentajes)**



*Nota:* Número de casos: CE de primaria, asistió a PLANCGED: 319; CE de primaria, no asistió a PLANCGED: 154; CE de secundaria, asistió a PLANCGED: 256; y CE de secundaria, no asistió a PLANCGED: 188.

*Fuente:* Directores de CE.

**Cuadro 2**  
**Participación de profesores y/o padres en la elaboración del PDI. Resultados**  
**de la muestra total de CE de primaria y secundaria**  
**(Porcentajes)**

Categorías	Primaria	Secundaria
Se ha reunido con profesores y padres	45,3	34,3
Se ha reunido con profesores y no con padres	23,7	30,7
Se ha reunido con padres y no con profesores	6,7	7,3
No se ha reunido con ninguno	24,3	27,6
<b>Total</b>	100,0	100,0
<b>Número de casos</b>	481	492

Fuente: Directores de CE.

### **3 PERSONAS QUE HAN PARTICIPADO EN LA ELABORACIÓN DEL PDI**

Por tratarse de una herramienta que busca fomentar un mayor compromiso de los agentes directamente involucrados con el funcionamiento del CE, el PDI debe elaborarse de manera conjunta, esto es, con la participación de profesores y padres de familia. El análisis de este acápite refiere sólo a aquellos CE donde el director manifestó que el PDI estuvo elaborado parcial o completamente<sup>11</sup>.

En primer lugar, se preguntó a los directores si se habían reunido con profesores y padres de familia para elaborar el PDI. En segundo lugar, se preguntó a los profesores en qué aspectos de la elaboración habían participado (definición de la identidad —visión y misión—, diagnóstico del centro, propuesta de gestión, propuesta pedagógica). Finalmente, se preguntó a los padres de familia con qué frecuencia habían participado en la elaboración del PDI. A continuación se presentan los resultados.

#### **PARTICIPACIÓN DE PROFESORES Y PADRES DE FAMILIA REPORTADA POR LOS DIRECTORES**

Como se puede observar en el cuadro 2, se ha encontrado que los porcentajes de directores que reportan haberse reunido con profesores y

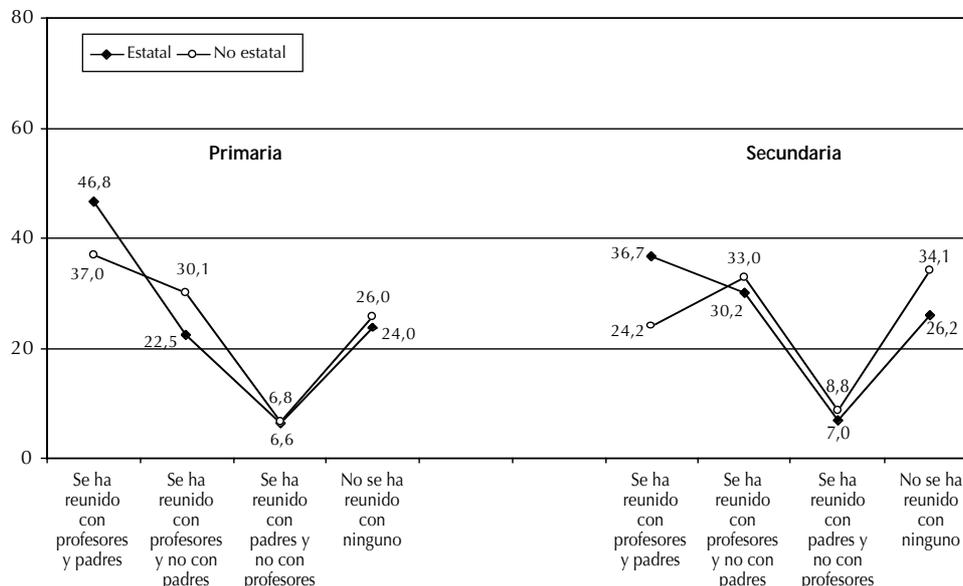
con padres de familia para la elaboración del PDI son los más altos en primaria y en secundaria. Esto es más evidente en primaria que en secundaria. Adicionalmente, alrededor de un cuarto de la muestra, en ambos niveles, se ha elaborado sin la participación de ninguno de dichos agentes educativos.

De esta manera se encuentra que, si bien el porcentaje de CE que ha logrado implicar tanto a padres como a profesores en el proceso de elaboración de su PDI es alto, existe un porcentaje importante de CE donde se ha elaborado el PDI sin la participación de alguno de dichos agentes educativos, e incluso sin la participación de ninguno de ellos (un cuarto de la muestra). Estos resultados nos alertan sobre la posibilidad de que en estos últimos CE surjan mayores dificultades en la aplicación del PDI.

Cuando se analiza la información tomando en cuenta el tipo de gestión de los CE, como se puede observar en el gráfico 3, en primaria y secundaria el porcentaje de directores de CE estatales que refieren haberse reunido con profesores y padres de familia es mayor que el de CE no estatales. En ambos niveles se aprecia una mayor participación de padres de CE estatales que de no estatales. No se encuentra la misma tendencia en la participación de profe-

<sup>11</sup> En primaria, 86,9% de los directores reportan haber elaborado el PDI total o parcialmente; en secundaria, 90%.

**Gráfico 3**  
**Participación de profesores y padres de familia en la elaboración del PDI según tipo de gestión del CE. Resultados de CE de primaria y secundaria (Porcentajes)**



*Nota:* Número de casos: CE estatales de primaria: 408; CE no estatales de primaria: 73; CE estatales de secundaria: 401; y no estatales de secundaria: 91.

*Fuente:* Directores de CE.

sos, que sólo es mayor en CE estatales de primaria.

Además de las diferencias observadas en la participación de profesores y padres de familia según tipo de gestión del CE, se consideró importante analizar las diferencias en los CE estatales según asistencia del director a PLANCGED. En el gráfico 4 se puede observar que la asistencia está relacionada con la participación de profesores y padres, en ambos niveles. Tanto en primaria como en secundaria es mayor el porcentaje de directores que asistió alguna vez a PLANCGED que refiere haberse reunido con profesores y padres. Al mismo tiempo, es menor el porcentaje de directores que asistió a PLANCGED alguna vez que refiere no haberse reunido con ninguno.

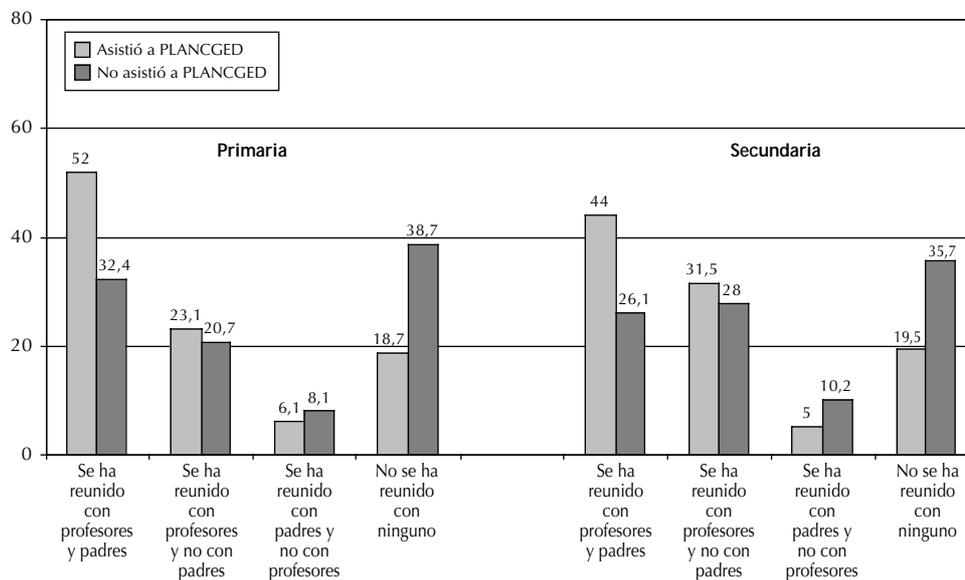
### **PARTICIPACIÓN DE LOS PROFESORES EN LAS DISTINTAS ACTIVIDADES DEL PDI**

Se preguntó a los profesores en qué aspectos de la elaboración del PDI habían participa-

do: definición de la identidad —visión y misión—, diagnóstico del CE, propuesta de gestión y/o propuesta pedagógica. En primaria, 79,9% de los profesores había participado por lo menos en una actividad; en secundaria, 76,8%. Aproximadamente un quinto de la muestra de profesores, en ambos niveles, refirió no haber participado en alguna de estas actividades.

Como se puede observar en el cuadro 3, se ha encontrado un mayor porcentaje de profesores que ha participado en la elaboración del diagnóstico. Porcentajes similares de maestros han colaborado en la propuesta pedagógica y en la definición de la visión y misión del CE. Por último, un menor porcentaje ha cooperado en la propuesta de gestión. Estos resultados eran previsibles, ya que la formulación del diagnóstico requiere la participación de un significativo número de profesores. Con esta finalidad se suelen formar múltiples comisiones de trabajo para analizar con profundidad los diferentes componentes de la situación actual del CE. Por otra parte, la par-

**Gráfico 4**  
**Participación de profesores o padres en la elaboración del PDI según participación del director en PLANCGED. Resultados de CE estatales de primaria y secundaria (Porcentajes)**



*Nota:* Número de casos: CE de primaria, asistió a PLANCGED: 294; CE de primaria, no asistió a PLANCGED: 111; CE de secundaria, asistió a PLANCGED: 241; y CE de secundaria, no asistió a PLANCGED: 157.  
*Fuente:* Directores de CE.

**Cuadro 3**  
**Participación de profesores en actividades del PDI. Resultados de la muestra total de CE de primaria y secundaria (Porcentajes)**

Categorías	Primaria	Secundaria
Elaborar el diagnóstico del CE	65,6	58,9
Formular la propuesta pedagógica	54,3	47,4
Definir la visión y la misión	53,0	46,4
Formular la propuesta de gestión	36,0	29,9

*Fuente:* Directores de CE.

*Nota:* Como las actividades no son excluyentes, los porcentajes suman más de 100%. El número de casos de cada categoría se puede observar en el anexo 3.

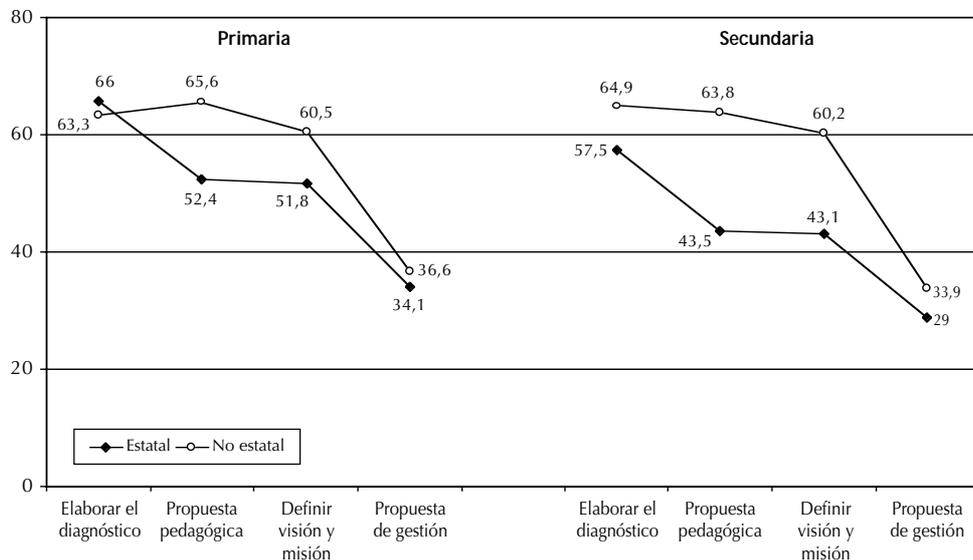
Participación de los profesores en la propuesta de gestión no es tan interesante, ya que para esto cuentan con menos herramientas.

Cuando se analiza la información tomando en consideración el tipo de gestión de los CE, como se puede observar en el gráfico 5, las tendencias encontradas en el cuadro anterior se mantienen. Sin embargo, se puede ob-

servar un mayor porcentaje de participación de profesores en los CE estatales de primaria frente a los de secundaria.

Además, en ambos niveles se aprecia que el porcentaje de participación de los profesores de CE no estatales es mayor que el de los estatales en todas las actividades, salvo en la elaboración del diagnóstico en primaria. En

**Gráfico 5**  
**Participación de profesores en actividades del PDI según tipo de gestión del CE.**  
**Resultados de CE de primaria y secundaria (Porcentajes)**



*Nota:* Como las actividades no son excluyentes, los porcentajes no suman 100%. El número de casos de cada categoría se puede observar en el anexo 3.  
*Fuente:* Profesores de los alumnos evaluados.

secundaria, la diferencia según tipo de gestión es más relevante que en primaria.

En el gráfico 6 se puede apreciar que, en primaria, la asistencia del director a PLANCGED está relacionada con una mayor participación de los profesores en las actividades del PDI. No sucede lo mismo en

secundaria, ya que no se observa una variación importante en los porcentajes de participación de los profesores (excepto en la propuesta pedagógica) según la asistencia del director a PLANCGED.

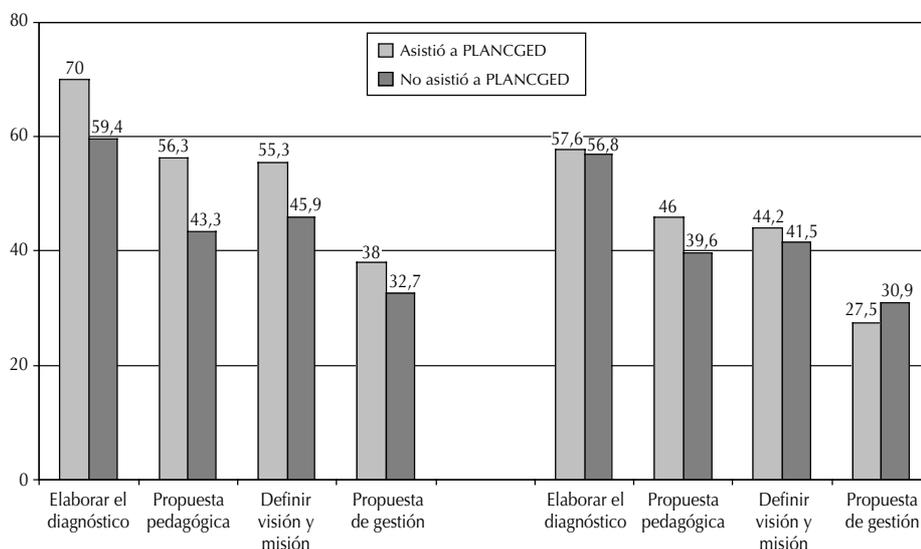
### FRECUENCIA DE PARTICIPACIÓN DE LOS PADRES DE FAMILIA EN EL PDI

Se preguntó a los padres de familia con qué frecuencia habían asistido al CE para colaborar con la elaboración del PDI<sup>12</sup>. En el cuadro 4 se puede observar que la mayoría de padres de familia de primaria y secundaria de los CE que han elaborado el PDI no ha participado de él. Este porcentaje es significativamente mayor en secundaria.

Cuando se analiza la información tomando en cuenta el tipo de gestión de los CE, como se puede observar en el gráfico 7, se encuentra que en primaria la participación de los padres de familia de CE estatales es más frecuente que la de los padres de CE no

12 Existe una diferencia aproximada de 10 puntos porcentuales entre la participación de los padres reportada por los directores y aquella proporcionada por los mismos padres de familia (véase la diferencia entre los cuadros 4 y 2). Esta diferencia puede ser explicada por varias razones: i) la pregunta formulada a los directores hacía referencia a reuniones con padres del nivel sin especificar el grado escolar de sus hijos; ii) los directores pudieron haber atribuido un significado distinto al término "participación"; algunos pudieron haber asumido que si los habían convocado, aunque no asistieron, sí habían "participado"; iii) se presentaron diferencias en los significados atribuidos a "participación" por padres y directores; y, iv) algunos directores podrían haber contestado sobre la base de la respuesta esperada.

**Gráfico 6**  
**Participación de profesores en actividades del PDI según participación del director en PLANCGED. Resultados de CE estatales de primaria y secundaria (Porcentajes)**



*Nota:* Como las actividades no son excluyentes, los porcentajes suman más de 100%. El número de casos de cada categoría se puede observar en el anexo 3.

*Fuente:* Profesores de los alumnos evaluados.

**Cuadro 4**  
**Frecuencia de participación de padres de familia en la elaboración del PDI. Resultados de la muestra total de CE de primaria y secundaria (Porcentajes)**

Categorías	Primaria	Secundaria
Nunca	56,7	70,5
Pocas veces	21,0	19,3
Muchas veces o siempre	22,3	10,2
<b>Total</b>	<b>100,0</b>	<b>100,0</b>
<b>Número de casos</b>	<b>19 368</b>	<b>20 004</b>

*Fuente:* Padres de los alumnos evaluados.

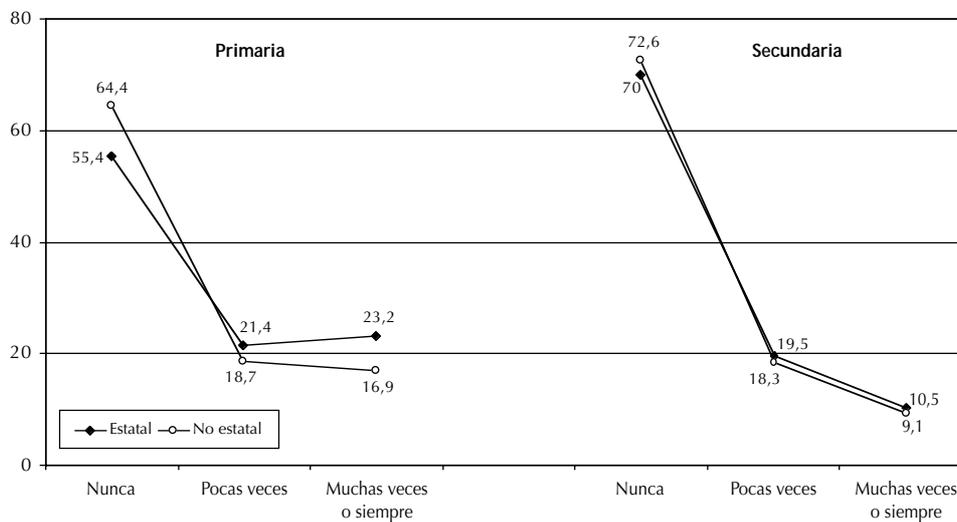
estatales. En cambio, en secundaria, las tendencias de CE estatales y no estatales son similares.

Por último, como se puede observar en el gráfico 8, no existen variaciones en la frecuencia de participación de padres de familia se-

gún la asistencia del director al PLANCGED. Estos resultados son similares a los presentados en el cuadro 4, donde se observa que la mayoría de padres de familia de primaria y secundaria no ha participado en la elaboración del PDI.

**Gráfico 7**

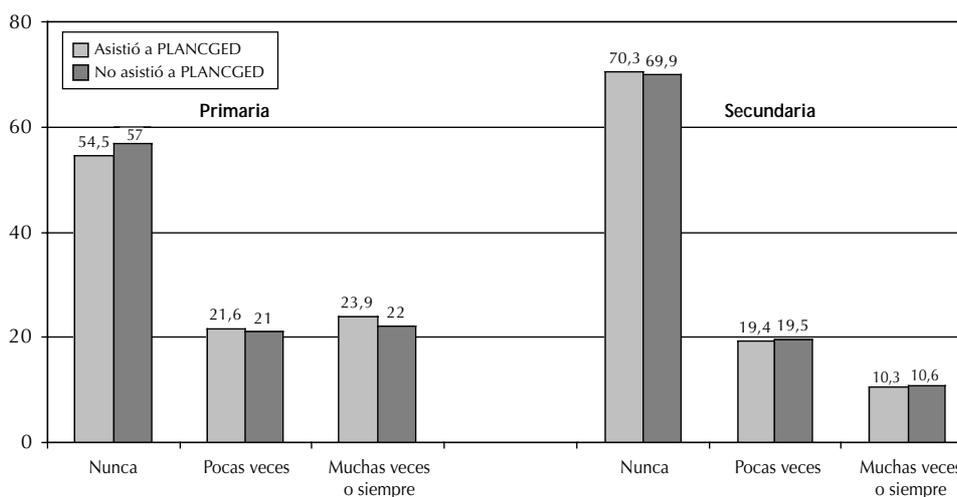
**Frecuencia de participación de padres de familia en la elaboración del PDI según tipo de gestión del CE. Resultados de la muestra total de CE de primaria y secundaria (Porcentajes)**



*Nota:* Número de casos: CE estatales de primaria: 16 500; CE no estatales de primaria: 2868; CE estatales de secundaria: 16 296; y CE no estatales de secundaria: 3708.  
*Fuente:* Profesores de los alumnos evaluados.

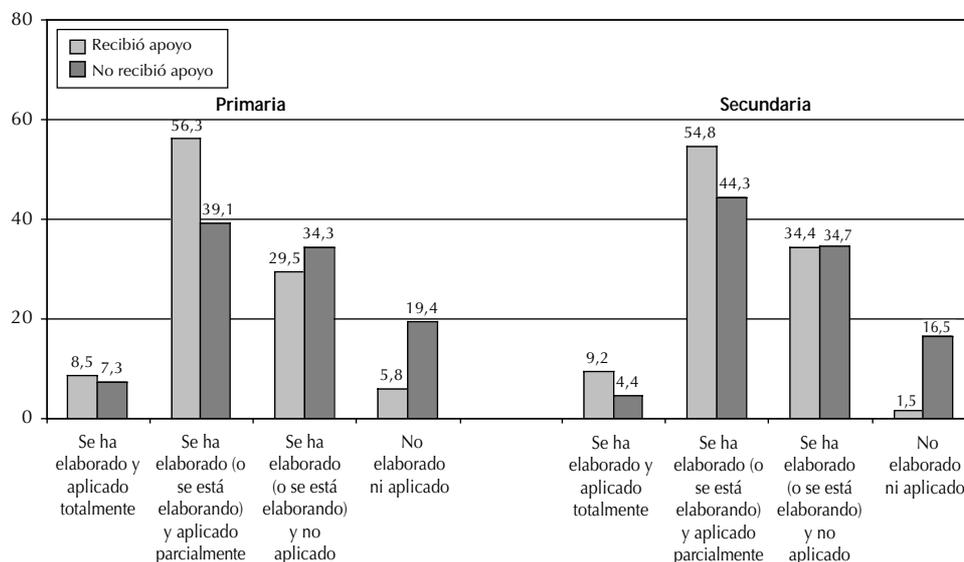
**Gráfico 8**

**Frecuencia de participación de padres de familia en la elaboración del PDI según participación del director en PLANCGED. Resultados de CE estatales de primaria y secundaria (Porcentajes)**



*Nota:* Número de casos: CE de primaria, asistió a PLANCGED: 11 792; CE de primaria, no asistió a PLANCGED: 4392; CE de secundaria, asistió a PLANCGED: 9794; y CE de secundaria, no asistió a PLANCGED: 6307.  
*Fuente:* Profesores de los alumnos evaluados.

**Gráfico 9**  
**Recibió apoyo de la instancia administrativa superior para elaborar el PDI.**  
**Resultados de CE estatales de primaria y secundaria (Porcentajes)**



Fuente: Directores de CE.

#### **4. APOYO DE LAS INSTANCIAS ADMINISTRATIVAS SUPERIORES PARA ELABORAR, DESARROLLAR O IMPLEMENTAR EL PDI**

Como instancias de ejecución descentralizadas que tienen la responsabilidad general de desarrollar y administrar la educación en CE y programas de su jurisdicción<sup>13</sup>, los órganos intermedios deberían ser instituciones de apoyo y soporte tanto en el proceso de formulación como en el de implementación del PDI. Por ello se preguntó a los directores si habían recibido apoyo de su instancia inmediata superior.

En primaria, 45,2% de los directores refiere haber recibido apoyo de su instancia; este porcentaje es de 47,5% para CE estatales y de 31,7% para no estatales. En secundaria, 42,7% de los directores encuestados refiere haber recibido apoyo: 44,3% de CE estatales y 35,9% de los no estatales.

En el gráfico 9 se puede observar, en ambos niveles, que el apoyo de la instancia inmediata superior está relacionado con la elaboración y aplicación del PDI. De los CE que

han recibido apoyo, sólo 5,8% en primaria y 1,5% en secundaria no han elaborado ni aplicado el PDI.

#### **5. RESUMEN Y CONCLUSIONES**

Los análisis de este informe han sido desarrollados de manera descriptiva, lo que nos ha permitido identificar algunas asociaciones interesantes entre las variables de estudio. Sin embargo, el tipo de análisis realizado no hace posible establecer relaciones causales entre dichas variables. Por otra parte, es posible que otras variables no consideradas en este documento sean importantes para entender los resultados.

Las observaciones anteriores deben ser tomadas en cuenta al momento de interpretar

<sup>13</sup> MED: *Reglamento de organización y funciones de las Direcciones de Educación de Lima y del Callao y de las Unidades de Servicios Educativos*. Lima: MED, 1996.

los resultados de este trabajo, los que a continuación se resumen.

En la mayoría de CE de primaria y secundaria que forman la muestra, la elaboración y/o implementación del PDI están en proceso. Esto dice mucho del alto grado de difusión y generalización que está alcanzando esta herramienta de gestión institucional. Dentro de este grupo de CE, el porcentaje de aquellos que han elaborado el PDI es superior al de los que lo han implementado. Estos resultados concuerdan con lo esperado, debido a que la secuencia de desarrollo del proyecto supone en primer lugar su elaboración.

En los CE de primaria y secundaria donde el PDI estuvo elaborado parcial o completamente, es mayor el porcentaje de directores que reporta haberse reunido con profesores y padres de familia para la elaboración de este proyecto. Se observa, así, que en un porcentaje importante de CE se ha logrado implicar en tareas vinculadas al proyecto institucional de la escuela a los principales agentes formadores, esto es, padres y profesores. Sin embargo, una cuarta parte de los directores de CE donde se ha elaborado el PDI refiere haber llevado a cabo este proceso sin la participación de ninguno de estos agentes. Esta situación incrementa el riesgo en esas escuelas de que el PDI, al no haber sido elaborado a partir de un esfuerzo compartido, sea visto por algunos como algo impuesto y poco adaptado a las necesidades reales de la comunidad educativa.

Por otro lado, se ha encontrado que la mayoría de profesores encuestados refiere haber participado al menos en alguna de las actividades del PDI. En contraste, se observa que los porcentajes de participación de padres de familia son mucho más bajos: la mayoría de ellos refiere no haber participado en la elaboración del PDI.

Asimismo, en ambos niveles educativos la participación de profesores en la formulación del PDI ha sido, en primer lugar, en actividades vinculadas con la elaboración del diagnóstico del CE. En segundo y tercer lugar respectivamente, los profesores han participado en la formulación de la propuesta

pedagógica y la definición de la visión y misión de la escuela. Y, por último, se han encontrado porcentajes menores de participación de profesores en la formulación de la propuesta de gestión. Este resultado no debe extrañar si se considera que esta última actividad está principalmente vinculada con aspectos administrativos e institucionales del CE, que tienen menor relación con las funciones que los docentes deben cumplir.

Por otro lado, la comparación de los resultados entre CE de primaria y secundaria muestra que en ambos niveles se ha elaborado o aplicado el PDI en porcentajes similares, aunque los niveles de participación de padres y profesores hayan sido superiores en los CE de primaria.

De otra parte, si se comparan los resultados de los CE según su tipo de gestión, se puede observar que es mayor el porcentaje de CE no estatales que ha elaborado y aplicado el PDI totalmente. Sobre los niveles de participación de profesores y padres, la información reportada en las diferentes encuestas parece poco consistente entre sí. Mientras el porcentaje de directores que manifiesta haberse reunido con profesores para tratar el tema del PDI es más alto en los CE estatales, en términos generales es mayor el porcentaje de profesores que manifiesta haber participado en las actividades relacionadas con la elaboración del PDI en los CE no estatales.

Sobre el impacto de la capacitación de PLANCGED en la elaboración y desarrollo del PDI, se observa, en primer lugar, que la asistencia a estas capacitaciones está asociada a porcentajes más altos de elaboración y/o aplicación del PDI en los colegios. Por otro lado, en primaria parece que la asistencia del director a PLANCGED está asociada a mayores niveles de participación e implicación de los profesores en el proceso de elaboración del PDI.

Finalmente, otro aspecto que parece favorecer la implementación del PDI en las escuelas es el apoyo de las instancias administrativas superiores. Así, pues, la mayoría de CE que han recibido dicho apoyo han iniciado la elaboración y/o aplicación del PDI.

# ANEXO 1

## PREGUNTAS ANALIZADAS

### ENCUESTA PARA DIRECTORES

En su centro educativo, ¿se ha elaborado el PDI?

- a) Sí
- b) Está en elaboración
- c) No

¿El PDI está siendo aplicado este año?

- a) Sí, totalmente
- b) Sí, parcialmente
- c) No

Después de Fiestas Patrias, ¿se ha reunido con algún profesor o profesores para elaborar el PDI?

- a) Sí
- b) No

A lo largo del año escolar, ¿se ha reunido con padres de familia de primaria para elaborar el PDI?

- a) Sí
- b) No

Este año, ¿su instancia administrativa superior ha apoyado en la elaboración, desarrollo o implementación del PDI?

- a) Sí
- b) No

### ENCUESTA PARA RESPONSABLES DE AULA (EN PRIMARIA) Y PARA PROFESORES DE MATEMÁTICA Y LENGUAJE Y LITERATURA (EN SECUNDARIA)

En su centro educativo, ¿ha participado usted en alguna de las siguientes actividades relacionadas con el PDI?

- |                                       |    |    |
|---------------------------------------|----|----|
| a) Elaborar el diagnóstico del centro | Sí | No |
| b) Definir la visión y misión         | Sí | No |
| c) Formular la propuesta pedagógica   | Sí | No |
| d) Formular la propuesta de gestión   | Sí | No |

## ENCUESTA PARA PADRES DE FAMILIA/APODERADOS DE PRIMARIA

Este año, ¿ha asistido a la escuela para colaborar en la elaboración del PDI?

Primaria

- a) Nunca
- b) Pocas veces
- c) Muchas veces
- d) Siempre

Secundaria

- a) Nunca
- b) Pocas veces
- c) Muchas veces o siempre
- d) No sabe qué es el PDI

## ANEXO 2

### MUESTRA Y PROCEDIMIENTOS

Este documento usa información recogida en el marco del proyecto de Medición del Rendimiento Estudiantil CRECER 1998, donde se aplicaron, adicionalmente a las pruebas de rendimiento<sup>1</sup>, encuestas dirigidas a directores, docentes, padres de familia y alumnos.

Los instrumentos fueron aplicados la última semana de noviembre y la primera de diciembre de 1998, últimos meses del año escolar. La aplicación estuvo a cargo de personal externo a los CE evaluados, especialmente preparado para manejar los instrumentos. Se puso énfasis en que las encuestas eran confidenciales y no serían revisadas por otros colegas del CE.

El proyecto CRECER se ejecutó en CE polidocentes completos<sup>2</sup> de zonas urbanas a nivel nacional; por ello sus resultados son representativos.

Los cuadros 1 y 2 muestran el número de sujetos considerados en los análisis.

Cuadro 1 Muestra de primaria			
	Directores	Docentes <sup>3</sup>	Familiares
Estatal	487	981	19 389
No estatal	86	172	3 361
<b>Total</b>	<b>573</b>	<b>1153</b>	<b>22 750</b>

*Nota:* La muestra de docentes está formada por 576 docentes de cuarto grado y 577 de sexto grado; la muestra de familiares, por 11 451 familiares de alumnos de cuarto grado y 11 299 de alumnos de sexto grado. Se eligió presentar los resultados en forma agregada (juntando los dos grados) porque los desagregados por grado no mostraban diferencias significativas.

Cuadro 2 Muestra de secundaria			
	Directores	Docentes <sup>4</sup>	Familiares
Estatal	455	932	18 528
No estatal	105	210	4 093
<b>Total</b>	<b>560</b>	<b>1 142</b>	<b>22 621</b>

*Nota:* La muestra de docentes está formada por docentes de cuarto grado; se obtuvo también información de quinto grado, pero no se reportó porque había docentes que enseñaban en ambos grados. Además, la información de quinto grado resultó muy parecida a la de cuarto. La muestra de familiares está formada por 11 351 familiares de alumnos de cuarto grado y 11 270 de alumnos de quinto grado. Se eligió presentar los resultados en forma agregada (juntando los dos grados), pues desagregados por grado no mostraban diferencias significativas.

1 Las áreas de desarrollo evaluadas en primaria fueron: en cuarto grado, Lógico-Matemática, Comunicación Integral, Personal Social, y Ciencia y Ambiente; en sexto grado, Matemática, Lenguaje, Ciencias Histórico-Sociales y Ciencias Naturales. Los cursos evaluados en secundaria fueron Matemática y Lengua y Literatura.

2 Esto es, en CE donde existe por lo menos un profesor para cada grado.

3 En primaria se encuestó a docentes responsables de aula.

4 En secundaria se encuestó a docentes de Lenguaje y de Matemática.

## ANEXO 3

### Nº MERO DE CASOS DEL CUADRO 3 Y GRÁFICOS 5 Y 6

	Primaria				
	Total	Estatal	No estatal	Estatal	
				Asistió a PLANCGED	No asistió a PLANCGED
Elaborar el diagnóstico del CE	960	817	143	583	217
Definir la visión y la misión	921	776	145	547	211
Formular la propuesta pedagógica	918	774	144	550	207
Formular la propuesta de gestión	902	763	139	544	204
	Secundaria				
	Total	Estatal	No estatal	Estatal	
				Asistió a PLANCGED	No asistió a PLANCGED
Elaborar el diagnóstico del CE	999	811	188	486	317
Definir la visión y la misión	974	788	186	475	306
Formular la propuesta pedagógica	981	793	188	478	308
Formular la propuesta de gestión	970	787	183	476	304

# RENDIMIENTO Y ACTITUDES HACIA LA MATEMÁTICA EN EL SISTEMA ESCOLAR PERUANO

JORGE BAZÁN  
GIULIANA ESPINOSA  
CHOLLY FARRO

El nuevo enfoque curricular propuesto desde el MED establece que la dimensión afectiva del aprendizaje resulta esencial para el logro de las competencias y propósitos educativos que el sistema escolar se plantea, razón por la cual esta dimensión debe ser atendida y desarrollada desde la escuela.

En esta línea, las actitudes hacia el aprendizaje y hacia las diversas asignaturas o áreas curriculares parecen constituir un aspecto esencial para entender los resultados académicos que los estudiantes consiguen. Sin embargo, la relación entre estas dos variables está lejos de ser clara.

Esta sección del documento es el resultado de una investigación llevada a cabo a partir de los resultados de la evaluación nacional del rendimiento estudiantil realizada en el Perú por la Unidad de Medición de la Calidad del Ministerio de Educación-UMC (CRECER 1998). Junto con las pruebas de rendimiento en Matemática, en este proceso de evaluación se aplicaron encuestas a alumnos en las que se indagó, entre muchas otras variables, por cuatro aspectos que se tomarán como punto de partida para este estudio: 1) la dificultad percibida para el aprendizaje de la Matemática; 2) el temor del alumno de participar en clase; 3) el gusto por la materia en cuestión; y, finalmente, 4) el nivel percibido de comprensión de las explicaciones de su profesor.

En esta sección se busca mostrar el grado de asociación entre el rendimiento y la

actitud hacia la Matemática de los estudiantes evaluados en CRECER 1998. Para ello, en el acápite de resultados se presentan, en primer lugar, los niveles de asociación hallados entre el rendimiento en Matemática y cada uno de los cuatro aspectos mencionados; en segundo lugar, se muestra la relación existente entre el rendimiento y los perfiles actitudinales construidos sobre la base de dichos aspectos; finalmente se presentan los rendimientos para cada perfil desagregados según tipo de gestión del centro educativo (CE) y género del alumno.

## **1 MARCO TEÓRICO: APROXIMACIÓN AL CONCEPTO DE ACTITUDES**

### **ACTITUDES EN EL CONTEXTO EDUCATIVO**

Las actitudes son definidas como la tendencia psicológica que se expresa a través de la evaluación favorable o desfavorable de una entidad en particular<sup>1</sup>. Dicha entidad puede ser un objeto, una persona, un suceso o cualquier evento capaz de ser valorado.

1 Eagly, A. y S. Chaiken: "Attitude Structure and Function", en D.T. Gilbert, S.T. Fiske y G. Lindzey: *The Handbook of Social Psychology*, vol. 1, pp. 269-322. New York: McGraw-Hill, 1998 (4ª edición).

Las actitudes son inferidas de lo que una persona manifiesta acerca del objeto actitudinal, puesto que no son directamente observables ni se traducen necesariamente en conductas<sup>2</sup>. De esta manera, aunque las actitudes predisponen a actuar en consonancia con la evaluación, es posible que, debido a la influencia del entorno, las conductas no siempre sean consistentes con aquéllas. Por ejemplo, en el caso de la escuela podríamos observar que en ocasiones algunos alumnos, pese a presentar una actitud desfavorable ante una asignatura, pueden invertir tiempo y esfuerzo en estudiarla debido a las exigencias académicas que plantea su profesor, o, más aún, a las presiones familiares.

Otra característica de las actitudes es que son adquiridas y persistentes en el tiempo, aunque pueden cambiar<sup>3</sup>. Este aspecto cobra especial importancia en el ámbito educativo, donde debe existir un esfuerzo intencionado por reforzar actitudes que favorezcan el aprendizaje y cambiar aquellas que lo perjudiquen.

Al respecto, Pozo y Gómez<sup>4</sup> postulan que la forma de organizar las actividades de enseñanza y aprendizaje selecciona y refuerza ciertas actitudes en los alumnos, aunque en la mayor parte de los casos no exista un propósito explícito de enseñarlas. Muchas veces el carácter implícito de este proceso lleva a transmitir actitudes contrarias a los propósitos que la educación se plantea, lo que ha sido llamado por algunos autores la transmisión del currículo oculto<sup>5</sup>. Por ejemplo, con frecuencia los criterios de evaluación que emplean algunos profesores contradicen los objetivos trazados por ellos mismos. Imaginemos el caso

del profesor que desea estimular la creatividad en la resolución de problemas en el área de Lógico-Matemática, pero que, pese a ello, en sus evaluaciones sólo plantea ejercicios en los que basta que el alumno aplique mecánicamente el algoritmo correspondiente. El mensaje que —sin querer— este profesor estaría enviando a sus alumnos sería el siguiente: “lo que realmente importa es que repitan mecánicamente lo que les enseñé”.

Hacer consciente dicho proceso y orientarlo a estimular de manera deliberada y eficaz actitudes favorables hacia el aprendizaje y las asignaturas requiere considerar y trabajar con los tres componentes básicos de toda actitud: cognitivo, afectivo y conductual<sup>6</sup>.

El *componente cognitivo* de la actitud está definido por los conocimientos y creencias de una persona sobre el objeto de la actitud. Por ejemplo, la creencia del alumno de que la Matemática es difícil de aprender sería una categorización sobre el objeto actitudinal (en este caso el aprendizaje de la Matemática) que propiciaría una actitud de evitamiento frente a él, especialmente si aquel alumno es de los que no se siente lo suficientemente competente para abordar con éxito una materia “difícil”.

El *componente afectivo*, por su parte, supone una valoración emocional del objeto actitudinal. Por ejemplo, el alumno que manifiesta gusto por la Matemática mostrará probablemente una tendencia a la aproximación y aceptación de esta materia.

Finalmente, el *componente conductual o conativo* de la actitud está definido por las acciones manifiestas y la declaración de intenciones de una persona sobre el objeto de la actitud. Así, si un alumno participa espontáneamente en la clase de Matemática puede estar mostrando una actitud favorable hacia dicha materia, que muy probablemente repercutirá de manera positiva en su nivel de aprovechamiento.

Si bien existe controversia sobre la manera en que estos tres componentes interactúan, los diversos autores que han desarrollado el tema del cambio de actitudes postulan que para llevarlo a cabo es necesario trabajar con los tres. Por eso, muchos programas de intervención fracasan en su intento de modificar actitudes al trabajar sólo con un componente<sup>7</sup>. Un claro ejemplo de esto se observa

2 Summers, G.: *Medición de actitudes*. México D.F.: Trillas, 1976.

3 *Ibidem*.

4 Pozo, J.I. y M.A. Gómez Crespo: *Aprender y enseñar ciencia*. Madrid: Ediciones Morata, 2000 (2ª edición).

5 Torres, J.: *El currículum oculto*. Madrid: Morata, 1994 (4ª edición).

6 Abordaremos el presente informe desde el enfoque tricomponential. Para conocer mejor este enfoque se puede revisar Triandis, H.: *Attitude and Attitude Change*. New York: Wiley & Sons, 1971.

7 Costa, M. y E. López: *Educación para la salud*. Madrid: Martínez Roca, 1996.

con frecuencia en las aulas donde el profesor intenta cambiar la actitud del alumno mediante comunicaciones verbales como exhortaciones o información verbal, pero descuidando las dimensiones afectiva y conativa de la actitud, lo que lo conduce al fracaso.

## ACTITUDES Y DESEMPEÑO ESCOLAR

Diversos estudios internacionales han mostrado que, en general, existe una relación significativa y directa entre las actitudes de los alumnos y el rendimiento en Matemática. Por ejemplo, en el estudio del TIMSS (Third International Math and Science Study) realizado entre los años 1994 y 1995 con la participación de 41 países, se observó una relación positiva entre el gusto por la Matemática y las puntuaciones obtenidas en las pruebas de esta asignatura, de tal manera que el puntaje promedio de aquellos alumnos que manifestaban gusto por dicha materia era superior que el de aquellos que reportaban que la Matemática no les gustaba.

Por su parte, los estudios del National Assessment of Education Progress (NAEP) realizados entre los años 1994 y 1996 en Estados Unidos revelaron que existe asociación entre el gusto por la Matemática y la disposición de los alumnos para estudiar más de esa materia. Esto está en relación con la elección de cursos avanzados durante la secundaria. Los resultados de este estudio sugirieron, además, que la mayoría de estudiantes de los diferentes grados evaluados manifestaban una actitud favorable hacia la Matemática; sin embargo, los porcentajes de aceptación declinaban conforme se avanzaba en la carrera escolar.

Si bien en los estudios mencionados —y, en general, en la literatura que trata sobre el tema— se resalta la asociación de las actitudes con el desempeño de los estudiantes, es preciso considerar que existe la posibilidad de que un alumno pueda alcanzar un nivel de rendimiento satisfactorio y, pese a ello, tener una actitud desfavorable frente a la materia, y viceversa. De esta forma, una actitud positiva no garantiza un mejor rendimiento, aunque sí eleva la probabilidad de que éste se dé.

Finalmente, cabe mencionar que la relación entre actitud y rendimiento es bidireccional y compleja. El bajo rendimiento, el fracaso repetido o una historia académica difícil pueden ser producto y a la vez origen de actitudes negativas hacia las materias y hacia la escuela en general<sup>8</sup>.

## ALGUNOS ASPECTOS QUE CONFIGURAN LAS ACTITUDES: PERCEPCIONES, TEMORES Y PREFERENCIAS

Las percepciones y creencias de un estudiante sobre sus habilidades y posibilidades personales configuran su sentido de autoeficacia, que en interacción con la evaluación que realice sobre la situación de aprendizaje favorecerá una actitud de aproximación o rechazo, según el caso. De esta forma, los estudiantes hacen interpretaciones personales sobre sus logros y fracasos y establecen metas basándose en dichas interpretaciones. De acuerdo con Bandura<sup>9</sup>, la gente tiende a evitar aquellas situaciones en las que anticipa que sus capacidades se verán excedidas por la dificultad de la tarea<sup>10</sup>.

Vinculado con esto, otro aspecto que configurará la actitud del estudiante frente a la situación de aprendizaje es la ansiedad que se genera en el aula. Muchas veces dicha ansiedad encuentra su mayor manifestación en el temor de participar en clase. Este temor puede tener su origen en la interacción de diversas variables. De un lado, como veíamos antes, el tipo de evaluación que haga el alumno sobre su propia competencia influirá en el nivel de ansiedad que experimente. Pero, además, la relación que se establezca entre él y su profesor, el tipo de actividades que éste le plantee en clase y, sobre todo, la manera en que el docente reciba y retroalimente sus participaciones serán, también,

8 Alonso, J.: *Motivación y aprendizaje en el aula: Cómo enseñar a pensar*. Madrid: Santillana, 1998.

9 Bandura, Albert: *Social Foundations of Thought and Action: a Social Cognitive Theory*. New York: Prentice Hall, 1986.

10 Tollefson, N.: "Classroom Applications of Cognitive Theories of Motivation", en *Educational Psychology Review* 12, pp. 63-83, 2000.

elementos que fomenten un clima determinado en el aula y que incrementen o reduzcan el nivel de seguridad experimentado por el estudiante.

Desde la psicología educativa se postula que la participación activa del alumno en clase favorece su involucramiento en el proceso educativo y, por tanto, su nivel de desempeño y logro. Por ello, al margen de cuál sea el origen del temor de participar en clase, su sola existencia plantea la posibilidad de que el nivel de compromiso del estudiante se vea reducido y que estos sentimientos negativos se generalicen a toda la situación de aprendizaje de la Matemática, con lo que se configuraría una actitud negativa hacia ella.

Percibir que no se entiende nada o casi nada de las explicaciones del profesor podría ser otro factor de riesgo, ya que puede llevar consigo un sentimiento de pérdida de control sobre el propio aprendizaje y, por ende, desmoralizar o frustrar. Con frecuencia los estudiantes realizan su valoración del contenido del área (en este caso Matemática) basándose en sus experiencias educativas personales. Si la comprensión de las explicaciones del profesor es deficiente, probablemente también lo sea el proceso de aprendizaje del alumno, y esto puede constituir la base para una actitud de rechazo o desagrado frente a las actividades de ese curso.

Finalmente, se ha considerado que el gusto o agrado por la clase de Matemática se asociará en gran medida al tipo de orientación que el estudiante muestre frente a dicha clase. Esta variable en sí misma denota una actitud (es decir, una evaluación con carga afectiva del objeto actitudinal) sobre la clase de Matemática.

---

11 Estos porcentajes del tipo de gestión del CE mantienen la proporción del universo.

12 La muestra de estudiantes efectivamente evaluados en cualquiera de las pruebas es en todos los casos ligeramente menor que la señalada debido a ausencias el día de la administración. En ningún caso esta pérdida es mayor del 5% de la muestra planificada.

13 Se estableció una serie de evidencias de validez. Éstas van desde aspectos cualitativos, como la opinión de expertos sobre los contenidos de las pruebas, hasta indicadores de confiabilidad y unidimensionalidad de las pruebas, pasando por in-

## **2. MÉTODO**

### **PARTICIPANTES**

Se evaluó una muestra representativa nacional de estudiantes de CE urbanos polidocentes completos de cuarto y sexto de primaria, y cuarto y quinto de secundaria.

La muestra estuvo constituida por 1142 CE, de los cuales 576 eran de primaria y 566 de secundaria. Para cada grado se planificó evaluar 17 370 estudiantes (14 820 de CE estatales y 2550 de no estatales, aproximadamente 85% y 15% respectivamente)<sup>11</sup>, de los cuales la mitad fueron hombres y la otra mitad mujeres<sup>12</sup>.

La muestra fue diseñada para permitir desagregaciones a escala departamental, entre hombres y mujeres y entre estudiantes de CE estatales y no estatales. El número de estudiantes por departamento fue variable, pero nunca por debajo de 500.

El método de muestreo utilizado fue probabilístico bietápico con estratos en la primera etapa y conglomerados en la segunda. La muestra incluyó 30 estudiantes por grado evaluado en cada CE. La selección de los CE, de las secciones (una por grado) y de los estudiantes dentro de cada sección fue aleatoria.

Para realizar los análisis de esta investigación se tomó en cuenta sólo la información de aquellos estudiantes que resolvieron las pruebas de rendimiento en Matemática y que contestaron las cuatro preguntas que indagaban sobre las actitudes hacia esta materia. Como consecuencia de esto, la muestra efectiva se redujo a 16 601 alumnos para cuarto grado de primaria, 16 425 para sexto grado de primaria, 16 085 para cuarto de secundaria y 15 678 para quinto de secundaria.

### **INSTRUMENTOS**

Las pruebas nacionales de Matemática y Lenguaje de cuarto y sexto grado de educación primaria y cuarto y quinto de secundaria presentan propiedades psicométricas óptimas para su empleo en el análisis de resultados<sup>13</sup>.

Sus puntajes se muestran en una escala relativa arbitraria basada en el modelo

Cuadro 1 Categorías de percepciones, temores y preferencias vinculadas a la materia o clase de Matemática					
Aspecto/ Nivel de comprensión percibido	Categorías				
	Muy bajo	Bajo	Medio	Alto	Muy alto
Nivel de agrado	Gusto	Desagrado			
Nivel de seguridad	Temor	Seguridad			
Capacidad de aprendizaje percibida	Percepción de dificultad	Percepción de competencia			

Rasch<sup>14</sup> de media 300 y desviación estándar 50. El uso de escalas relativas para la presentación de resultados de rendimiento se sustenta en el enfoque de normas seguido en la construcción de estos instrumentos. Las escalas absolutas —como el puntaje total o porcentaje— no resultan apropiadas para la presentación de resultados, pues podrían prestarse a interpretaciones erróneas sobre lo que saben o no los estudiantes.

Por su parte, las encuestas fueron diseñadas para recoger información sobre los factores asociados con el rendimiento. Para el análisis se seleccionaron las siguientes preguntas de las encuestas aplicadas a alumnos:

- De las cosas que te explica tu profesor/a de Matemática, entiendes:
  - a) Nada
  - b) Casi nada
  - c) Sólo algunas cosas
  - d) Casi todo
  - e) Todo
- ¿Te gustan las clases de Matemática?
  - a) Sí    b) No
- ¿Tienes miedo de preguntar en la clase de Matemática?/¿Te sientes nervioso cuando tienes que hablar en clase de Matemática?<sup>15</sup>
  - a) Sí    b) No
- ¿Aprender Matemática es difícil para ti?
  - a) Sí    b) No

En el cuadro 1 se muestran las categorías que se han desprendido de las preguntas anteriores y que se emplearán a lo largo de esta sección para describir los resultados.

## PROCEDIMIENTO

Para la administración de las pruebas y encuestas a los alumnos se formó una Red Administrativa Nacional que incluía monitores (de la sede central), supervisores y coordinadores (de los órganos intermedios), y examinadores (docentes locales). Se elaboraron manuales de procedimientos que fueron explicados en capacitaciones que se llevaron a cabo tanto en Lima como en las sedes de administración de las pruebas durante la última semana de noviembre y la primera de diciembre de 1998.

Los análisis estadísticos fueron realizados utilizando el programa SPSS. Para el análisis de las preguntas de las encuestas y su relación con el rendimiento se emplearon análisis

---

dicadores más cuantitativos en el análisis de las preguntas, como índices de dificultad, discriminación, correlación punto biserial y no respuesta adecuados.

14 Un modelo de la Teoría de Respuesta a los Ítemes que privilegia el análisis de las preguntas sobre el análisis de la prueba. Para mayores detalles puede revisarse Muñiz, J.: *Teoría de respuesta a los ítemes: Un nuevo enfoque en la evaluación psicológica educativa*. Madrid: Ediciones Pirámide, 1990.

15 La primera frase aparecía en la encuesta de primaria; la segunda, en la de secundaria.

Cuadro 2 Perfiles de actitudes hacia la Matemática sobre la base de las categorías de respuesta de los alumnos			
Matemática			
	Perfil	Primaria	Secundaria
Actitudes <i>no favorables</i> para el aprendizaje	A	<ul style="list-style-type: none"> <li>• Temor</li> </ul>	<ul style="list-style-type: none"> <li>• Temor</li> </ul>
	B	<ul style="list-style-type: none"> <li>• Percepción de dificultad</li> <li>• Desagrado</li> <li>• Comprensión (de muy baja a alta)</li> </ul>	<ul style="list-style-type: none"> <li>• Percepción de dificultad</li> <li>• Desagrado</li> <li>• Comprensión (de muy baja a media)</li> </ul>
Actitudes <i>favorables</i> para el aprendizaje	C	<ul style="list-style-type: none"> <li>• Seguridad</li> </ul>	<ul style="list-style-type: none"> <li>• Seguridad</li> <li>• Comprensión alta</li> </ul>
	D	<ul style="list-style-type: none"> <li>• Percepción de competencia</li> <li>• Gusto</li> <li>• Comprensión muy alta</li> </ul>	<ul style="list-style-type: none"> <li>• Percepción de competencia</li> <li>• Gusto</li> <li>• Comprensión muy alta</li> </ul>

sis de distribuciones y tablas cruzadas. Para la elaboración del indicador de actitud se usó el análisis factorial de correspondencias múltiples, llamado también análisis de homogeneidad (HOMALS) o *dual scaling*<sup>16</sup>. En el anexo se detalla el procedimiento realizado para la construcción del indicador y de los perfiles actitudinales que se presentan a continuación.

El cuadro 2 revela las actitudes favorables y desfavorables subyacentes a las categorías de las preguntas a través de cuatro perfiles.

En este cuadro, cada una de las celdas incluye una o más categorías que configuran un perfil actitudinal (A, B, C o D) para cada nivel educativo (primaria o secundaria). Los dos primeros expresan actitudes desfavorables hacia la Matemática, mientras que los últimos definen actitudes favorables hacia dicha materia y su enseñanza.

16 Abascal, E. e I. Grande: *Métodos multivariantes para la investigación comercial: Teoría, aplicaciones y programación Basic*. Barcelona: Ariel, 1989; Visauta, B.: *Análisis estadístico con SPSS para Windows*. Madrid: McGraw-Hill, 1998.

### 3. RESULTADOS

#### RELACIÓN ENTRE EL RENDIMIENTO EN MATEMÁTICA Y LOS NIVELES DE COMPRESIÓN, DE AGRADO, DE SEGURIDAD Y DE CAPACIDAD PERCIBIDA PARA APRENDER DICHA MATERIA

El cuadro 3 muestra los porcentajes de alumnos que se ubican en cada una de las categorías que indagan sobre el nivel de comprensión percibido, grado de seguridad experimentada para participar en clase, percepción de competencia y nivel de agrado, así como las medias de rendimiento para cada categoría. Los resultados pueden apreciarse para los cuatro grados evaluados.

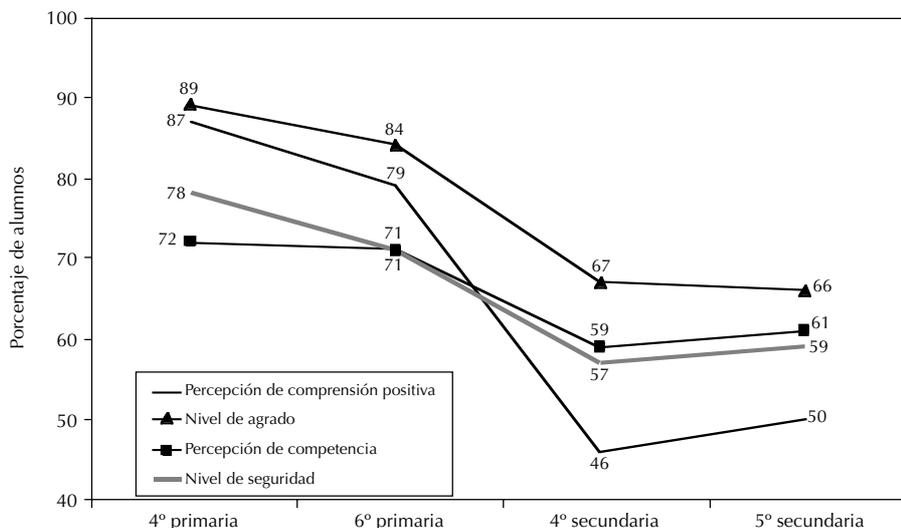
En el cuadro se observa que los mayores porcentajes de alumnos se encuentran en las categorías consideradas aquí como más favorables hacia la Matemática: comprensión positiva (muy alta y alta), percepción de competencia, gusto y seguridad. Esta tendencia se observa en todos los grados evaluados. Sin embargo, se aprecia también que dichos porcentajes son sensiblemente menores en secundaria que en primaria (véase el gráfico 1). Esto

**Cuadro 3**  
**Rendimiento en Matemática de los alumnos evaluados en CRECER 1998 según sus percepciones, preferencias y temores vinculados a la Matemática y su aprendizaje**

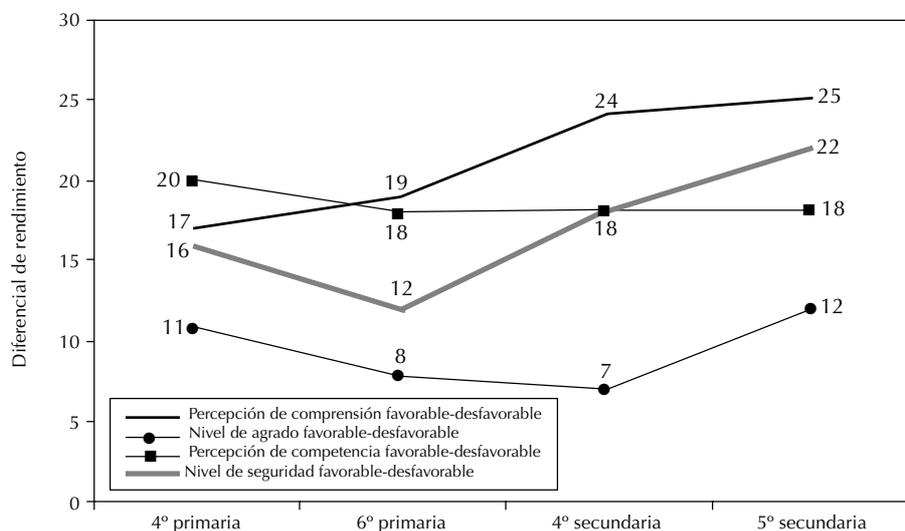
Aspecto	Categorías	Educación primaria				Educación secundaria			
		4° grado		6° grado		4° grado		5° grado	
		%	Rendimiento	%	Rendimiento	%	Rendimiento	%	Rendimiento
Nivel percibido de comprensión de la Matemática	Muy bajo	1	265	1	281	2	288	2	283
	Bajo	1	269	1	278	6	290	5	290
	Medio	11	288	19	286	45	289	43	287
	<b>Alto</b>	<b>25</b>	<b>301</b>	<b>38</b>	<b>300</b>	<b>35</b>	<b>307</b>	<b>37</b>	<b>305</b>
	<b>Muy alto</b>	<b>62</b>	<b>303</b>	<b>41</b>	<b>307</b>	<b>11</b>	<b>333</b>	<b>13</b>	<b>331</b>
Capacidad percibida de aprendizaje	Percepción de dificultad	28	286	29	287	41	290	39	289
	<b>Percepción de competencia</b>	<b>72</b>	<b>306</b>	<b>71</b>	<b>305</b>	<b>59</b>	<b>308</b>	<b>61</b>	<b>307</b>
Nivel de agrado	Desagrado	11	290	16	293	33	296	34	292
	<b>Gusto</b>	<b>89</b>	<b>301</b>	<b>84</b>	<b>301</b>	<b>67</b>	<b>303</b>	<b>66</b>	<b>304</b>
Nivel de seguridad experimentado para participar	Temor	22	288	29	292	43	290	41	287
	<b>Seguridad</b>	<b>78</b>	<b>304</b>	<b>71</b>	<b>304</b>	<b>57</b>	<b>308</b>	<b>59</b>	<b>309</b>
Muestra nacional		100	300	100	300	100	300	100	300

\* Se resaltan en negritas las categorías que, se presume, son favorables hacia la Matemática.

**Gráfico 1**  
**Percepción de competencia, comprensión positiva, seguridad y agrado hacia la Matemática según los grados escolares**



**Gráfico 2**  
**Diferencia en el rendimiento en Matemática según respuestas favorables y respuestas desfavorables hacia el aprendizaje de la Matemática**



podría ser un indicio de que con el paso de los años las actitudes positivas hacia la Matemática en el sistema escolar decrecen.

Por otro lado, el cuadro permite apreciar que el rendimiento de los alumnos en Matemática está relacionado positivamente con sus respuestas en las preguntas consideradas para el estudio. Así, por ejemplo, el grupo de alumnos que manifiestan un muy alto nivel de comprensión de las explicaciones de su profesor presenta un mayor rendimiento promedio que el resto de los grupos. En cuanto a los otros tres aspectos evaluados se encontró que, de acuerdo con lo que se esperaba, los mejores rendimientos medios los obtuvieron los grupos en los que se manifestaba sentido de competencia para el aprendizaje de la Matemática (frente al grupo que manifestaba percibir dificultad para aprender), gusto por dicha materia (frente al grupo que manifiesta desagrado) y seguridad para participar de manera activa en las clases de

ésta (frente al grupo de aquellos con temor a participar).

El gráfico 2 se ha obtenido restando al promedio de rendimiento en Matemática de los estudiantes que tienen respuestas favorables, el de aquellos que presentan respuestas desfavorables. Esta operación se realizó para cada pregunta en cada grado evaluado. Por ejemplo, para la pregunta sobre el nivel de agrado por la Matemática en cuarto grado de primaria se obtuvo que los alumnos con respuesta favorable (agrado) presentan en promedio 11 puntos por encima de aquellos con respuesta desfavorable (desagrado).

Las diferencias en el rendimiento de aquellos que expresan respuestas favorables respecto de los que expresan respuestas desfavorables en cada aspecto y en todos los grados son mostradas en el gráfico 1. En los 16 casos estas diferencias son estadísticamente significativas ( $p < 0,01$ )<sup>17</sup>.

Como se aprecia en el cuadro 2 y el gráfico 2, con el transcurso de los años escolares se muestra una tendencia al aumento de la diferencia de rendimiento entre los grupos que manifiestan diferentes grados de comprensión de las explicaciones de su profesor. Por ejemplo, mientras que en cuarto de primaria el gru-

<sup>17</sup> Se empleó la prueba T de Student para muestras independientes. Las diferencias son significativas al 1%, pues el error estándar de cada comparación no sobrepasa los 2 puntos.

**Cuadro 4**  
**Rendimiento en Matemática según los perfiles actitudinales de los alumnos hacia dicha materia**

	Primaria		Secundaria	
	Porcentaje	Media	Porcentaje	Media
Perfil A desfavorable (temor)	15	285	16	290
Perfil B desfavorable (dificultad, desagrado, baja comprensión)	28	289	32	282
Perfil C favorable (seguridad)	50	311	43	308
Perfil D favorable (competencia, gusto, buena comprensión)	8	295	9	296
Nacional	100	300	100	300

po que revela tener un buen grado de comprensión de dichas explicaciones obtiene 17 puntos porcentuales por encima del grupo con bajos niveles de comprensión, en quinto de secundaria esta diferencia asciende a 25 puntos. Una tendencia similar se observa en la pregunta relacionada con el nivel de seguridad, en el que la diferencia va aumentando progresivamente, salvo por el hecho de que en sexto grado de primaria ésta se reduce con relación a la de cuarto. Esta variación ascendente en ambos casos podría significar que en los últimos grados de la secundaria la comprensión de las explicaciones del profesor y la seguridad para participar en clase tienen una asociación mayor con el rendimiento que en los grados de primaria.

En el caso de la pregunta que indaga por el nivel de competencia percibido, se observa una diferencia en el rendimiento que se mantiene constante con el paso del tiempo. Finalmente, en el caso del nivel de agrado no se observa ni estabilidad ni tendencia ascendente de acuerdo con los grados.

#### **RELACIÓN ENTRE LOS PERFILES ACTITUDINALES EN MATEMÁTICA Y EL RENDIMIENTO**

A continuación se evalúa la relación entre el rendimiento en Matemática y los perfiles ac-

titudinales mostrados en el acápite “Procedimiento”.

En el cuadro 4 se presentan los promedios de rendimiento en Matemática en primaria y secundaria para los diferentes perfiles identificados, así como la distribución porcentual de alumnos para cada perfil.

Allí se puede observar que son mayores los porcentajes de alumnos con perfiles de actitudes favorables (perfiles C y D) en primaria y secundaria, lo que muestra coherencia con los resultados obtenidos a partir del análisis de las preguntas aisladas (véase el cuadro 3). Sin embargo, aquellos porcentajes son más bajos en primaria que en secundaria. Asimismo, los resultados muestran que los alumnos con actitudes favorables tienen en promedio mejores rendimientos que los que presentan perfiles desfavorables. Por otro lado, los promedios de rendimiento en los diferentes perfiles presentados no difieren significativamente entre sí<sup>18</sup>.

Tal como se aprecia en el cuadro, los grupos de estudiantes de los grados evaluados en primaria se ordenan del siguiente modo de acuerdo con su rendimiento en Matemática:

18 A partir de las comparaciones múltiples mediante la prueba estadística de Scheffé.

1. *Grupo de mayor rendimiento:* Se muestran seguros para participar en clase (Perfil C).
2. *Grupo intermedio favorable:* Gustan de la clase, se perciben competentes para aprender el curso y manifiestan alcanzar un nivel de comprensión muy alto de las explicaciones del (de la) profesor/a (Perfil D).
3. *Grupo intermedio no favorable:* No les agrada la clase, perciben dificultad para aprender el curso, su nivel de comprensión del profesor/a fluctúa entre bajo y muy bajo (Perfil B).
4. *Grupo de menor rendimiento:* Alumnos que tienen temor de participar en clase (Perfil A).

Los resultados de secundaria, como los de primaria, ratifican la coherencia de lo obtenido con este indicador, pues a perfiles favorables, mayor rendimiento. Las comparaciones múltiples de los promedios de rendimiento entre los diferentes perfiles presentados manifiestan que éstos difieren significativamente entre sí<sup>19</sup>.

Tal como se aprecia en el cuadro 4, en secundaria los grupos de estudiantes se ordenan del siguiente modo de acuerdo con su rendimiento en Matemática:

1. *Grupo intermedio favorable:* Se sienten seguros para participar en clase y refieren un nivel de comprensión alto o muy alto (Perfil C).
2. *Grupo de mayor rendimiento:* Gustan de la clase, se perciben competentes para aprender el curso y tienen un nivel de comprensión del (de la) profesor/a muy alto (Perfil D).
3. *Grupo intermedio no favorable:* Sienten desagrado por las clases de Matemática, perciben dificultad para aprender el curso, y el nivel de comprensión que tienen de las explicaciones del (de la) profesor/a es bajo o muy bajo (Perfil A).
4. *Grupo de menor rendimiento:* Manifiestan temor para intervenir en clase (Perfil B).

19 A partir de la prueba estadística de Scheffé.

## RELACIONES ENTRE EL RENDIMIENTO Y LAS ACTITUDES SEGÚN GESTIÓN DE LA ESCUELA Y GÉNERO DEL ALUMNO

Las relaciones entre el rendimiento y las actitudes han sido abordadas en el acápite precedente. En éste se presenta una mirada de los perfiles de actitudes en Matemática según la gestión de la escuela y el género del alumno evaluado. Como antes, se detallan los promedios de rendimiento y el porcentaje de alumnos que se ubican en un determinado perfil.

En los cuadros 5 y 6 se observa que, tanto en primaria cuanto en secundaria, el rendimiento de los alumnos de escuelas de gestión no estatal es más alto que el de aquellos que estudian en escuelas estatales, independientemente del perfil actitudinal que se tenga. Por otro lado, se observan diferencias estadísticamente significativas en la distribución de porcentajes de estudiantes con perfiles actitudinales de escuelas estatales y no estatales. En éstas hay mayor porcentaje de estudiantes con actitudes favorables que en aquéllas.

En el caso del género, en primaria, cuando los perfiles de actitudes hacia la Matemática son favorables (perfiles C y D), los rendimientos de varones son mayores que los de las mujeres. En el perfil B (dificultad, desagrado, baja a media comprensión), de actitud desfavorable, también los varones tienen un rendimiento superior, aunque sólo por un punto. Es en el perfil A (temor) donde las niñas destacan, con un rendimiento 6 puntos superior al que alcanzan sus pares varones. En lo que respecta a la distribución de los porcentajes de estudiantes según los diferentes perfiles actitudinales, se observa que los patrones identificados para el rendimiento se mantienen. Hay diferencias estadísticamente significativas por género. Son los varones los que presentan actitudes más favorables hacia la Matemática, y las mujeres las que presentan actitudes desfavorables.

En secundaria se puede apreciar que el rendimiento de los varones es mayor que el de las mujeres para todos los perfiles de actitudes hacia la Matemática, sean favorables o no. Sin embargo, los porcentajes de estudiantes con perfiles de actitudes favorables y desfavorables hacia la Matemática no siguen el mismo patrón; así, por ejemplo, en el perfil C

**Cuadro 5**  
**Relación entre rendimiento en Matemática y las actitudes según gestión de la escuela y género del alumno en primaria<sup>20</sup>**

Actitudes hacia la Matemática	Gestión				Género			
	Estatal		No estatal		Masculino		Femenino	
	%	Media	%	Media	%	Media	%	Media
Perfil A	15	278	14	326	12	282	17	288
Perfil B	28	284	25	323	27	290	28	289
Perfil C	49	303	54	346	53	313	47	308
Perfil D	8	289	7	330	8	296	8	295
Nacional	100	293	100	336	100	302	100	298

**Cuadro 6**  
**Relación entre el rendimiento en Matemática y las actitudes según gestión de la escuela y género del alumno en secundaria**

Actitudes hacia la Matemática	Gestión				Género			
	Estatal		No estatal		Masculino		Femenino	
	%	Media	%	Media	%	Media	%	Media
Perfil A	15	281	18	323	16	292	15	288
Perfil B	34	278	23	308	27	285	37	279
Perfil C	41	297	51	348	49	310	38	305
Perfil D	10	291	8	327	8	301	10	292
Nacional	100	291	100	343	100	305	100	296

los varones alcanzan mejor rendimiento y son más, pero, en cambio, en el perfil B, pese a que los varones presentan un rendimiento más alto, son menos que las mujeres.

#### **4 DISCUSIEN**

El presente estudio halló que la mayor parte de los alumnos encuestados optó por las categorías de respuesta que denotan actitudes más favorables hacia la Matemática: comprensión percibida alta, percepción de competencia para el aprendizaje, gusto hacia esta materia y seguridad para participar en clase. Esto

parece indicar que, para los grados estudiados, la mayoría de los alumnos de CE urbanos del Perú presenta actitudes favorables hacia dicha materia, al menos en los cuatro aspectos aquí considerados. Estos resultados son semejantes a los obtenidos en NAEP para diferentes países, y a los del Paraguay según un es-

<sup>20</sup> En las comparaciones por gestión y género de las medias de rendimiento en cada perfil actitudinal se ha empleado la prueba t de Student. En las comparaciones por gestión y género de la distribución de los porcentajes de estudiantes que caen en los perfiles actitudinales se ha empleado la prueba Chi cuadrado.

tudio del SNEPE (Sistema Nacional de Evaluación del Proceso Educativo)<sup>21</sup>. Asimismo, al igual que en dichos estudios, en esta ocasión se encontró, en términos generales, que en los últimos años de la secundaria se incrementan los porcentajes de estudiantes con respuestas que denotan actitudes desfavorables. Algunos estudios en nuestro medio han encontrado que esta tendencia puede llegar a extenderse a los estudios superiores<sup>22</sup>.

En lo que se refiere a la relación del rendimiento con los cuatro aspectos encuestados, se ha encontrado que el nivel de comprensión de la clase, la percepción de competencia, el gusto y la percepción de seguridad para participar son muy relevantes en el rendimiento de los alumnos. Tal como se esperaba, las respuestas que denotan una actitud favorable se relacionan de manera directa con el nivel de logro académico alcanzado en la Matemática y las respuestas desfavorables lo hacen de manera inversa.

De los resultados se desprende, entonces, que la recepción de los contenidos y la comprensión de la información recibida a través de las explicaciones del profesor son fundamentales para el desempeño y aprendizaje de una materia, al menos de la Matemática. Asimismo, es importante que el estudiante se sienta a gusto y seguro frente a una situación de aprendizaje determinada. Esta seguridad se sostiene en el sentimiento de competencia para el aprendizaje y se manifiesta a través de una participación activa y libre de temores.

Es importante, sin embargo, considerar que si bien todos estos aspectos han mostrado una asociación significativa con el rendimiento, la magnitud del impacto de ellos sobre esta variable difiere. De esta manera se aprecia, por ejemplo, que la diferencia entre el grupo de estudiantes que manifiesta comprender y el que manifiesta no comprender es superior a la diferencia en rendimiento entre el grupo que expresa gusto por la Matemática y aquel que expresa desagrado.

Esto respaldaría la idea de que para lograr mejorar el rendimiento no sólo es fundamental contar con un clima de aula que reciba de forma apropiada las participaciones de los estudiantes y que los estimule mediante una metodología motivadora. Se requiere, principalmente, asegurar que los estudiantes se sientan competentes para aprender y, sobre todo, que entiendan los contenidos que se trabajan en el aula.

Por otra parte, es importante considerar que estos aspectos evaluados interactúan para configurar actitudes diversas. Por eso, el presente estudio ha identificado cuatro perfiles actitudinales hacia la Matemática basados en la interacción de estos aspectos. Estos perfiles, similares en primaria y secundaria, se pueden dividir en dos grupos: perfiles favorables (C y D) y perfiles desfavorables (A y B)<sup>23</sup>.

Utilizando estos perfiles se ha encontrado que, tanto en las escuelas estatales cuanto en las no estatales, los rendimientos medios se ordenan siguiendo el siguiente patrón: el grupo de mayor rendimiento es el del perfil D. Éste se caracteriza por manifestar gusto por la materia, percepción de comprensión muy alta y percepción de competencia para aprenderla. El grupo de rendimiento medio superior es el que responde al perfil C. Estos estudiantes expresan seguridad para participar activamente en clase. El grupo que presenta un rendimiento medio inferior es el del perfil B, cuyos estudiantes perciben dificultad para aprender la Matemática y sienten desagrado por las clases de este curso. Por último, el grupo de rendimiento más bajo es el que responde al perfil A. Los estudiantes que se ubican aquí manifiestan temor o falta de seguridad para participar en clase.

Asimismo, con relación a la gestión de la escuela se observa que si se comparan los ren-

21 SNEPE: *Informe de resultados: Tercer grado*. Asunción: Ministerio de Educación y Cultura del Paraguay, 1998.

22 Aparicio, A. y J. Bazán: "Actitudes hacia la Matemática en ingresantes a la UNALM", en "Niñez y adolescencia en el Perú". *Más Luz: Revista de Psicología y Pedagogía*, 3 (2), 1997, pp. 351-380. Número monográfico.

23 Entre otros aspectos, la construcción de estos perfiles como indicador de actitudes queda avalada por el hecho de que los perfiles favorables se relacionan en general con rendimientos más altos en Matemática que en Lenguaje. Por el contrario, los perfiles de actitudes desfavorables hacia la Matemática se relacionan con rendimientos más bajos en Matemática y más altos en Lenguaje. No obstante, cabe señalar que estas asociaciones difieren poco entre sí.

dimientos medios de los estudiantes de escuelas estatales y no estatales, estos últimos obtienen siempre resultados superiores, lo que se cumple para cada tipo de perfil actitudinal. También se ha encontrado que el porcentaje de estudiantes que tienen perfiles actitudinales favorables es mayor en las escuelas no estatales que en las estatales.

Con relación a la variable género se encontró que para los dos niveles evaluados (primaria y secundaria), si se comparan los rendimientos medios de cada perfil, los estudiantes varones obtienen mejores resultados que sus pares mujeres, con la única excepción del perfil A (desfavorable) en primaria, donde los niños han obtenido rendimientos inferiores que los de las niñas. También se observa que el porcentaje de varones con perfiles actitudinales favorables es superior al de sus pares mujeres.

Si bien los resultados encontrados con el indicador de actitud son similares a los hallados en otros estudios que emplean escalas de actitudes<sup>24</sup>, el indicador que se ha desarrollado en este estudio no debe sustituir al uso de una escala de actitudes.

Los resultados encontrados en este trabajo sugieren la importancia de la dimensión afectiva del aprendizaje sobre el rendimiento académico que logran los estudiantes. Así, las actitudes, percepciones y temores constituyen aspectos que deben ser considerados y estudiados con mayor profundidad si se aspira a alcanzar una educación de calidad. Replantear el papel del profesor y considerar esta dimensión desde la propia estructura curricular son algunas de las medidas que deben ser tomadas en cuenta. En consecuencia, el reto consiste en formar docentes que sean capaces de estimular en sus estudiantes el gusto por la materia, que puedan hacer del aula un espacio donde la participación sea alentada y, sobre todo, que contribuyan a que sus estudiantes se sientan competentes para aprender.

---

24 Por ejemplo, Rodríguez Feijoo, N.: "Estudio de las actitudes hacia la Matemática y la Estadística", en *Revista Intercontinental de Psicología y Educación*, 4, 1991, pp. 369-383; Aliaga, J. y J. Pecho: "Evaluación de la actitud hacia la Matemática en estudiantes secundarios", en *Paradigmas*, 1 (1-2), 2000, pp. 61-78.

# ANEXO

## CONSTRUCCIÓN DEL INDICADOR DE ACTITUD

El análisis de correspondencias múltiples\* tiene como objetivo obtener los pesos de las preguntas en las dimensiones componentes para identificarlas adecuadamente, así como las coordenadas de las categorías que permiten caracterizar los grupos de alumnos sobre la base de su cercanía.

Para los propósitos de este trabajo se considera relevante ofrecer de manera independiente los análisis referidos al indicador de actitud para el cuarto grado de educación primaria y para el quinto de secundaria. La razón de esta selección es que tanto el sexto grado de primaria cuanto el cuarto de secundaria presentan perfiles muy similares a los de su correspondiente nivel.

El análisis realizado para primaria y secundaria revela —como se observa en los gráficos 1 y 2— dos dimensiones. Los porcentajes de varianza explicada por cada análisis de correspondencia son del 66% y 88% para primaria y secundaria respectivamente.

Como se desprende del gráfico 1, destaca un grupo de estudiantes que manifiesta temor (grupo A), en tanto que un grupo contrario (está al extremo izquierdo de éste) lo forman, por un lado, aquellos que se perciben competentes para aprender Matemática, presentan alta comprensión y les gusta el curso (grupo D) y, por otro, los alumnos que no manifiestan temor (grupo C). El cuadrante del grupo B no está tan claramente definido como los otros, ya que los puntos se encuentran más dispersos y se ubican a lo largo del eje positivo.

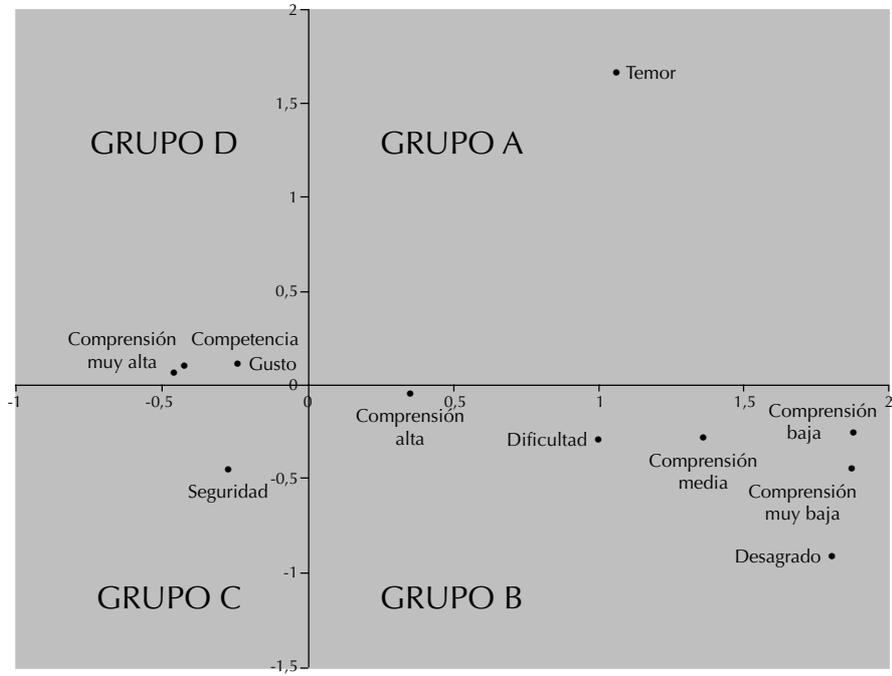
En el caso de secundaria se observa que el grupo A es el más claramente definido, y está formado por aquellos que manifiestan temor de preguntar en clase. El grupo B presenta una tendencia a concentrarse hacia la derecha y está formado por aquellos alumnos ubicados en el rango de comprensión de muy baja a media, considera que el curso es difícil y además le desagrada. El grupo C es ambiguo porque, como se puede observar, un punto se ubica en el extremo superior del eje vertical, en tanto que hay otro punto que se encuentra lejos de éste, casi al extremo izquierdo del eje horizontal; es decir, no existe una concentración de características. El eje vertical es fácil de nombrar, pues expresa dos sentimientos extremos: la seguridad (grupo C) en la puntuación positiva, y el temor (grupo A) en la puntuación negativa del eje.

Un análisis de las respuestas a los diferentes aspectos para ambos grados revela que aunque rotados en primaria y secundaria, los grupos C y D son los que presentan actitudes más favorables hacia la Matemática, mientras que los grupos A y B, rotados también en primaria y secundaria, son los que manifiestan actitudes menos favorables.

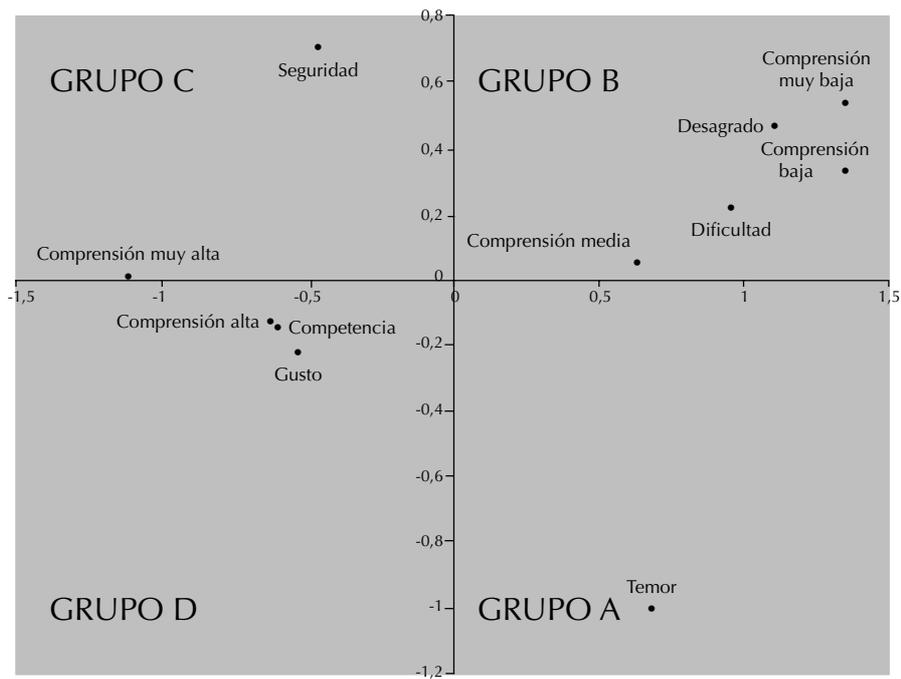
---

\* El análisis de correspondencias múltiples es una técnica exploratoria que determina la representación multidimensional de la asociación entre las categorías expresadas en una tabla de contingencia. Esta técnica convierte las categorías de la tabla de contingencia (filas y columnas) en un menor número de dimensiones, indicando qué porcentaje del valor de la asociación puede ser explicado por éstas. Al respecto puede revisarse la siguiente dirección URL: [http://www3.uniovi.es/~Psi/Dpto\\_Psicologia/metodos/hardcopy/98\\_1.html](http://www3.uniovi.es/~Psi/Dpto_Psicologia/metodos/hardcopy/98_1.html)

**Gráfico 1**  
**Primaria: Coordenadas de las categorías de respuesta hacia la Matemática sobre sus dimensiones componentes**



**Gráfico 2**  
**Secundaria: Coordenadas de las categorías de respuesta hacia la Matemática sobre sus dimensiones componentes**



Estos resultados se muestran en el subacápite “Procedimiento” del acápite “Método”.

Con este análisis hemos podido representar once categorías sobre dos dimensiones. Si usamos los cuadrantes que se forman, estas dos dimensiones subyacentes determinan cuatro perfiles. Para asignar un perfil a un alumno se han usado sus puntajes en las dos dimensiones, sean éstos positivos o negativos. Así, un perfil es (+) (+), otro es (+) (-), un tercero es (-) (+) y un cuarto es (-) (-). Para interpretar estos perfiles se ha empleado la ubicación de las categorías en los cuadrantes. Esto significa que la asignación no es excluyente, porque no se basa en las respuestas originales de los alumnos sino en su ubicación en el nuevo sistema de coordenadas dado por las dimensiones subyacentes. En esto, el análisis de correspondencias múltiples es muy similar al análisis factorial y supone una pérdida de información que se justifica por una ganancia conceptual dada por la reducción de categoría en cuatro perfiles.

# ¿QUÉ PIENSAN LOS DOCENTES DE SUS ALUMNOS?

FERNANDO ANDRADE\*

En 1998 se administraron una serie de pruebas y encuestas como parte del proyecto CRECER (Crecer con Calidad y Equidad en la Educación y el Rendimiento) de la Unidad de Medición de Calidad Educativa (UMC) del Ministerio de Educación (MED). Las encuestas estuvieron dirigidas a conocer algunas características de importantes actores en el sistema educativo (estudiantes, padres, docentes y directores), de modo que luego éstas pudieran ser analizadas tanto en sus interrelaciones como en sus asociaciones con los resultados de las pruebas de rendimiento. El objetivo final era generar información útil para conocer mejor el sistema y evaluar sus logros para estar en capacidad de sugerir políticas educativas.

Esta sección del documento aborda el tema de las percepciones de los docentes respecto de sus alumnos en lo que concierne a la identificación de factores que estarían limitando su rendimiento escolar. Según investigaciones realizadas en diferentes partes del mundo, uno de los aspectos más importantes e influyentes en el rendimiento de los alumnos es la percepción del docente sobre ellos; en otras palabras, lo que piensa y espera el docente de sus alumnos determina la manera como él o ella va a tratar a sus estudiantes, tanto individual como grupalmente, y esto puede redundar de manera directa en el rendimiento de sus alumnos<sup>1</sup>.

En este contexto se preguntó a los docentes qué opinión tenían de sus estudiantes en dos grandes aspectos: el primero se refiere a

la apreciación del docente sobre la capacidad y esfuerzo de sus alumnos; el segundo, a las dificultades que el docente piensa que tiene la mayoría de sus alumnos<sup>2</sup>. Éstas se han clasificado en cuatro temas<sup>3</sup>. El primero ata-

---

\* El autor, quien trabaja en el Grupo de Análisis para el Desarrollo (GRADE), desea expresar su agradecimiento a Santiago Cueto y a José Rodríguez por sus comentarios y sugerencias. Este artículo ha sido elaborado en forma conjunta con la UMC.

- 1 Tauber, Robert T.: "Good or Bad, What Teachers Expect from Students They Generally Get!". Eric Digests. Web ED426985, 1998; Lumsden, Linda: "Expectations for Students". Eric Digests. Web ED409609, 1997; Snow, Richard; Lyn Corno y Douglas Jakobson III: "Individual Differences in Affective and Conative Functions", en David Berliner y Robert C. Calfee, editores: *Handbook of Educational Psychology*. New York: Simon and Schusters Macmillan, 1976, pp. 243-310; Shaffer, David: *Social and Personality Development*. California: Wadsworth, 1994.
- 2 Cabe destacar que estos problemas no son necesariamente una explicación de la percepción de la capacidad y esfuerzo que el profesor tiene sobre sus alumnos.
- 3 Los problemas que los profesores observan en la mayoría de sus alumnos se reunieron primero por criterios conceptuales, es decir, agrupando problemas que teóricamente deberían estar relacionados. Es interesante notar que un análisis factorial de estos problemas arrojó prácticamente los mismos grupos que habían sido definidos conceptualmente (el análisis factorial es una técnica estadística de agrupación de variables basada en la correlación que muestran una a una, hasta lograr conjuntos, cada uno de los cuales es denominado "factor").

ñe a problemas relacionados con el estudio (poco interés, malos hábitos, escasa concentración, poca comprensión de lo que leen y poca percepción de la utilidad de lo que se estudia). El segundo tema es el relativo a la asistencia a la escuela y las tardanzas. El tercero se refiere al poco tiempo que el docente cree que sus alumnos tienen para estudiar debido a que trabajan. Finalmente, el cuarto tema concierne a la influencia de los problemas familiares en el rendimiento escolar.

Para el presente estudio se diseñó una muestra de modo que los resultados fueran representativos de los estudiantes de centros educativos (CE) polidocentes completos estatales y no estatales de las zonas urbanas del país. Se incluyeron en la muestra estudiantes de cuarto y sexto de primaria y cuarto y quinto de secundaria. Sin embargo, dado que los resultados de los dos grados de un mismo nivel son muy similares, en esta sección se incluyen sólo los resultados de sexto y quinto (último grado de los respectivos niveles).

Finalmente, se debe notar que en todos los casos se presentan datos para CE estatales y no estatales. La responsabilidad principal del MED es para con los estudiantes del sistema estatal (que constituyen más del 80% del estudiantado en cualquier grado de primaria o secundaria); sin embargo, a menudo resulta ilustrativo tener a los del sistema privado como un referente.

En términos generales, se han encontrado diferencias entre los CE estatales y no estatales sobre la percepción que tienen los docentes de sus alumnos, y se ha concluido que los docentes de CE estatales encuentran problemas en sus estudiantes en mayor proporción que sus colegas de los no estatales. Estas diferencias podrían obedecer a factores vinculados a la forma en que se trabaja en los CE estatales y no estatales (incluyen-

do infraestructura, características de los docentes, programación curricular, etcétera); a las diferencias en la población de estudiantes que atienden (por ejemplo, en cuanto al nivel educativo de los padres del sistema estatal y privado, diferencias culturales y lingüísticas, etcétera), o a la combinación de estos dos factores. Al analizar las diferencias que se presentan en estos reportes se debe tener esto en consideración, de modo que no se llegue a conclusiones simplistas y rápidas sobre la calidad de los estudiantes o del sistema educativo estatal.

Para poner énfasis en lo mencionado, la primera parte de esta sección proporciona una breve descripción de los alumnos que asisten a los CE estatales y no estatales. Esta descripción está basada en algunas de las características recogidas en las encuestas CRECER 1998<sup>4</sup>.

En la segunda parte se presentan los resultados de la percepción de los docentes sobre sus alumnos y los problemas que aquéllos observan en la mayoría de éstos.

Finalmente, en un tercer acápite exponemos las conclusiones del estudio y las explicaciones técnicas de la muestra.

## **1 PADRES Y ESTUDIANTES DE LOS CE ESTATALES Y NO ESTATALES**

De las muchas características recogidas en las encuestas CRECER 1998, se han seleccionado algunas que tipifican a los padres y estudiantes de los CE no estatales y estatales.

Estas características se refieren al trabajo del alumno, los materiales que éste tiene en casa para estudiar (libros y computadora), el grado de instrucción de los padres y los servicios de agua y desagüe en el hogar.

### **¿QUÉ PORCENTAJE DE ALUMNOS TRABAJA?**

En el Perú existe un porcentaje de niños que tiene que trabajar además de estudiar. La pregunta específica que se hizo en las encuestas dirigidas a los alumnos es: "¿Qué trabajo realizas?"<sup>5</sup>. En el presente análisis se ha con-

4 Se reportan aquí datos de las encuestas a padres de familia y alumnos.

5 Las alternativas de respuesta para esta pregunta eran: a) Trabajar en la chacra o cuidar ganado; b) Trabajar en la calle (vender caramelos, periódicos, lustrar zapatos, cuidar carros, etcétera); c) Trabajar en un negocio (ayudar en el puesto de un mercado, atender en la bodega, hacer artesanías, etcétera); d) Trabajar en un hogar particular como emplea-

siderado que un alumno trabaja cuando realiza una o más de las siguientes actividades<sup>6</sup>: trabaja en la chacra o cuidando ganado; trabaja en la calle; trabaja en un negocio; trabaja en un hogar<sup>7</sup>; realiza otras actividades laborales. No se ha considerado como trabajo la ayuda que el alumno puede brindar en los quehaceres domésticos de su propio hogar.

Para el caso de los alumnos de sexto de primaria se ha encontrado que en los CE estatales el porcentaje que trabaja es relativamente alto, más de la cuarta parte, mientras que en los CE no estatales lo hace menos de la quinta parte (véase el cuadro 1). Para el

Cuadro 1 6° grado de educación primaria		
Tipo de gestión	Público %	Privado %
El alumno trabaja	29,9	15,3
<b>N° total</b>	<b>13 545</b>	<b>2712</b>
<b>Materiales en casa para hacer tareas</b>		
Libro de Matemática	71,7	77,2
Libro de Lenguaje	74,7	76,1
Libro de CC.HH.SS.	66,7	71,2
Libro de CC.NN.	65,7	72,6
Computadora	11,0	27,2
<b>N° total</b>	<b>13 734</b>	<b>2762</b>

caso de quinto de secundaria, el porcentaje sube notablemente con respecto a primaria en los CE estatales, mientras que sólo existe un ligero aumento en los no estatales (véase el cuadro 2).

### ¿TIENEN MATERIALES EN CASA PARA HACER TAREAS?

Los materiales que el alumno puede necesitar para hacer sus tareas en casa son diversos; por ejemplo, útiles de escritorio, un escritorio o mesa, cuadernos de trabajo, libros, diccionarios, entre otros. En las encuestas

Cuadro 2 5° grado de educación secundaria		
Tipo de gestión	Público %	Privado %
El alumno trabaja	43,7	19,3
<b>N° total</b>	<b>12 422</b>	<b>3652</b>
<b>Materiales en casa para hacer tareas</b>		
Libro de Matemática	60,1	72,1
Libro de Lengua y Literatura	74,9	82,0
Computadora	7,0	35,4
<b>N° total</b>	<b>12 507</b>	<b>3666</b>
<b>Libros en casa</b>		
25 libros o menos	65,7	30,0
26 libros o más	34,3	60,0
<b>N° total</b>	<b>12 507</b>	<b>3675</b>

CRECER 1998 se preguntó específicamente: "¿Qué materiales usas en tu casa para hacer tus tareas?", y algunas de las alternativas fueron: libros de Matemática, de Lenguaje y computadora. En el caso de sexto de primaria se consideraron adicionalmente los libros de Ciencias Histórico-Sociales y Ciencias Naturales<sup>8</sup>.

En cuanto a la tenencia de libros, el porcentaje que cuenta con éstos es mayor en general entre estudiantes de CE no estatales. Sin embargo, se podría decir que las diferencias son pequeñas en primaria (menos del 10%) y que aumentan ligeramente en secundaria.

da/o; e) Ayudar en las actividades de la casa (cuidar a los hermanos, cocinar, etcétera); f) Otro. Las alternativas fueron las mismas para sexto de primaria y quinto de secundaria.

6 Remuneradas o no.

7 Como empleado o empleada domésticos.

8 Puede sorprender que no se hayan utilizado términos del nuevo currículo para las áreas curriculares (por ejemplo, *Comunicación Integral* en vez de *Lenguaje*). Téngase en cuenta que al momento de administrar la encuesta (fines de 1998), el nuevo currículo aún no se había generalizado en sexto de primaria.

Queda claro que la mayoría de estudiantes cuenta con textos escolares. (No obstante, el porcentaje debería ser 100%, es decir, todos los alumnos deberían contar con dichos textos; seguramente esta situación ha cambiado en los años 1999 y 2000, cuando se repartieron masivamente textos escolares en CE estatales de primaria.)

En lo que respecta a la tenencia de computadora, las diferencias entre los dos tipos de CE son notorias (véase el cuadro 2). Cabe señalar que para el caso de los CE no estatales el porcentaje de tenencia de computadora aumenta de sexto de primaria a quinto de secundaria. Es probable que la computadora sea percibida como más necesaria para los estudios en secundaria que en primaria.

Sólo para el caso de secundaria se preguntó: "Aproximadamente, ¿cuántos libros hay en tu casa? (No cuentes los libros del colegio, revistas ni periódicos)". Para efectos de este documento se han agrupado las alternativas de respuesta en dos grandes categorías<sup>9</sup>: 25 libros o menos y 26 libros o más. Así, se encuentra que en los CE estatales el porcentaje es mayor en la primera categoría, mientras que en los CE no estatales ocurre lo contrario. En otras palabras, los estudiantes de CE no estatales viven en ambientes con más libros. Esto forma parte de lo que algunos autores denominan "capital cultural de la familia".

9 Los rangos de las respuestas son: de 0 a 10 libros, de 11 a 25 libros, de 26 a 50 libros, más de 50 libros.

10 Sin instrucción; primaria incompleta; primaria completa; secundaria incompleta; secundaria completa; superior técnica incompleta; superior técnica completa; superior universitaria incompleta; superior completa.

11 Se recogió información sobre materiales del techo, piso, paredes y alumbrado de la vivienda; si se disponía en el hogar de agua y desagüe, entre otros temas.

12 En cuanto a esto, se preguntó por el tipo de abastecimiento de agua de la vivienda. Las categorías de respuesta fueron: instalación de red pública dentro de la vivienda; pilón de uso público; pozo; camión cisterna u otro; río, acequia, manantial; y otro.

13 Se preguntó si la vivienda tenía servicio de desagüe, con alternativas de respuesta "sí" y "no".

## ¿CUÁL ES EL NIVEL EDUCATIVO DE LOS PADRES?

En las encuestas aplicadas a los padres de familia CRECER 1998 se les preguntó por su grado de educación. Las preguntas que recogían esta información eran, para el caso de la madre, "¿Cuál es el máximo nivel educativo alcanzado por la madre del alumno/a?". Para el caso del padre esta pregunta era: "¿Cuál es el máximo nivel educativo alcanzado por el padre del alumno/a?".

De todas las categorías sobre las que se preguntó<sup>10</sup>, para abreviar los cuadros se han seleccionado las correspondientes a las de estudios superiores incompletos y completos combinados. En los cuadros 3 y 4 se nota claramente que los logros educativos de los padres de estudiantes de CE no estatales son mayores que los de sus pares estatales.

## ¿TIENEN AGUA Y DESAGÜE EN EL HOGAR?

En las encuestas CRECER 1998 dirigidas a los padres de familia se recogió información sobre la vivienda de los alumnos<sup>11</sup>. Para el presente estudio se decidió trabajar con dos de estos aspectos, que son el abastecimiento de agua<sup>12</sup> y desagüe<sup>13</sup> en el hogar. En el caso del

**Cuadro 3**  
**6° grado de educación primaria**

Tipo de gestión	Público %	Privado %
<b>Padres: Estudios superiores completos o incompletos</b>		
Padre	23,0	58,0
Madre	17,3	52,2
<b>N° total</b>	<b>9372</b>	<b>1884</b>
<b>Servicio de agua y desagüe en viviendas</b>		
Abastecimiento de agua		
por red pública	73,4	91,2
Servicio de desagüe	67,7	90,8
<b>N° total</b>	<b>9410</b>	<b>1887</b>

Cuadro 4 5° grado de educación secundaria		
Tipo de gestión	Público %	Privado %
<b>Padres: Estudios superiores completos o incompletos</b>		
Padre	23,9	65,5
Madre	15,6	54,1
<b>N° total</b>	<b>8685</b>	<b>2539</b>
<b>Servicio de agua y desagüe en viviendas</b>		
Abastecimiento de agua por red pública	78,0	94,6
Servicio de desagüe	73,1	95,2
<b>N° total</b>	<b>8706</b>	<b>2553</b>

abastecimiento de agua se escogió trabajar sólo con el porcentaje de viviendas que cuentan con servicio de red pública dentro de la vivienda. Las preguntas utilizadas fueron: “¿Cómo se abastece de agua la vivienda donde vive el alumno?”, y “¿La vivienda donde vive el alumno tiene servicio de desagüe?”.

Los alumnos de los CE no estatales acceden en mayor proporción que los alumnos de los estatales a los servicios de agua y desagüe (véase los cuadros 3 y 4).

Los datos presentados en dichos cuadros muestran notables diferencias —que podríamos generalizar como “socioeconómicas”— entre los estudiantes que asisten a CE estatales y no estatales. Estas diferencias serán consideradas cuando se analicen los datos que reportan los docentes sobre la mayoría de sus alumnos, que son presentados en el siguiente acápite.

## 2 PERCEPCIÓN DE LOS DOCENTES SOBRE SUS ALUMNOS

En la base de la relación que el docente establece con sus alumnos se encuentran las creencias u opiniones que el primero tiene sobre los segundos. Uno de los estudios clásicos sobre este tema es el trabajo de Rosenthal y Jacobson (1968)<sup>14</sup>. Estos autores deno-

minaron “Pigmalión en el aula” al efecto de las expectativas que los docentes tienen sobre el rendimiento de sus alumnos. Según Tauber<sup>15</sup>, la importancia de estos estudios está en que al conocer el mecanismo del “efecto Pigmalión” o “profecía autocumplida” se puede utilizar este efecto como una herramienta pedagógica para generar expectativas positivas y, más importante aún, evitar expectativas negativas en los docentes. Las expectativas se forman sobre la base de las creencias acerca de las características físicas del alumno, género, raza, etnia, nombre o apellido, atractivo y nivel socioeconómico, entre otras.

El efecto Pigmalión o profecía autocumplida consiste en lo siguiente: el docente genera expectativas sustentadas en creencias u opiniones que tiene sobre sus alumnos. Luego, basado en estas expectativas, actúa de determinada manera diferenciando su trato. Luego, la manera como el docente trata a cada alumno le dice a éste qué comportamiento y qué rendimiento espera de él o ella. Si este trato del maestro es consistente, entonces tenderá a moldear el comportamiento y rendimiento del alumno. Lo que el estudio encontró es que, con el tiempo, el comportamiento del alumno y su rendimiento se acercarán mucho a las expectativas del docente, reafirmando de este modo sus creencias iniciales<sup>16</sup>.

Como ya se mencionó, la percepción del docente sobre sus alumnos se ha dividido en dos temas: el primero es la percepción del docente sobre la capacidad y esfuerzo de sus alumnos; el segundo trata de los problemas que el docente percibe en la mayoría de sus alumnos. A continuación se revisarán estos dos temas y los resultados respectivos.

### ¿LOS ALUMNOS SE ESFUERZAN? ¿TIENEN CAPACIDAD PARA APRENDER?

Se le preguntó al docente: “¿Qué afirmación refleja mejor la opinión que tiene respecto de

14 Citados en Tauber, R.: “Good or Bad...”, *op. cit.*, 1998.

15 *Ibid.*

16 *Ibid.*

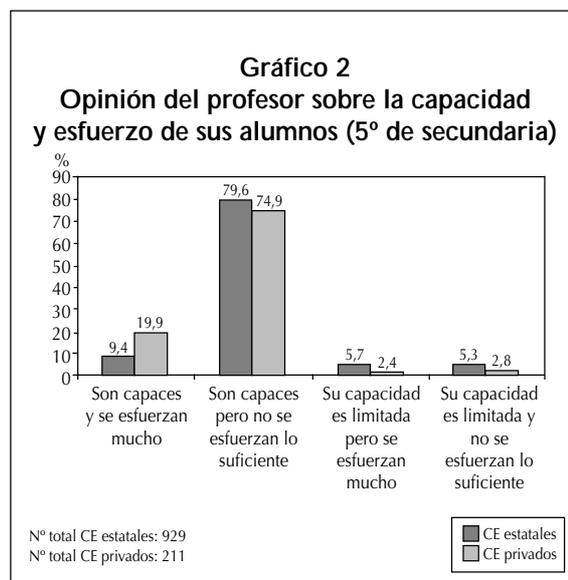
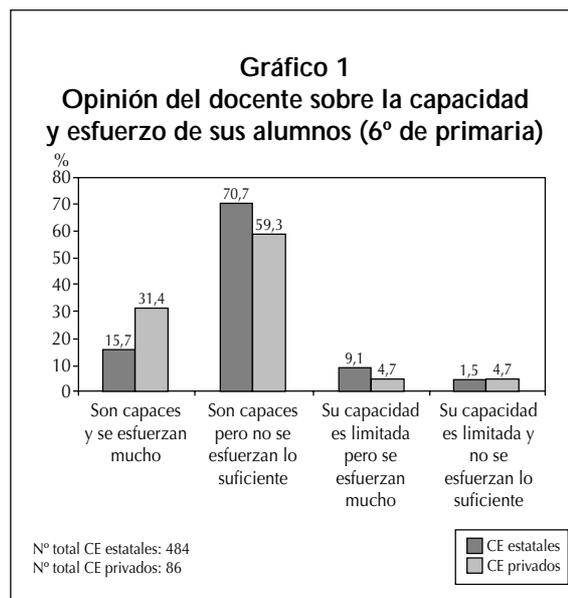
la mayoría de los alumnos a su cargo?"<sup>17</sup>. Las categorías de respuestas fueron:

- Son capaces y se esfuerzan mucho.
- Son capaces pero no se esfuerzan lo suficiente.
- Su capacidad es limitada pero se esfuerzan mucho.
- Su capacidad es limitada y no se esfuerzan lo suficiente.

Se encontró, tanto en primaria como en secundaria y en los dos tipos de CE, que la mayoría de docentes piensa que sus alumnos "son capaces pero no se esfuerzan lo suficiente". Esta alternativa sugiere que existe la posibilidad de lograr, a través de diferentes estrategias pedagógicas, que los alumnos estén más motivados para incrementar su esfuerzo<sup>18</sup>.

En términos ideales, se esperaría que la mayoría de los docentes se refieran a sus alumnos como "son capaces y se esfuerzan mucho", pero esto sucede sólo parcialmente, en especial con los docentes de sexto grado de los CE no estatales.

Más crítico es el caso de las categorías "Su capacidad es limitada pero se esfuerzan mucho" y "Su capacidad es limitada y no se esfuerzan lo suficiente". En ambos casos las expectativas de los docentes con respecto al comportamiento y rendimiento de sus alumnos tienden a ser negativas y con poca esperanza de cambio. Afortunadamente, el porcentaje de docentes en estas categorías es relativamente bajo (véase los gráficos 1 y 2).



17 Encuestas para docentes de cuarto y quinto de secundaria CRECER 1998 y encuestas CRECER 1998 para responsables de aula de cuarto y sexto de primaria.

18 Es obvio que no basta que los docentes se formen expectativas positivas sobre el nivel de esfuerzo de sus alumnos, ya que existen otros factores que pueden interferir en la exigencia que éstos les planteen (por ejemplo, que el alumno trabaje); así, gran parte de los alumnos de los CE estatales (29,9% en el caso de sexto de primaria y 43,7% en quinto de secundaria) trabaja, lo que seguramente les resta tiempo y energías para el estudio.

19 Citado en Michener, Andrew H. et al.: *Social Psychology*. Washington D.C.: Harcourt Brace Janovich, 1990.

Como se ha visto, la mayoría de docentes se refieren a sus alumnos como capaces y que no se esfuerzan lo suficiente. Para Weiner<sup>19</sup>, las percepciones de cada individuo sobre el éxito o fracaso se pueden explicar en función de dos dimensiones: el grado de estabilidad y el *locus* de control o atribución interna-externa. La dimensión de grado de estabilidad tiene dos categorías: estable e inestable, y según esta división el rasgo inestable es más susceptible al cambio que el rasgo estable. La dimensión de control o atribución interna-externa se refiere al tipo de control que una persona tiene sobre cualquier situación. Así, si el

control o la atribución es interno significa que la situación depende del individuo; y, por el contrario, si la atribución o el control es externo, la situación depende del entorno.

Entonces, desde el punto de vista de Weiner, la capacidad es a menudo percibida como un rasgo estable y, en cambio, el esfuerzo es visto como un rasgo inestable. Ambas características son de control interno, es decir, dependen del individuo. Que el esfuerzo sea un rasgo inestable quiere decir que puede ser más fácilmente modificable que la capacidad<sup>20</sup>.

En términos de política educativa, a partir de los resultados se podría pensar en programas dirigidos a incentivar el esfuerzo (o algo similar: por ejemplo, la motivación) que los alumnos dediquen a sus estudios. Estos programas deberían estar basados en las causas de este poco esfuerzo que perciben los docentes. Algunas de estas causas podrían estar reflejadas en los problemas que los docentes perciben en la mayoría de sus alumnos. Estos datos se presentan en el siguiente acápite.

### **¿CUÁLES SON LOS PROBLEMAS DE LA MAYORÍA DE LOS ALUMNOS?**

Las preguntas planteadas a los docentes fueron: “¿Qué problemas o dificultades afectan a *más de la mitad* de los alumnos del aula donde usted es el profesor responsable?”; debían responder “Sí” o “No” a cada problema. Las respuestas fueron agrupadas en cuatro temas.

Un primer tema está relacionado con el estudio y abarca los siguientes problemas: a) poco interés por los estudios; b) malos hábitos de estudio; c) poca concentración para atender la clase; d) escasos hábitos de lectura; e) dificultad para comprender lo que leen (los anteriores fueron comunes para quinto de secundaria y sexto de primaria); y, f) dificultades para encontrarle sentido práctico a lo que estudian (sólo para quinto de secundaria).

El segundo tema se refiere a la asistencia a clases, y tiene dos categorías: a) asistencia irregular; y, b) tardanzas.

El tercer tema atañe al poco tiempo que el docente cree que sus alumnos tienen para estudiar debido a que trabajan.

Finalmente, el cuarto tema es el relativo a los problemas en la familia del alumno. En este tema existen tres categorías: a) familia con problemas económicos; b) conflictos o peleas familiares; y, c) poco apoyo de los padres con las tareas.

### **¿Problemas con los estudios?**

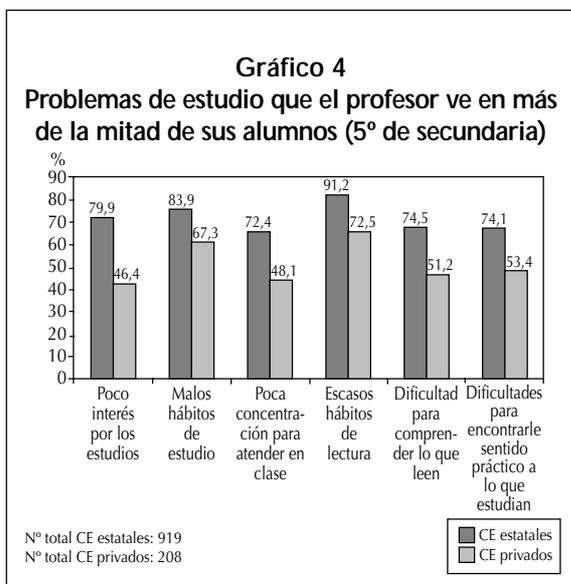
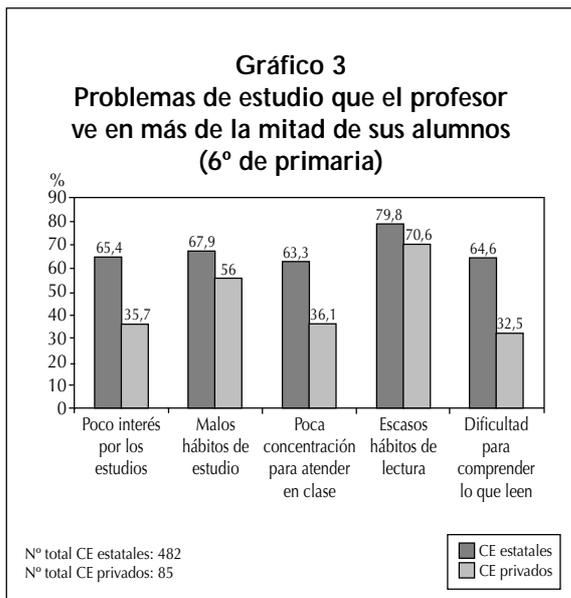
Para el caso de sexto de primaria, el problema de estudio más observado por los docentes de CE estatales y no estatales es el escaso hábito de lectura (véase el gráfico 3).

El segundo problema, en sexto de primaria, en orden decreciente de porcentajes, es que los alumnos tienen malos hábitos de estudio (67,9% y 56% en CE estatales y no estatales respectivamente). Otros de los problemas reportados para este grado son: poco interés por el estudio, poca concentración en la clase y dificultad para comprender lo que leen. Todos son reportados por más de la mitad de los docentes de CE estatales y menos de la mitad en CE no estatales (véase el gráfico 3).

Para el caso de quinto de secundaria aumenta el porcentaje de los docentes que observan problemas de estudio en sus alumnos con relación a sexto de primaria. Los escasos hábitos de lectura (91,2% y 72,5% en CE estatales y no estatales respectivamente) y los malos hábitos de estudio (83,9% y 67,3% para CE estatales y no estatales respectivamente), al igual que en sexto de primaria, son los dos problemas que la mayoría de los docentes han observado en sus alumnos (véase el gráfico 4).

Como en sexto de primaria, el poco interés por el estudio, la poca concentración en la clase y la dificultad para comprender lo que leen son los otros problemas que el

20 No pretendemos discutir teóricamente si la capacidad es un rasgo estable o inestable, y entendemos que muchas personas podrían argumentar por una posición o la otra. Sólo queremos destacar que, de acuerdo con Weiner, a menudo la capacidad es percibida por los individuos como un rasgo que difícilmente puede ser cambiado, mientras que el esfuerzo es percibido como relativamente más fácil de cambiar.



docente refiere sobre sus alumnos (véase el gráfico 4).

Sólo en quinto de secundaria se preguntó por la dificultad que los estudiantes tienen para encontrarle sentido práctico a las cosas que estudian; en ambos tipos de CE, más de la mitad de los docentes reportan que éste es un problema que observan en la mayoría de los alumnos (véase el gráfico 4).

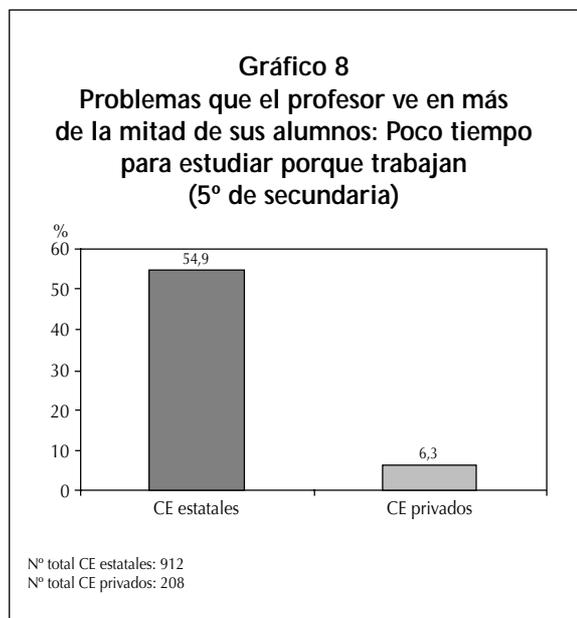
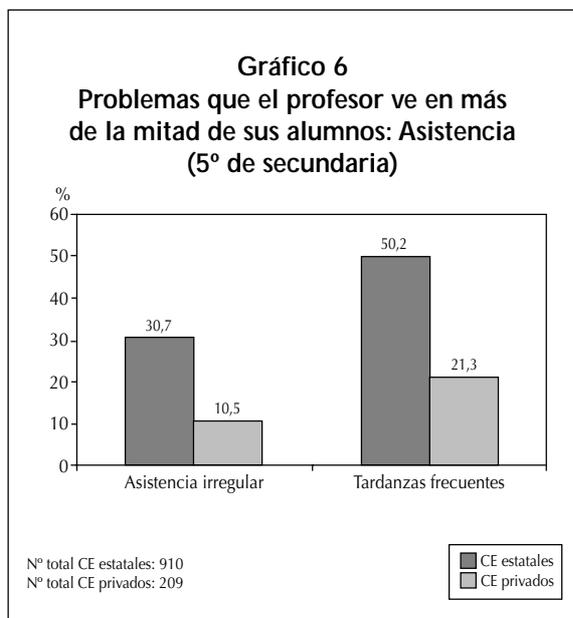
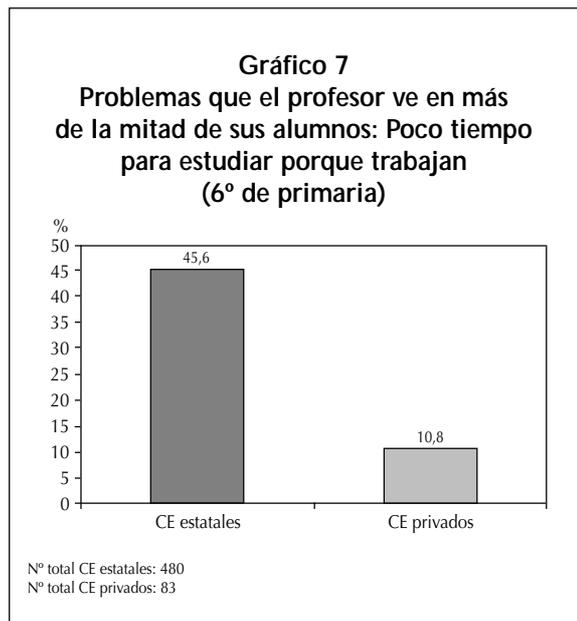
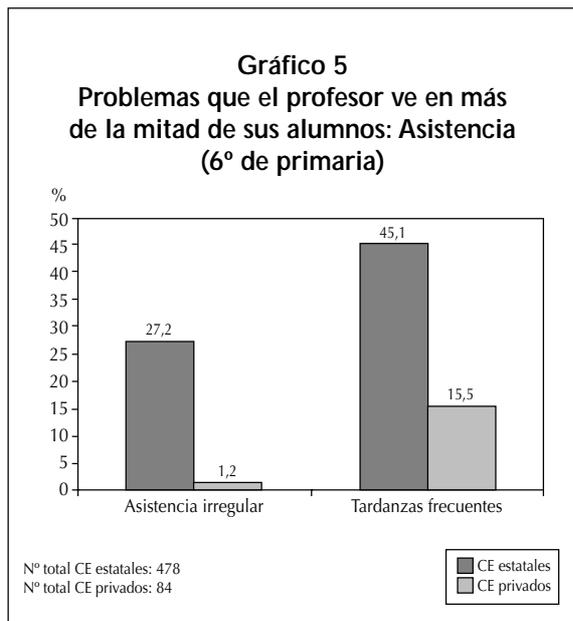
Los datos anteriores podrían ser tomados como un diagnóstico de los docentes sobre algunas carencias de sus estudiantes, y muestran que existen muchos problemas bastante

generalizados, sobre todo en CE estatales y en secundaria. Esto podría ser útil para identificar puntos específicos sobre los cuales el sistema educativo podría actuar para mejorar el aprendizaje de los estudiantes. Afortunadamente, usando la terminología de Weiner mencionada antes, se trata de rasgos inestables, es decir, que podrían ser modificados con relativa facilidad.

### ¿Problemas en la asistencia a clases?

Otro de los aspectos importantes relacionados con la educación escolar es la asistencia a clases. En general, si un alumno no está en clase pierde oportunidad de aprender y se pone en desventaja frente a aquellos que sí asisten. En las encuestas CRECER 1998 se preguntó al docente sobre dos aspectos: asistencia y tardanza. La pregunta específica fue: "¿Qué problemas o dificultades afectan a *más de la mitad* de los alumnos del aula donde usted es el docente responsable?". Las alternativas utilizadas fueron: "Asistencia irregular de los alumnos al centro educativo" y "Tardanzas frecuentes de los alumnos".

- a) Asistencia irregular: En el caso de los CE estatales, el porcentaje de docentes que reporta asistencia irregular en sus estudiantes es casi la tercera parte del total. El porcentaje aumenta para quinto de secundaria, donde nuevamente la asistencia irregular aparece como un problema relativamente mayor. Ello es consistente con el hecho de que hay un alto porcentaje de estudiantes que trabaja, especialmente en los CE estatales y quinto de secundaria. Esto permite especular respecto de que posiblemente la inasistencia irregular se deba, entre otros factores, a que los alumnos tengan que compartir parte de su tiempo de estudio con el trabajo.
- b) Tardanzas: Los resultados muestran un patrón parecido al de asistencia, pero con porcentajes mayores. Para el caso de CE estatales, el porcentaje de docentes que refiere tardanzas en sus alumnos llega casi a la mitad del total en sexto de primaria y al 50% en quinto de secundaria (véase los gráficos 5 y 6).



### ¿Tienen los alumnos poco tiempo para estudiar debido a que trabajan?

Los resultados muestran, al igual que en los dos problemas ya mencionados, que los docentes de los CE estatales son los que mayores problemas de trabajo observan en sus estudiantes (véase los gráficos 7 y 8). Evidentemente, el hecho de que los estudiantes trabajen es percibido como un problema en los CE estatales y no en los no estatales. Esto podría tener que ver con el número de horas que los

estudiantes trabajan (lamentablemente no disponemos de esos datos, pero es una pregunta interesante por explorar en futuras encuestas CRECER).

Con seguridad, sería interesante correlacionar el trabajo de los estudiantes con los problemas mencionados aquí por los docentes y, eventualmente, con su rendimiento escolar. Es de esperar que la correlación sea negativa. Existe abundante bibliografía sobre el tema del trabajo infantil y adolescente. Al respecto, la OIT/IPEC y Unicef postulan su

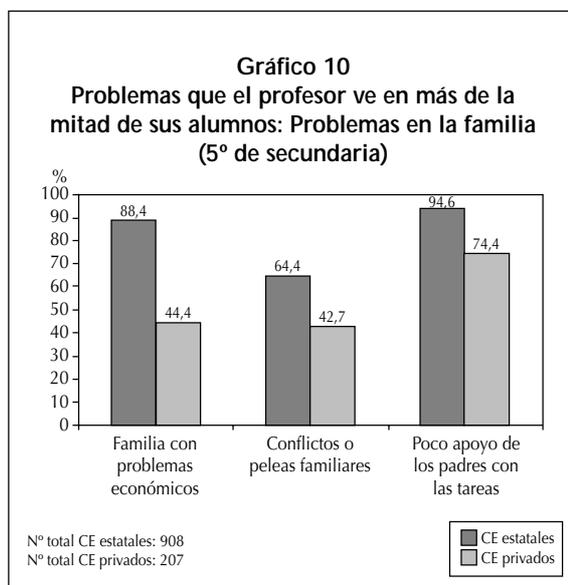
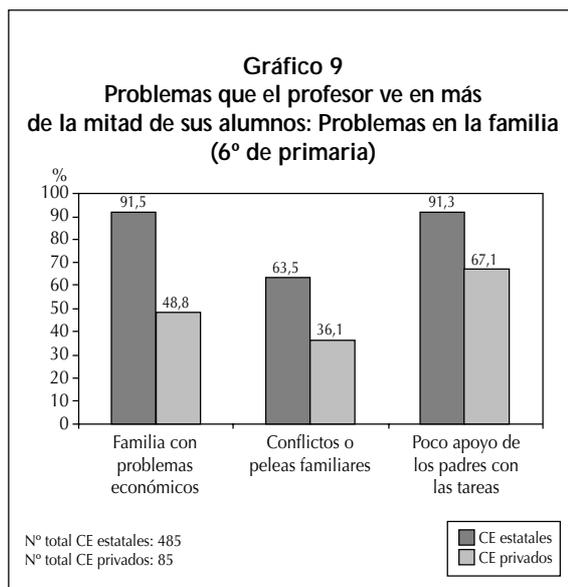
erradicación<sup>21</sup>. Cabe destacar que en nuestro país el asunto es motivo de discusión, ya que el nuevo Código de los Niños y Adolescentes ha fijado como edad mínima para el trabajo infantil los 12 años, mientras que el estándar fijado por la OIT en Latinoamérica es de 14 años. Unicef ha planteado revisar el tema pues, en su opinión, el trabajo infantil atenta contra el normal desarrollo y derechos esenciales del niño y, aún más, no contribuye a reducir la cantidad de niños que dejan de estudiar para dedicarse a trabajar<sup>22</sup>.

Por otro lado, al parecer, la opinión de aproximadamente la mitad de los docentes de CE estatales coincide con el reporte de las tardanzas como problema que presentan los estudiantes de dichos CE.

### ¿Tiene el alumno problemas familiares?

Se preguntó a los docentes si la mayoría de sus alumnos tenían problemas en casa<sup>23</sup>. Los tres temas tratados fueron: “problemas económicos en la familia”, “poco apoyo de los padres con las tareas” y “conflictos o peleas familiares”.

Los resultados muestran que casi todos los docentes de CE estatales reportan que la mayoría de las familias de sus alumnos tiene problemas económicos (91,5% en sexto de primaria y 88,4% en quinto de secundaria), y que los padres no brindan apoyo a los hijos con las tareas (91,3% en sexto de primaria y 94,6% en quinto de secundaria; véase los gráficos 9 y 10). En otras palabras, los padres de familia de menores recursos económicos optan por la gratuidad de la enseñanza (como se mostró al inicio de esta sección). Asimismo, son los padres de familia de CE estatales quienes tienen me-



nor grado de instrucción. Este hecho es una más de las posibles razones que dificultaría el que puedan ayudar a sus hijos en las tareas escolares, especialmente en secundaria.

Los conflictos o peleas familiares parecen ser un problema relativamente menor comparado con los dos anteriores, aunque más de la mitad de los docentes de CE estatales lo reporta.

### 3 A MODO DE CIERRE

En conclusión, se ha observado, al igual que en muchos otros estudios, que los alumnos de

21 Oficina Internacional del Trabajo (OIT): *Trabajo infantil en los países andinos: Bolivia, Colombia, Ecuador, Perú y Venezuela*. Ginebra: OIT, 1998. Documento de Trabajo N° 75.

22 “Niñez trabajadora”. *El Comercio*, 14 de agosto del 2000, p. A15.

23 La pregunta es la misma que se usó para los otros puntos de esta parte del documento: “¿Qué problemas o dificultades afectan a más de la mitad de los alumnos del aula donde usted es el docente responsable?”.

CE estatales provienen de estratos socioeconómicos más bajos que los de CE no estatales (se destaca el relativo menor nivel educativo de los padres); que la mayoría de los docentes de todos los CE percibe que sus alumnos son capaces, pero no se esfuerzan lo suficiente; que los principales problemas de estudio mostrados por los estudiantes (sobre todo en CE estatales en secundaria), a juicio de sus docentes, son: escaso hábito de lectura, malos hábitos de estudio y poco interés por los estudios; que en CE estatales los estudiantes presentan un alto porcentaje de inasistencia y tardanzas; que el trabajo de los estudiantes constituye un problema para el estudio en casa; y que en casa se presentan problemas de recursos económicos y poca ayuda de los padres para hacer las tareas con sus hijos. Los problemas anteriores se presentan en mayor porcentaje en CE estatales que en los no estatales, y en quinto de secundaria más que en sexto de primaria.

Tal vez más importante que las diferencias existentes entre los dos tipos de CE es que aparentemente los docentes son una fuente valiosa<sup>24</sup> para realizar diagnósticos sobre los problemas y necesidades de los estudiantes peruanos. Así, es posible que el “poco esfuerzo” que perciben los docentes respecto de sus estudiantes esté relacionado con los problemas que advierten en ellos, que fueron presentados en la segunda parte de esta sección. Es plausible también pensar en estudiantes que no se esfuerzan y faltan o llegan tarde al CE porque no ven la utilidad de lo que estudian, porque no les gusta leer, no saben estudiar y no tienen mayor ayuda en casa para resolver tareas escolares.

Afirmaciones como éstas no significan que ignoremos el contexto socioeconómico en que viven los estudiantes de CE estatales, que aparentemente fuerza a trabajar a un gran porcentaje de los alumnos. Sin embargo, esperamos que los datos aquí presentados puedan sugerir nuevas líneas de investigación o programas que busquen conocer mejor y

atenuar los problemas reportados por los docentes.

De todos los problemas anteriores, los más difíciles de enfrentar desde el sistema educativo serían el de los estudiantes que trabajan y los problemas familiares. Los más fáciles de abordar serían, qué duda cabe, los problemas de estudio y, parcialmente, los de asistencia y tardanza, aunque éstos podrían estar vinculados al trabajo. Es tal vez en estos problemas que se podría pensar inicialmente desde el MED: cómo ayudar a los docentes a aliviarlos de modo que el aprendizaje de los estudiantes mejore. En futuros trabajos sería interesante investigar la relación entre los problemas presentados en este documento y el rendimiento escolar.

## MUESTRA Y PROCEDIMIENTOS

La muestra es de CE de zonas urbanas del país. Se decidió trabajar sólo con sexto de primaria y quinto de secundaria, ya que los resultados no varían mucho entre los grados de un mismo nivel educativo, y estos grados fueron elegidos por tratarse de los últimos de cada nivel.

En sexto de primaria se trabajó con 491 docentes de CE estatales y 86 de CE no estatales; 60% fueron hombres y 40% mujeres. En quinto de secundaria se trabajó con 930 docentes de CE estatales y 211 de no estatales; 42% pertenecen al sexo masculino y 58% al femenino.

Los datos se obtuvieron de las encuestas CRECER 1998, tomadas en el marco de la segunda evaluación del rendimiento escolar CRECER 1998 realizada en noviembre de ese año. Los datos fueron recogidos por examinadores capacitados y no por docentes del aula ni del centro educativo.

---

24 De todos modos habría que demostrar esto con estudios empíricos.

# PARA EXPLICAR LAS DIFERENCIAS EN EL RENDIMIENTO EN MATEMÁTICA DE CUARTO GRADO EN EL PERÚ URBANO: ANÁLISIS DE RESULTADOS A PARTIR DE UN MODELO BÁSICO

MARTÍN BENAVIDES

La investigación educativa en América Latina ha llegado a definir un conjunto básico de variables que podrían explicar diferencias en el rendimiento académico. En esta sección se toman dichas variables y se analiza su relación con el rendimiento obtenido en una prueba de Matemática de cuarto de primaria. Para ello se utilizará la muestra CRECER 1998, en la que se administró una prueba de Matemática elaborada sobre la base del criterio de normas. Se presentarán, en primer lugar, algunas estadísticas descriptivas que incluyen como ejemplos otros grados y el caso de Comunicación Integral. En segundo lugar, los resultados a partir de un modelo de análisis multivariado y multinivel, pero sólo para el cuarto grado de primaria y en Matemática<sup>1</sup>. Investigaciones posteriores deberán complementar los resultados que aquí se presentan incorporando variables adicionales y realizando otros análisis.

## **1 LAS TRADICIONES DE INVESTIGACIÓN, LA EXPLICACIÓN DEL RENDIMIENTO Y LOS PROBLEMAS FUNDAMENTALES**

Se podría afirmar que los estudios sobre los problemas educativos en América Latina se agrupan en tres tipos de tradiciones: en primer lugar, aquellos de corte más cuantitativo y denominados comúnmente “análisis de fun-

ciones de producción educativa”; en segundo lugar, los estudios de corte más culturalista, que suelen utilizar metodologías de análisis cualitativo; finalmente, los de factores asociados, corriente que toma de los primeros la orientación hacia los estudios empíricos, mientras recoge de los segundos tanto la necesidad de enmarcar teóricamente los resultados cuanto la crítica al establecimiento de nexos causales.

Esta distinción es importante en la medida en que cada tipo de estudio levanta hipótesis diferentes en relación con los efectos democratizadores de la escuela. Mientras los primeros y los terceros sostienen hipótesis relacionadas con la posibilidad de los Estados para producir cambios educativos desde las escuelas, no se puede decir lo mismo de la tradición culturalista, que postula la incapacidad de la escuela para producir dichos resultados.

En relación con la explicación del rendimiento educativo y quizá por su naturaleza empírica, los estudios han sido básicamente orientados desde sus inicios por el paradigma de la función de producción, bajo la sombra de la investigación realizada en los países desarrollados, particularmente en los Estados

---

1 El autor quiere expresar su agradecimiento a la Unidad de Medición de la Calidad Educativa (UMC) por su apoyo, en especial a José Rodríguez y Tami Aritomi.

Unidos. Diversos analistas buscaron contrastar las hipótesis desarrolladas por Coleman<sup>2</sup>, entre otros. Estas investigaciones mostraban el limitado poder de los estados para controlar las escuelas. La sociedad producía determinaciones que los estados y sus políticas no podían —o no querían— modificar. De esa manera, las experiencias familiares terminaban siendo más importantes que las propias políticas educativas de las escuelas en la orientación de los niños.

Dichas hipótesis fueron sometidas a prueba en realidades diferentes de aquellas de los países desarrollados. Así, llegaron a ser cuestionadas por el estudio de Heyneman y Lowley, en el que se reivindicaban los efectos democratizadores de la escuela en los países no desarrollados, algunos de ellos latinoamericanos<sup>3</sup>. El argumento era que en estos países había una mayor homogeneidad en el vínculo entre las familias y la educación, así como una mayor heterogeneidad en el acceso a recursos escolares, lo que se percibía en ese momento por la expansión limitada de la escuela pública.

Esta relación entre factores extraescolares y escolares con el rendimiento ha seguido siendo un problema clave para los más recientes estudios de funciones de producción, así como de factores asociados. Si bien estos últimos no han desarrollado en la mayoría de

los casos modelos estadísticos apropiados, en la práctica han terminado asumiendo los mismos principios analíticos de los primeros.

De los diferentes estudios multivariados realizados hasta la actualidad se pueden recoger entonces las siguientes problemáticas fundamentales, que serán analizadas en esta sección.

## LA DISTRIBUCIÓN DE VARIANZA ENTRE FACTORES ESCOLARES Y NO ESCOLARES

Una primera duda que se ha intentado resolver teórica y empíricamente es cuánto de las diferencias en el rendimiento son diferencias relacionadas con características de las escuelas y cuánto de dichas diferencias se deben a características de los estudiantes y sus familias. Esta es una pregunta clave, pues indicará el campo de acción sobre el cual pueden actuar medidas de política orientadas hacia las escuelas. En Argentina, por ejemplo, se comprobó que en Matemática 53,4% y en Lengua 42% de las diferencias operan a través de las escuelas<sup>4</sup>. Por otro lado, en el Perú, un estudio del Banco Mundial que utilizó información de una prueba aplicada en 1996 encontró que 54% de dicha variabilidad se debía a diferencias entre las escuelas<sup>5</sup>.

## LA DESCOMPOSICIÓN DE LOS EFECTOS FAMILIARES

Otra problemática ha consistido en resolver la pregunta de cómo se relacionan con el rendimiento las variables vinculadas con las familias. Uno de los asuntos más importantes es responder a la pregunta de si existen efectos relacionados con las familias que sean independientes de su nivel socioeconómico. Si bien algunos estudios llegan a afirmar la importancia de las expectativas familiares y las características demográficas (número de miembros en el hogar o presencia de los familiares), existen otros que afirman que esas relaciones se deben básicamente a diferencias socioeconómicas entre las familias. De ser así, ni la presencia o número de familiares en el hogar, ni las expectativas familiares, mostrarían efectos significativos independientes<sup>6</sup>.

2 Coleman, J.; E. Campbell; C. Hobson; J. McPartland; A. Mood; F. Weinfeld y R. York: *Equality of Educational Opportunity*. Washington D.C.: Department of Health, Education and Welfare, 1966.

3 Heyneman, S. y W. Lowley: "The Effect of Primary-School Quality on Academic Achievement Across Twenty-Nine High-and Low-Income Countries", en *American Journal of Sociology*, 88(2), 1983, pp. 1162-1194.

4 Cervini, R.: "Los factores del rendimiento de Matemática en la escuela primaria urbana de Argentina", en *Seminario Iberoamericano Aspectos cualitativos y cuantitativos en la evaluación educativa. Una aproximación a los factores vinculados con el rendimiento escolar*. Buenos Aires: OEL y Ministerio de Cultura y Educación de Argentina, 1997.

5 The World Bank: *Peru, Education at a Crossroads. Challenges and Opportunities for the 21st Century*. Report N° 19066-PE. Washington: The World Bank, 1999.

6 Vélez, E.; E. Schiefelbein y J. Valenzuela: "Factors Affecting Achievement in Primary Education: a Re-

Por el contrario, otros estudios afirman que dos dimensiones que sí deben ser separadas de la socioeconómica son aquellas relacionadas con el capital cultural y el capital social. Capital cultural es el medio a través del cual las desigualdades de origen social son reproducidas inclusive en términos de logros educativos<sup>7</sup>. Las familias de clases altas tienen mayores posibilidades de vincularse con expresiones de la "cultura", lo que a su vez permite mejores logros académicos. Sin embargo, en Latinoamérica se ha llegado a mostrar empíricamente cómo el número de libros con los que se cuenta en el hogar ejerce un efecto independiente al socioeconómico<sup>8</sup>. Esto se explicaría por el hecho de que el gusto o disgusto por la cultura posiblemente esté distribuido homogéneamente entre diferentes grupos sociales.

Por otro lado, el capital social en educación tiene que ver con el seguimiento que hacen los padres de las diversas actividades educativas de los hijos. Dichos seguimientos se convierten en redes de apoyo social al estudiante<sup>9</sup>. De igual manera, el capital social podría comportarse de forma independiente del nivel socioeconómico de las familias, ya que puede llegar a producirse por incentivos propios de las escuelas, o características de los niños (por ejemplo, que estudiantes con problemas de aprendizaje necesiten ayuda extra). En Latinoamérica se ha llegado a mostrar su importancia en la determinación del rendimiento<sup>10</sup>.

Tales descomposiciones permitirán analizar los efectos netos del nivel socioeconómico de las familias. En este trabajo, y siguiendo las recomendaciones de Lookheed, Fuller y Nyrongo para países en desarrollo<sup>11</sup>, se usan índices compuestos por variables vinculadas al capital humano de las familias, pero también relacionados con las características de la vivienda y el acceso a servicios.

## LA PRESENCIA DE LOS ANTECEDENTES INDIVIDUALES

La presencia de los antecedentes individuales es probablemente una de las dimensiones menos problemáticas en cuanto a su relación con el rendimiento. Por ejemplo, se ha llega-

do a probar la importancia de la habilidad del niño para controlar los efectos del nivel socioeconómico<sup>12</sup>. Por otro lado, y a diferencia de lo que ocurre en países en desarrollo, en Latinoamérica los hombres siguen teniendo mejores resultados significativos que las mujeres<sup>13</sup>.

Aparte de ello, resultados anteriores en el Perú han mostrado la importancia relativa en el rendimiento de variables como la motivación y la realización de tareas<sup>14</sup>. Finalmente, compendios de otros estudios en Latinoamérica incorporan también las condiciones para el estudio y la distancia al centro educativo (CE) como variables explicativas del rendimiento<sup>15</sup>.

## LAS VARIABLES INSTITUCIONALES Y SOCIALES DE LAS ESCUELAS

Una vez establecidos los efectos relacionados con variables extraescolares, se deben analizar las variables vinculadas con las escuelas. En primer lugar, un problema clave es conocer

---

view of the Literature for Latin America and the Caribbean". Washington D.C.: The World Bank, 1993 (mimeo); Lareau, A.: "Social Class Differences in Family-School Relationships: the Importance of Cultural Capital", en *Sociology of Education*, 60 (2), 1987, pp. 73-85.

- 7 Katsillis, J. y R. Rubinson: "Cultural Capital, Student Achievement, and Educational Reproduction: the Case of Greece", en *American Sociological Review*, 55 (2), 1990, pp. 270-279.
- 8 Vélez *et al.*: "Factors Achievement...", *op. cit.*, 1993; The World Bank: *Peru Education...*, *op. cit.*, 1999.
- 9 Coleman, J.: "Social Capital in the Creation of Human Capital", en *American Journal of Sociology*, 94 (Supplement), 1988, pp. 95-120.
- 10 Vélez *et al.*: "Factors Achievement...", *op. cit.*, 1993.
- 11 Lockheed, M.; B. Fuller y R. Nyrongo: "Family Effects on Student's Achievement in Thailand and Malawi", en *Sociology of Education*, 62(4), 1989, pp. 239-256.
- 12 Katsillis y Rubinson: "Cultural Capital...", *op. cit.*, 1990.
- 13 Valverde, G. y M. Benavides: "Factores asociados al rendimiento en América Latina". Documento no publicado y preparado para PREAL, 1999; The World Bank: *Peru Education*, *op. cit.*, 1999.
- 14 The World Bank: *Peru Education...*, *op. cit.*, 1999.
- 15 Mella, O.: *Factores del rendimiento escolar: Estado del arte y análisis de datos*. Santiago de Chile: Simce, 1999; Vélez *et al.*: "Factors Achievement...", *op. cit.*, 1993.

cuáles son los efectos de las variables relacionadas con la gestión educativa. Diversos estudios descriptivos en Latinoamérica muestran las diferencias entre escuelas públicas y privadas<sup>16</sup>. En el Perú fue comprobada su importancia en un modelo multivariado y multinivel. Las escuelas privadas tenían mejores resultados que las escuelas públicas<sup>17</sup>. Ello no obstante, es necesario realizar análisis más precisos sobre la naturaleza de tal vínculo. En particular, cuánto de dicho efecto puede ser capturado por variables sociales de las escuelas, como sus recursos socioeconómicos, sus características geográficas u otras variables relacionadas con las escuelas y profesores. La inquietud final en este caso será la de llegar a saber si existe un efecto independiente relacionado con la gestión o si éste se diluye ante la presencia de las otras variables mencionadas.

En el caso de los recursos socioeconómicos de las escuelas, algunos estudios han mostrado ya su importancia relativa. Entre los más importantes están aquellos que privilegian la variable recursos económicos de la escuela como variable mediadora. La idea es que el mayor gasto por alumno, tanto en la instrucción cuanto en la administración de la escuela, permite lograr un mayor *ratio* en la proporción profesor/alumno, lo que, a su vez, produce una mayor cohesión en la escuela y, consecuentemente, redundando en mejores resultados académicos<sup>18</sup>.

16 Valverde y Benavides: "Factores asociados...", *op. cit.*, 1999.

17 The World Bank: *Peru Education...*, *op. cit.*, 1999.

18 Wenglinsky, H.: "How Money Matters: the Effect of School District Spending on Academic Achievement", en *Sociology of Education*, 70 (3), 1997, pp. 221-237.

19 Ravela, P.: "La búsqueda de escuelas productoras de conocimientos en el marco de la evaluación nacional de aprendizajes en Uruguay". *Seminario Iberoamericano Aspectos cualitativos...*, *op. cit.*, 1997.

20 Valverde y Benavides: "Factores asociados...", *op. cit.*, 1999.

21 The World Bank: *Peru Education...*, *op. cit.*, 1999.

22 Valverde y Benavides: "Factores asociados...", *op. cit.*, 1999.

23 Fuller, B.: "What School Factors Raise Achievement in the Third World", en *Review of Educational Research*, 57(3), 1987, pp. 255-292.

24 Vélez *et al.*: "Factors Achievement...", *op. cit.*, 1999.

Utilizando otros indicadores, Ravela ha llegado a estimar el efecto de la composición socioeconómica de las escuelas en el rendimiento escolar. Su estudio muestra que las diferencias institucionales relacionadas con el tipo de gestión disminuyen notablemente una vez que se incluye en el análisis la composición socioeconómica de las escuelas<sup>19</sup>.

Finalmente, estudios multivariados realizados en Honduras y Costa Rica mostraron la importancia de la ubicación geográfica de la escuela en la predicción del rendimiento<sup>20</sup>. En el Perú se han reportado ya diferencias significativas entre la costa y la selva y entre la sierra y la selva, ambas en perjuicio de los colegios de la selva<sup>21</sup>.

## LAS VARIABLES ESCOLARES Y DE LOS PROFESORES

La contribución de este tipo de variables en el rendimiento puede, en líneas generales, resumirse en las siguientes categorías: las características personales y profesionales de los docentes y directores, la pedagogía en el aula y la exposición del currículo, los recursos de infraestructura y materiales educativos, y el clima institucional.

En cuanto a las variables relacionadas con los profesores y directores, el estudio ya citado del Banco Mundial en el Perú reporta que sólo los años de servicio y la lengua materna del profesor fueron encontrados como variables significativas. El género del profesor es un factor asociado al rendimiento en Honduras<sup>22</sup>, mientras las variables relacionadas con la formación de los docentes no han tenido resultados sistemáticos<sup>23</sup>. El hallazgo más contundente es el reportado por Vélez y colegas, quienes luego de revisar diferentes estudios concluyen en su trabajo de 1993 que la capacitación formal es más importante que aquella que se logra en ejercicio.

En cuanto a las prácticas pedagógicas, se ha señalado en diversos estudios la diferencia entre los métodos activos y los pasivos<sup>24</sup>. Los primeros permiten mejorar el rendimiento de los estudiantes, lo que se corrobora en estudios realizados en Honduras y el Perú. Por ejemplo, se ha señalado la importancia de la explicación de los materiales, y el tipo de ta-

reas que se encargan<sup>25</sup>. Por otro lado, el tiempo dedicado al objetivo y el grado de cobertura del currículo se han mostrado como positivamente asociados con el rendimiento<sup>26</sup>. Ambas variables se constituirían en lo que se ha denominado oportunidades para aprender. Este concepto sigue el supuesto básico de que los estudiantes no podrán aprender a menos que tengan la oportunidad de hacerlo<sup>27</sup>.

En el caso de los recursos de infraestructura y de materiales de enseñanza, se ha mostrado que el acceso a los libros y a otros materiales de instrucción es importante para incrementar el rendimiento académico. Esto ha ocurrido en el Perú y Honduras, según los informes anteriormente citados. Por otro lado, mientras en Latinoamérica la provisión de infraestructura está positivamente asociada con el rendimiento, estudios en otros países encontraron que los recursos físicos y económicos de la escuela se relacionan con el aprendizaje sólo en siete casos<sup>28</sup>.

Finalmente, se ha llegado a considerar el clima de las escuelas como una importante dimensión explicativa del rendimiento. La idea básica es que la calidad de la escuela puede ser también definida por su clima o atmósfera. Para medirlo se usan variables como la colaboración entre administrativos y docentes, entre docentes, el tipo de liderazgo del director y la situación disciplinaria de las escuelas.

## **2 PREGUNTAS DE INVESTIGACIÓN, MUESTRA, VARIABLES DEL ESTUDIO Y MÉTODOS**

### **PREGUNTAS DE INVESTIGACIÓN**

La encuesta aplicada en 1998 permite analizar las relaciones entre el rendimiento y las variables clave anteriormente mencionadas. Las preguntas que se intenta responder son las siguientes:

- ¿Cuánto de la variación total en el rendimiento educativo es atribuible a diferencias entre las escuelas? Es decir, ¿qué margen de acción tienen las políticas educativas para influir en los logros educativos?
- ¿Cuál es el impacto relativo de las variables relacionadas con las características personales y familiares de los alumnos? ¿Qué importancia tiene en ese grupo la variable socioeconómica?
- Controlando<sup>29</sup> por dichas variables relacionadas con los estudiantes y sus familias, ¿qué efecto neto tiene la gestión pública o privada de la escuela, sus características socioeconómicas y su ubicación geográfica?
- ¿Cómo varía la importancia relativa de dichas variables cuando se compromete en el modelo el resto de variables relacionadas con los profesores y escuelas? Nos referimos a los efectos sobre el rendimiento de las características personales y profesionales de los docentes y directores, la pedagogía en el aula y la exposición del currículo, los recursos de infraestructura y materiales educativos, y el clima institucional.
- ¿Existe evidencia de una variabilidad importante de las relaciones entre las variables individuales y el rendimiento entre escuelas? Es decir, por ejemplo, ¿la relación entre el género y el rendimiento es significativamente diferente entre las escuelas? De ser así, ¿es posible explicar dicha relación con variables de las escuelas?

### **MUESTRA, VARIABLES Y MÉTODOS**

#### **La muestra**

La muestra que se analizará es la de CRECER 1998. En la primera parte se presentará la información descriptiva de la totalidad de la

25 The World Bank: *Peru Education...*, *op. cit.*, 1999; Valverde y Benavides: "Factores asociados...", *op. cit.*, 1999.

26 Valverde y Benavides, *ibidem*; Vélez et al: "Factors Achievement", *op. cit.*, 1999.

27 Riordan, C.: *Equality and Achievement: an Introduction to the Sociology of Education*. New York: Longman, 1997.

28 Hanushek, E.: "Money Might Matter Somewhere: a Response to Hedges, Laine, and Greenwald", en *Educational Researcher*, 23 (4), 1994, pp. 5-8.

29 Este término se refiere al control o represión de ciertas variables para así apartarlas de otras y poder observar los efectos de su ausencia.

muestra, en tanto que en la segunda, la parte analítica, se trabajará con la muestra reducida. Se ha hecho varias correcciones a la muestra original, las que se describen a continuación:

- Fueron eliminados los casos individuales que no son comunes o son muy diferentes del resto. Esto con el objetivo de no alterar las estimaciones estadísticas ni las pruebas que se realicen.
- Se descartaron algunas escuelas que no tenían suficiente información de algunas de las variables individuales.
- Se realizaron imputaciones por departamento y gestión, debido a la gran cantidad de casos de escuelas sin información en algunas de las variables de dicho nivel.

De esa manera, para la parte analítica, la muestra final de alumnos es de 12 855 y la de escuelas, 467.

### Las variables

Las variables que han sido utilizadas aparecen en el cuadro de las páginas 89 y 90.

### El método

En esta investigación la información será estudiada de dos maneras. En la primera parte se presentarán resultados descriptivos sobre la base de algunas de las variables. Se analizará información tanto de Matemática cuanto de Comunicación Integral de los cuatro años: cuarto y sexto de primaria y cuarto y quinto de secundaria. El segundo procedimiento consistirá en realizar un análisis multivariado, pero sólo para Matemática en cuarto grado. Así podremos saber a ciencia cierta si los efectos que aparecen en los resultados descriptivos se sostienen una vez que el modelo se hace más complejo. El método utilizado será el HLM o

modelo jerárquico lineal<sup>30</sup>. Este método tiene virtudes teóricas y empíricas. En primer lugar, permite distinguir efectos netos relacionados con las escuelas de aquellos relacionados con las familias y los estudiantes. Así, posibilita calcular qué variables explican las diferencias en los promedios de los colegios, al mismo tiempo que estima los determinantes individuales y familiares de la variación en el rendimiento que ocurre en las escuelas. Además de ello, el HLM 0 —a diferencia de la regresión lineal clásica— considera los efectos de la dependencia ocasionados cuando la información tiene dos niveles; en este caso, alumnos al interior de escuelas<sup>31</sup>.

A pesar de las correcciones realizadas en los casos individuales, aún persisten problemas pequeños de heterogeneidad y normalidad ocasionados por los pocos casos que tienen algunas escuelas. HLM puede corregir esos problemas, razón por la cual se decidió no eliminar esas escuelas, pero sí realizar estos análisis usando errores gruesos para poder dar mayor credibilidad a los resultados.

## 3 ALGUNOS RESULTADOS DESCRIPTIVOS

En este acápite algunas de las variables de los estudiantes, sus familias, profesores y escuelas, identificadas como importantes factores asociados al rendimiento por la investigación preliminar en América Latina, serán relacionadas con el puntaje obtenido en la prueba CRECER 1998. Se han seleccionado aquellas que muestran resultados más definidos en su relación con el rendimiento. Adicionalmente, se presentarán resultados según región y departamento. Se analizarán Matemática y Comunicación Integral tanto en primaria cuanto en secundaria. De la misma manera, se presentarán los resultados para escuelas públicas y privadas. Es necesario llamar la atención sobre la naturaleza preliminar de las relaciones que aquí se sugieren. Muchas de ellas pueden debilitarse o fortalecerse cuando se realicen análisis más complejos que tomen en cuenta un conjunto más amplio de variables. En el cuadro 1 se presentan los estadísticos de las variables utilizadas en este primer análisis.

30 Bryk, A. y S. Raudensbush: *Hierarchical Linear Models: Applications and Data Analysis Methods*. London: Sage, 1992.

31 Por razones del diseño de la muestra, la información de los profesores y directores fue agrupada en el nivel de la escuela.

Agenero	Género del alumno	Mujer = 1. Hombre = 0
Totmiem	Número de personas que viven con el alumno en la casa, incluyendo al alumno.	
Avivcon	Si el alumno vive con otros familiares u otros adultos.	Sí vive con otros adultos = 1. Otro caso = 0
Libmate	Si tiene el cuaderno de trabajo <i>Aprendemos Matemáticas</i> .	Sí tiene cuaderno = 1. No tiene cuaderno = 0
Asiste	Si asistió todos los días durante las últimas dos semanas.	Sí asistió = 1. No asistió = 0
Camino	Si se demora más de una hora en llegar al colegio.	Más de 1 hora = 1. Menos de 1 hora = 0
Tayuda1	Nadie lo ayuda con las tareas.	Si nadie lo ayuda = 1. Si alguien lo ayuda = 0
Tareama	Número de días que hizo tareas de Matemática.	
Taream2	Tareama al cuadrado.	
Trambie	Si estudia en ambiente adecuado.	
Gustlen	Si le gusta el curso de Lenguaje.	
Gustmat	Si le gusta el curso de Matemática.	
Mrendim	Tasa de acierto en la prueba de Matemática de los alumnos * 100	
Madreviv	Si la madre del alumno está viva.	
Pexp1	Expectativas de los padres.	Primaria completa = 1. Superior universitaria completa = 0
Pexp2	Expectativas de los padres.	Secundaria completa = 1. Superior universitaria completa = 0
Pexp3	Expectativas de los padres.	Superior no universitaria completa = 1. Superior universitaria completa = 0
Repite	Número de años que ha repetido el alumno.	
Repite2	Repite al cuadrado.	
Prinido	Si el alumno asistió a un preescolar privado.	
Pubnido	Si el alumno asistió a un preescolar público o a PRONOEI.	
Hhlibros	Número de libros en la vivienda del alumno.	De 0 a 10 = 1, de 11 a 25 = 2, de 26 a 50 = 3, de 50 a más = 4
Cuartos	Número de habitaciones en la vivienda del alumno.	
Innsse2	Índice socioeconómico. Promedio compuesto de número de activos estandarizado, acceso a servicios estandarizado y promedio de educación de los padres estandarizado.	
Ppmotiv	Cuán motivado se siente el responsable del aula en ese CE.	Muy descontento = 1, descontento = 2, contento = 3, muy contento = 4

► Ppgener	Género del responsable del aula.	Mujer = 1. Hombre = 0
Otrabaj2	Si el responsable del aula tiene otro trabajo diferente del de profesor.	
Pptitul3	Si el responsable del aula tiene título pedagógico de ISP.	
Pptitul4	Si el responsable del aula tiene título pedagógico universitario.	
Ppservi	Número de años de experiencia del responsable del aula como docente.	
Ppservi2	Ppservi al cuadrado.	
Matmate	Si tiene materiales de Matemática en la clase.	
Totclas	Tamaño de la clase.	Número de matriculados – número de retirados.
Ppncurr	Si el responsable del aula tiene un ejemplar del nuevo currículo.	
Ppreunp	Si el responsable del aula se reunió con otros profesores.	
Ppreund	Si el responsable del aula se reunió con el director.	
Material	Si el responsable del aula tiene libro de texto, libro de consulta, guías didácticas u otras lecturas.	
Probce4	Si el responsable del aula piensa que hay problemas de disciplina.	
Probc12	Si el responsable del aula piensa que el estilo de liderazgo del director es autoritario.	
Pphrpcl	Número de horas que usa el responsable del aula para preparar su clase, durante una semana.	
Hrsemat	Número de horas pedagógicas que enseña Matemática a la semana.	
Biblio	Si el aula tiene biblioteca o similar.	
Gestion	Gestión educativa.	Público = 1. Privado = 0
Costa	Región costa.	Costa = 1. Otra región = 0
Sierra	Región sierra.	Sierra = 1. Otra región = 0
Selva	Región selva.	Selva = 1. Otra región = 0
Dmobilia	Si el director considera que el colegio tiene problemas con el mobiliario.	
Dinfras	Si el director considera que el colegio tiene infraestructura deficiente.	
Dmater	Si el director considera que el colegio tiene materiales de educación desactualizados o no los tiene.	
Dapyfam	Si el director considera que el colegio tiene problemas de falta de apoyo de los padres de familia.	
Ctecho	Si el colegio presenta techo adecuado.	
Cinnse2	Promedio por colegio estandarizado del nivel socioeconómico de los alumnos.	
Ppntrad <sup>32</sup>	Índice que mide actividades pedagógicas no tradicionales de los profesores en clase. A mayor puntaje, mayor uso de técnicas no tradicionales. Está compuesto de variables normalizadas que corresponden a tres grupos: (i) tiempo que se dedica a objetivos en clase, reportados por el mismo profesor (11 variables); (ii) tipo de tareas con las que trabaja el profesor, reportadas por el profesor (4 variables); y, (iii) características del profesor, reportadas por el alumno (2 variables).	
Gxppntr	Interacción entre Ppntrad y Gestion.	
Nutri	Porcentaje de alumnos desnutridos en el colegio.	

32 Esta variable fue diseñada por Tami Aritomi y Claudia Galindo de la UMC.

**Cuadro 1**  
**Principales estadísticos de las variables empleadas en los análisis**

Variable	Cuarto grado de primaria					Sexto grado de primaria					Cuarto de secundaria					Quinto de secundaria											
	Obs.	Media	Desv. est.	Min.	Max.	Obs.	Media	Desv. est.	Min.	Max.	Obs.	Media	Desv. est.	Min.	Max.	Obs.	Media	Desv. est.	Min.	Max.	Obs.	Media	Desv. est.	Min.	Max.		
Mrendim	13 072	54,543	15,553	3,33	100	15 447	48,903	18,929	0	100	15 293	46,380	17,781	0	100	14,796	46,680	18,865	0	100							
Lrendim	13 072	53,758	18,517	0	100	15 447	54,559	19,966	0	96,77	15 506	44,188	15,652	0	94,74	15 054	54,121	15,851	0	97,44							
Agenero	13 225	0,500	0,500	0	1	15 602	0,501	0,500	0	1	15 519	0,536	0,499	0	1	15 083	0,525	0,499	0	1							
Gustmat	13 147	0,885	0,319	0	1	15 616	0,849	0,358	0	1	15 520	0,678	0,467	0	1	15 094	0,673	0,469	0	1							
Gustlen	13 177	0,886	0,318	0	1	15 649	0,897	0,304	0	1	15 486	0,883	0,321	0	1	15 072	0,885	0,319	0	1							
Tareama	13 189	2,811	1,022	1	4	15 650	2,659	0,903	1	4	15 538	2,060	0,712	1	4	15 108	1,997	0,701	1	4							
Tareale	13 162	2,657	0,986	1	4	15 649	2,556	0,898	1	4	15 529	2,154	0,764	1	4	15 109	1,983	0,706	1	4							
Nido	13 251	0,636	0,651	0	2	15 681	0,646	0,675	0	2	15 587	0,664	0,698	0	2	15 162	0,644	0,690	0	2							
Decinse																											
(Deciles																											
socioeconómicos)																											
Pexpec	8 927	5,494	2,871	1	10	10 553	5,494	2,876	1	10	15 492	5,490	2,874	1	10	15 055	5,489	2,875	1	10							
(Expectativas																											
de los padres)	8 878	3,508	0,808	1	4	10 510	3,553	0,756	1	4	10 395	2,721	0,537	1	3	10 121	2,729	0,508	1	3							
Tayuda1	13 187	0,190	0,392	0	1	15 639	0,328	0,470	0	1	15 537	0,586	0,493	0	1	15 104	0,661	0,473	0	1							
Gestion	13 251	0,882	0,322	0	1	15 681	0,851	0,356	0	1	15 587	0,820	0,384	0	1	15 162	0,820	0,384	0	1							
Ctecho	13 251	0,521	0,500	0	1	15 681	0,526	0,499	0	1	15 587	0,621	0,485	0	1	15 162	0,626	0,484	0	1							
Matmate	13 251	0,451	0,498	0	1	15 681	0,399	0,490	0	1	15 587	0,752	0,432	0	1	15 162	0,716	0,451	0	1							
Mititulo	13 251	1,781	0,852	0	3	8 367	1,711	0,922	0	3	15 587	2,037	1,016	0	3	15 162	2,066	0,981	0	3							
Lititulo	13 079	1,779	0,844	0	3	8 658	1,806	0,891	0	3	15 587	2,228	0,974	0	3	15 162	2,238	0,920	0	3							
Mapostg	13 251	0,029	0,167	0	1	8 367	0,043	0,202	0	1	15 587	0,050	0,217	0	1	15 162	0,040	0,197	0	1							
Lepostg	13 079	0,029	0,168	0	1	8 658	0,025	0,156	0	1	15 587	0,068	0,252	0	1	15 162	0,069	0,253	0	1							
Css	13 251	1,618	0,725	1	3	15 681	1,639	0,754	1	3	15 587	1,628	0,735	1	3	15 162	1,627	0,732	1	3							
Dept	13 251	13,086	6,669	1	25	15 681	13,224	6,598	1	25	15 587	13,020	6,466	1	25	15 162	12,955	6,458	1	25							

## LAS VARIABLES DE LOS ESTUDIANTES

### El género: ¿borrando o cambiando las diferencias?

A diferencia de lo que usualmente se supone, los resultados por género no serían tan contundentes. En primer lugar, las diferencias entre hombres y mujeres son más importantes en Matemática que en Comunicación Integral. En segundo lugar, mientras en la mayoría de los casos las diferencias en Matemática favorecen a los hombres, las diferencias en Comunicación Integral favorecen generalmente a las mujeres. Así, en cuarto grado de primaria los hombres tienen mejores puntajes promedio en Matemática, mientras las mujeres los obtienen en Comunicación Integral. Es interesante destacar que mientras la diferencia en Matemática a favor de los hombres se da básicamente en las escuelas públicas, la diferencia en Comunicación Integral a favor de las mujeres se da en las escuelas privadas (véase el cuadro 2).

En el caso de sexto grado, y tanto en escuelas públicas como privadas, hay diferencias en Matemática a favor de los hombres. Ello no obstante, diferencias a favor de la mujer aparecen en las escuelas privadas y en Comunicación Integral. En cuarto año de secundaria los hombres tienen mejores notas en Matemática, mientras las mujeres las obtienen en Comunicación Integral. Esa tendencia se da de la misma forma en escuelas públicas y privadas. En quinto de secundaria dichas tendencias son básicamente las mismas, salvo en las escuelas privadas, donde los hombres tienen mejores promedios en Comunicación Integral que las mujeres.

### El gusto por la materia

Los resultados son diferentes según el área de desarrollo que se analice. Mientras el rendimiento satisfactorio en Matemática requiere de estudiantes motivados, no ocurre lo mismo en Comunicación Integral. Aquellos a quienes les gusta la Matemática tienen en promedio mejores puntajes que aquellos a quienes no les gusta. Eso se da de la misma forma en primaria y en secundaria, así como

en colegios públicos y privados. Por el contrario, en los dos años de secundaria y en sexto de primaria, aquellos estudiantes que afirman no tener predilección por la Comunicación Integral obtienen mejores promedios que aquellos que dicen que sí les gusta. Esto sucede de la misma forma en escuelas públicas y privadas. Sólo en cuarto grado de primaria ocurre lo esperable; es decir, que aquellos a quienes les gusta la Comunicación Integral obtienen mejor puntaje promedio, lo cual se da especialmente en escuelas públicas (véase el cuadro 3).

### ¿Más tareas, mejor rendimiento?

La relación entre la variable número de días que hace tareas a la semana y el rendimiento muestra cosas interesantes. A diferencia de lo que sucede en los grados de primaria, en los dos años de secundaria pareciera ser que realizar tareas tiene un objetivo remediador. Efectivamente, en los dos años de secundaria y en las dos áreas, aquellos que hacen más tareas a la semana tienen menores puntajes en promedio. Esto sucede más frecuentemente en las escuelas públicas que en las privadas. Por el contrario, el comportamiento de esta relación para estudiantes de primaria es menos lineal. Al parecer, en este nivel aquellos que se ubican en los extremos (no hace tareas o hace muchas tareas) son los que, en promedio, tienen peores resultados, mientras que aquellos que se ubican en el medio de ambos grupos tienen aparentemente los más altos promedios (véase el cuadro 4).

### La importancia del preescolar

Estos resultados sugieren que los estudiantes que han asistido a un preescolar privado tienen mejores puntajes en promedio que aquellos que han asistido a uno público o aquellos que no han asistido a ningún tipo de preescolar. Cabe resaltar también que las diferencias entre estos dos últimos grupos en las escuelas públicas se inclinan a favor de aquellos que han asistido al preescolar público en comparación con aquellos que no han asistido. Por el contrario, en las escuelas privadas

Cuadro 2												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Hombres	64,64	54,00	55,26	65,88	46,76	49,62	62,72	45,08	48,13	65,76	45,08	48,42
Mujeres	64,84	52,36	53,83	64,68	45,29	48,19	57,92	41,90	44,92	58,26	41,92	45,15
<b>Total</b>	<b>64,74</b>	<b>53,18</b>	<b>54,54</b>	<b>65,28</b>	<b>46,03</b>	<b>48,90</b>	<b>60,04</b>	<b>43,39</b>	<b>46,41</b>	<b>61,44</b>	<b>43,45</b>	<b>46,70</b>
<b>Lenguaje</b>												
Hombres	66,04	51,41	53,14	69,80	51,83	54,51	56,23	41,33	43,87	66,17	51,18	53,61
Mujeres	70,60	52,21	54,38	71,94	51,56	54,61	57,46	41,53	44,52	65,19	52,02	54,62
<b>Total</b>	<b>68,31</b>	<b>51,81</b>	<b>53,76</b>	<b>70,87</b>	<b>51,69</b>	<b>54,56</b>	<b>56,92</b>	<b>41,43</b>	<b>44,21</b>	<b>65,61</b>	<b>51,61</b>	<b>54,14</b>

Cuadro 3												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
No le gusta	61,76	48,76	50,57	63,15	42,54	46,64	56,14	41,55	44,60	57,19	40,63	43,92
Le gusta	65,15	53,75	55,06	65,89	46,60	49,31	62,28	44,19	47,24	63,91	44,76	48,05
<b>Total</b>	<b>64,69</b>	<b>53,19</b>	<b>54,55</b>	<b>65,34</b>	<b>46,03</b>	<b>48,91</b>	<b>60,01</b>	<b>43,38</b>	<b>46,40</b>	<b>61,50</b>	<b>43,45</b>	<b>46,71</b>
<b>Lenguaje</b>												
No le gusta	70,41	48,96	51,29	74,97	51,83	56,74	60,15	44,74	49,04	68,64	54,16	58,94
Le gusta	68,03	52,21	54,10	70,20	51,69	54,32	56,20	41,02	43,54	64,81	51,35	53,51
<b>Total</b>	<b>68,28</b>	<b>51,84</b>	<b>53,78</b>	<b>70,89</b>	<b>51,70</b>	<b>54,57</b>	<b>56,92</b>	<b>41,40</b>	<b>44,18</b>	<b>65,62</b>	<b>51,61</b>	<b>54,13</b>

Cuadro 4												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Ningún día	65,62	49,09	51,18	64,91	42,63	46,26	60,64	44,33	48,11	63,56	44,04	48,41
De 1 a 2 días	65,35	53,16	54,66	64,49	45,44	48,25	59,72	43,39	46,16	59,82	43,58	46,34
De 3 a 4 días	66,08	55,16	56,71	67,10	48,74	51,76	59,64	43,33	46,23	63,53	43,09	46,53
De 5 a más días	62,52	53,17	54,07	63,56	44,37	46,72	62,27	38,68	41,72	63,84	38,26	41,58
<b>Total</b>	<b>64,77</b>	<b>53,19</b>	<b>54,55</b>	<b>65,30</b>	<b>46,02</b>	<b>48,90</b>	<b>59,99</b>	<b>43,38</b>	<b>46,39</b>	<b>61,49</b>	<b>43,44</b>	<b>46,70</b>
<b>Lenguaje</b>												
Ningún día	72,16	50,67	53,56	74,33	48,53	53,54	59,36	42,24	46,10	67,80	52,38	55,91
De 1 a 2 días	69,70	52,89	55,07	70,81	52,16	54,99	56,56	41,74	44,46	65,03	52,28	54,44
De 3 a 4 días	67,66	52,61	54,28	70,83	53,69	56,21	55,78	41,04	43,21	64,16	49,17	51,43
De 5 a más días	64,42	50,13	51,59	67,89	49,09	51,38	53,04	37,43	39,35	62,08	45,99	48,68
<b>Total</b>	<b>68,36</b>	<b>51,84</b>	<b>53,79</b>	<b>70,85</b>	<b>51,71</b>	<b>54,57</b>	<b>56,92</b>	<b>41,42</b>	<b>44,20</b>	<b>65,63</b>	<b>51,63</b>	<b>54,15</b>

Cuadro 5												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
No asistió	64,37	51,93	53,13	63,41	44,82	47,01	58,66	42,43	44,79	60,02	42,20	44,79
CE estatal	62,41	53,46	54,27	60,99	46,20	47,72	57,28	43,30	45,02	58,44	43,82	45,74
CE privado	68,00	59,67	62,57	70,89	52,98	61,08	63,46	49,17	56,06	65,80	49,13	57,05
<b>Total</b>	<b>64,73</b>	<b>53,18</b>	<b>54,54</b>	<b>65,28</b>	<b>46,03</b>	<b>48,90</b>	<b>59,99</b>	<b>43,37</b>	<b>46,38</b>	<b>61,44</b>	<b>43,43</b>	<b>46,68</b>
<b>Lenguaje</b>												
No asistió	68,21	49,90	51,67	69,42	49,88	52,17	55,93	40,18	42,45	64,58	50,25	52,33
CE estatal	64,56	52,44	53,55	66,39	52,32	53,77	53,45	41,48	42,93	63,00	52,05	53,48
CE privado	72,94	60,35	64,73	76,19	59,98	67,32	60,61	48,21	54,15	69,13	57,71	63,09
<b>Total</b>	<b>68,30</b>	<b>51,81</b>	<b>53,76</b>	<b>70,87</b>	<b>51,69</b>	<b>54,56</b>	<b>56,91</b>	<b>41,41</b>	<b>44,19</b>	<b>65,59</b>	<b>51,60</b>	<b>54,12</b>

aquellos que no han asistido tienen mejores puntajes que aquellos que han ido a un pre-escolar público. Esto sucede de la misma manera en los dos niveles y materias (véase el cuadro 5).

## LAS FAMILIAS DE LOS ALUMNOS

### El nivel socioeconómico y las expectativas de los padres

¿Cuán importante es ser hijo o hija de padres con mayores recursos sociales y económicos? Según los resultados preliminares, los estudiantes que proceden de las familias de los deciles socioeconómicos más altos siguen obteniendo mejores resultados tanto en primaria cuanto en secundaria en Matemática y en Comunicación Integral. Es decir, a mayor nivel socioeconómico, mejores resultados en las pruebas. Esta tendencia se reproduce de esa manera en escuelas públicas y privadas (véase el cuadro 6).

En el caso de las expectativas, el resultado más consistente es que los mejores promedios los tienen aquellos alumnos cuyos padres esperan que estudien educación superior. Esto sucede en los cuatro grados analizados en Matemática y en Comunicación Integral, tanto en escuelas públicas cuanto en privadas. Por otro lado, es interesante destacar

que no existe una relación lineal entre esta variable y el rendimiento. En algunos casos, los hijos o hijas de padres cuya expectativa es que sus hijos lleguen hasta la secundaria técnica tienen menos promedio que aquellos cuyos padres aspiran que lleguen sólo hasta secundaria. En otros casos sucede que los hijos o hijas de padres que tienen como expectativa que sus hijos lleguen a secundaria, alcanzan promedios más bajos que aquellos cuyos padres sólo aspiran a que lleguen a la primaria (véase el cuadro 7).

### El apoyo familiar en las tareas

Nuevamente, destaca el hecho de que en algunos grados y situaciones las tareas tienen un papel remediador. En cuarto de primaria, aquellos alumnos o alumnas que reciben apoyo de algún familiar tienen menores puntajes en Matemática y Comunicación Integral que aquellos que no reciben ayuda. Esta diferencia obedece básicamente al comportamiento que se da en las escuelas públicas (véase el cuadro 8). En sexto de primaria, por otro lado, los alumnos de escuelas privadas que reciben apoyo tienen menores promedios en Matemática y Comunicación Integral, mientras lo contrario ocurre en las escuelas públicas. En el caso de los años de secundaria, las diferencias no existen o son muy pequeñas. En los

Cuadro 6												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Decil 1	52,18	47,03	47,18	46,53	38,52	38,89	38,77	37,76	37,78	43,79	38,26	38,40
Decil 2	57,02	47,73	48,02	54,44	40,02	40,69	46,19	38,64	38,93	50,44	40,37	40,76
Decil 3	53,52	48,53	48,73	49,76	41,66	42,08	49,59	40,97	41,39	50,08	41,18	41,61
Decil 4	55,86	50,92	51,11	56,14	43,36	43,98	50,94	41,88	42,56	49,12	41,61	42,01
Decil 5	55,22	53,01	53,15	57,46	44,85	45,58	53,21	43,07	43,80	51,48	41,71	42,58
Decil 6	58,90	54,42	54,77	57,43	46,11	47,16	54,93	44,77	45,88	54,26	44,19	45,39
Decil 7	59,92	55,14	55,58	62,01	49,04	50,84	55,22	45,96	47,49	58,47	46,73	48,70
Decil 8	61,87	58,12	58,66	63,20	51,76	53,97	55,94	48,06	49,86	58,01	47,33	50,02
Decil 9	66,79	59,85	61,32	67,00	56,17	59,12	60,84	50,85	54,83	62,26	50,28	54,85
Decil 10	69,98	64,40	67,03	73,80	59,48	67,67	67,17	52,03	62,02	69,06	53,03	63,31
<b>Total</b>	64,38	53,19	54,52	65,74	45,99	48,99	59,98	43,42	46,43	61,44	43,49	46,74
<b>Lenguaje</b>												
Decil 1	53,33	43,75	44,03	54,10	42,51	43,04	39,41	35,21	35,32	51,35	45,11	45,28
Decil 2	53,81	45,85	46,10	55,51	42,90	43,48	47,24	36,62	37,02	55,13	48,02	48,29
Decil 3	56,67	45,74	46,17	54,37	45,55	46,01	46,13	38,16	38,55	55,62	49,11	49,42
Decil 4	54,14	48,27	48,49	60,91	49,19	49,76	48,22	39,94	40,54	58,94	49,45	49,94
Decil 5	56,29	50,57	50,91	61,13	51,08	51,66	52,16	40,36	41,21	58,24	50,42	51,10
Decil 6	58,67	53,79	54,17	63,47	52,55	53,56	51,75	43,22	44,13	61,74	53,61	54,57
Decil 7	59,79	55,53	55,92	66,73	56,35	57,79	53,65	44,67	46,14	61,32	55,15	56,18
Decil 8	64,11	57,24	58,22	68,89	59,80	61,55	52,66	46,33	47,76	63,80	56,06	58,00
Decil 9	69,54	61,33	63,07	72,05	63,69	65,96	57,71	49,77	52,89	66,66	58,42	61,59
Decil 10	75,93	67,86	71,67	79,33	67,44	74,25	63,26	51,16	59,14	71,05	60,98	67,41
<b>Total</b>	67,48	51,99	53,82	70,96	51,78	54,69	56,95	41,46	44,24	65,61	51,65	54,17

Cuadro 7												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Primaria completa	60,76	48,63	49,63	60,49	43,52	45,78						
Secundaria completa	58,45	48,33	48,94	62,81	40,01	42,16	55,78	40,01	41,91	56,17	39,78	41,05
Superior técnica completa	58,04	51,03	51,51	56,05	42,52	43,50	45,69	39,89	40,23	52,19	39,83	40,83
Superior universitaria completa	65,68	54,89	56,43	67,07	48,09	51,54	61,11	44,59	48,18	62,63	44,91	48,72
<b>Total</b>	64,38	53,19	54,51	65,73	46,01	49,00	60,00	43,35	46,39	61,58	43,54	46,83
<b>Lenguaje</b>												
Primaria completa	63,79	48,30	49,58	62,49	45,01	47,34						
Secundaria completa	62,47	45,78	46,78	67,74	43,50	45,80	55,41	36,69	38,92	60,26	46,39	47,49
Superior técnica completa	57,07	49,26	49,79	59,75	48,21	49,05	45,00	37,92	38,34	57,63	48,00	48,77
Superior universitaria completa	69,20	54,07	56,23	72,55	54,42	57,71	57,57	42,73	45,91	66,39	53,18	56,00
<b>Total</b>	67,46	52,00	53,83	70,96	51,80	54,71	56,72	41,39	44,15	65,50	51,73	54,23

Cuadro 8												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Recibe ayuda	64,72	52,77	54,23	64,54	46,31	49,20	59,18	43,72	46,63	60,87	43,67	46,78
No recibe ayuda	65,02	54,85	55,90	67,15	45,49	48,32	60,56	43,14	46,21	61,79	43,33	46,67
<b>Total</b>	<b>64,77</b>	<b>53,18</b>	<b>54,55</b>	<b>65,29</b>	<b>46,03</b>	<b>48,91</b>	<b>59,97</b>	<b>43,38</b>	<b>46,39</b>	<b>61,48</b>	<b>43,45</b>	<b>46,70</b>
<b>Lenguaje</b>												
Recibe ayuda	68,32	51,45	53,50	70,57	52,21	55,12	56,37	41,70	44,44	64,51	51,66	53,97
No recibe ayuda	68,54	53,39	54,94	71,62	50,73	53,46	57,32	41,22	44,03	66,19	51,59	54,22
<b>Total</b>	<b>68,36</b>	<b>51,82</b>	<b>53,78</b>	<b>70,87</b>	<b>51,71</b>	<b>54,57</b>	<b>56,91</b>	<b>41,42</b>	<b>44,20</b>	<b>65,62</b>	<b>51,61</b>	<b>54,14</b>

casos en que aquellas existen, ocurren principalmente en escuelas privadas, donde se muestra que los alumnos o alumnas que reciben apoyo tienen menores puntajes en Matemática y en Comunicación Integral.

## LAS ESCUELAS Y LOS PROFESORES

En los cuadros anteriores se han podido notar las diferencias existentes entre colegios públicos y privados. Los niños que estudian en escuelas privadas tienden a obtener mejores resultados en sus pruebas que aquellos que estudian en escuelas públicas. Esto sucede tanto en primaria como en secundaria, así como en Matemática y Comunicación Integral. Sería interesante evaluar si esta relación se mantiene o desaparece cuando uno analiza las variables de las escuelas y profesores.

### La infraestructura

Estos resultados preliminares sugieren que es importante contar con escuelas con techo adecuado. Tanto en primaria como en secundaria y en Matemática y en Comunicación Integral, los estudiantes de escuelas que, según el director, tienen techo adecuado, obtienen mejores resultados en promedio que aquellos estudiantes cuyas escuelas, según su director, no tienen un techo adecuado. Esto sucede en

colegios públicos y privados. Es necesario llamar la atención sobre la naturaleza preliminar de estos datos. Este “efecto” puede en realidad estar capturando características de otras variables, lo que se podría concluir a partir de análisis multivariados más complejos (véase el cuadro 9).

### El acceso a materiales de Matemática

Los alumnos o alumnas cuyos profesores dicen que cuentan con materiales de Matemática suelen tener mejores puntajes en promedio que aquellos cuyos profesores afirman no contar con ellos. Esta tendencia se reproduce en los dos niveles analizados, así como en colegios públicos y privados. En estos últimos, sin embargo, la diferencia aludida suele ser más importante (véase el cuadro 10).

### La formación de los docentes

La relación entre la titulación del profesor y el rendimiento de los alumnos no es sistemática. Destaca el hecho de que en las escuelas privadas de cuarto grado los alumnos o alumnas cuyos profesores no se han titulado son los que tienen mejores promedios en Matemática y Comunicación Integral. No ocurre lo mismo en las escuelas públicas, donde son los alumnos o alumnas cuyos profesores se

Cuadro 9												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Inadecuado	56,95	50,24	50,61	57,02	42,47	43,26	53,84	40,94	41,73	57,18	40,89	41,86
Adecuado	66,93	56,32	58,20	67,00	49,99	53,99	60,89	45,22	49,20	62,04	45,35	49,56
<b>Total</b>	<b>64,73</b>	<b>53,18</b>	<b>54,54</b>	<b>65,28</b>	<b>46,03</b>	<b>48,90</b>	<b>59,99</b>	<b>43,37</b>	<b>46,38</b>	<b>61,44</b>	<b>43,43</b>	<b>46,68</b>
<b>Lenguaje</b>												
Inadecuado	59,08	48,05	48,64	60,23	46,92	47,64	49,71	38,08	38,79	60,43	48,31	49,04
Adecuado	70,91	55,83	58,51	73,08	57,02	60,79	57,98	43,95	47,48	66,32	54,06	57,14
<b>Total</b>	<b>68,30</b>	<b>51,81</b>	<b>53,76</b>	<b>70,87</b>	<b>51,69</b>	<b>54,56</b>	<b>56,91</b>	<b>41,41</b>	<b>44,19</b>	<b>65,59</b>	<b>51,60</b>	<b>54,12</b>

Cuadro 10												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
No accede	62,22	52,61	53,18	62,68	43,86	45,47	56,61	42,74	44,35	56,66	42,71	44,26
Accede	65,70	54,00	56,21	62,22	47,71	49,61	60,63	43,60	47,06	62,45	43,75	47,64
<b>Total</b>	<b>64,73</b>	<b>53,18</b>	<b>54,54</b>	<b>62,40</b>	<b>45,76</b>	<b>47,56</b>	<b>59,99</b>	<b>43,37</b>	<b>46,38</b>	<b>61,44</b>	<b>43,43</b>	<b>46,68</b>

han titulado en universidades o institutos pedagógicos los que tienen mejores promedios. Por otro lado, en sexto grado, mientras en las escuelas públicas los promedios más bajos en Matemática y Comunicación Integral los tienen los estudiantes cuyos profesores no se han titulado, los promedios más altos los alcanzan aquellos cuyos profesores se han titulado en la universidad. Sin embargo, en las escuelas privadas los promedios más altos en Matemática y Comunicación Integral los tienen los alumnos o alumnas de profesores que se han titulado en institutos pedagógicos (véase el cuadro 11).

En el caso de cuarto de secundaria, vuelve a destacar el hecho de que en las escuelas privadas los estudiantes cuyos profesores no se han titulado tienen mejores promedios en Matemática que aquellos cuyos profesores se han titulado en institutos pedagógicos o a través de la profesionalización docente. Solamente son superados en promedio por aque-

llos estudiantes cuyos profesores se han titulado en universidades. En las escuelas públicas, las mejores notas las tienen los estudiantes de profesores titulados en universidades, seguidos por aquellos cuyos profesores se han titulado en institutos pedagógicos. En Matemática, los que obtienen menor promedio son los alumnos de docentes titulados a través de profesionalización docente, mientras en Comunicación Integral son los estudiantes de profesores no titulados.

En quinto de secundaria se da un patrón similar. En los colegios privados, los estudiantes de profesores no titulados tienen el mejor promedio en Comunicación Integral o están detrás de aquellos estudiantes con profesores titulados en institutos pedagógicos. Ello no sucede en las escuelas públicas, donde los estudiantes de profesores no titulados obtienen los promedios más bajos, tanto en Matemática como en Comunicación Integral.

Cuadro 11												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Sin título	67,25	52,10	53,91	59,06	42,43	44,56	60,30	40,28	47,72	61,65	41,05	48,05
Profesionalización docente	62,67	49,98	50,89	63,33	44,73	45,65	57,77	43,08	46,12	60,11	40,56	43,13
ISP	63,91	53,67	54,80	63,77	46,04	48,29	61,43	42,87	45,42	60,12	42,87	45,39
Universidad	65,98	54,94	56,99	60,06	48,76	49,70	58,85	44,78	46,91	62,70	45,16	48,16
<b>Total</b>	<b>64,73</b>	<b>53,18</b>	<b>54,54</b>	<b>62,40</b>	<b>45,76</b>	<b>47,56</b>	<b>59,99</b>	<b>43,37</b>	<b>46,38</b>	<b>61,44</b>	<b>43,43</b>	<b>46,68</b>
<b>Lenguaje</b>												
Sin título	70,48	50,93	52,69	63,45	47,80	49,19	59,42	35,45	40,93	65,09	48,44	50,64
Profesionalización docente	68,76	49,38	50,77		49,76	49,76	55,04	39,78	41,86	64,42	51,35	52,06
ISP	65,19	52,42	53,78	71,84	50,53	53,62	55,21	41,68	43,50	66,35	51,15	53,76
Universidad	69,73	52,17	54,77	70,81	52,42	56,64	57,19	42,66	45,61	65,21	52,67	55,32
<b>Total</b>	<b>66,96</b>	<b>51,81</b>	<b>53,42</b>	<b>70,81</b>	<b>50,37</b>	<b>53,12</b>	<b>56,91</b>	<b>41,41</b>	<b>44,19</b>	<b>65,59</b>	<b>51,60</b>	<b>54,12</b>

Cuadro 12												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Sin posgrado	64,85	53,20	54,56	62,40	45,51	47,43	59,70	43,06	46,01	60,91	43,33	46,43
Con posgrado	61,55	52,46	53,90		50,65	50,65	63,77	49,84	53,42	69,49	46,30	52,74
<b>Total</b>	<b>64,73</b>	<b>53,18</b>	<b>54,54</b>	<b>62,40</b>	<b>45,76</b>	<b>47,56</b>	<b>59,99</b>	<b>43,37</b>	<b>46,38</b>	<b>61,44</b>	<b>43,43</b>	<b>46,68</b>
<b>Lenguaje</b>												
Sin posgrado	67,11	51,79	53,40	70,45	50,26	52,96	56,10	41,26	43,74	64,63	51,48	53,64
Con posgrado	63,56	52,66	54,39	83,98	55,06	59,23	62,18	43,92	50,31	71,03	53,83	60,64
<b>Total</b>	<b>66,96</b>	<b>51,81</b>	<b>53,42</b>	<b>70,81</b>	<b>50,37</b>	<b>53,12</b>	<b>56,91</b>	<b>41,41</b>	<b>44,19</b>	<b>65,59</b>	<b>51,60</b>	<b>54,12</b>

Cuadro 13												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Costa	65,81	56,28	57,65	67,03	49,76	53,20	62,06	44,98	48,83	62,86	44,73	48,82
Sierra	65,69	50,82	52,24	63,81	44,27	46,31	57,52	42,29	44,85	61,30	43,17	46,12
Selva	54,60	48,03	48,53	54,02	38,91	40,02	46,95	40,84	41,18	44,43	40,21	40,48
<b>Total</b>	<b>64,73</b>	<b>53,18</b>	<b>54,54</b>	<b>65,28</b>	<b>46,03</b>	<b>48,90</b>	<b>59,99</b>	<b>43,37</b>	<b>46,38</b>	<b>61,44</b>	<b>43,43</b>	<b>46,68</b>
<b>Lenguaje</b>												
Costa	70,67	56,22	58,30	72,37	56,64	59,78	59,32	44,05	47,46	68,04	53,98	57,15
Sierra	66,50	47,60	49,41	70,28	47,07	49,48	53,44	38,85	41,26	62,19	48,80	50,95
Selva	57,26	46,42	47,25	59,51	46,27	47,24	45,67	38,68	39,08	53,63	50,17	50,38
<b>Total</b>	<b>68,30</b>	<b>51,81</b>	<b>53,76</b>	<b>70,87</b>	<b>51,69</b>	<b>54,56</b>	<b>56,91</b>	<b>41,41</b>	<b>44,19</b>	<b>65,59</b>	<b>51,60</b>	<b>54,12</b>

## Los estudios de posgrado

Por otro lado, en el caso de los estudios de posgrado, salvo en el caso de cuarto grado de primaria, donde los estudiantes de colegios privados cuyos profesores han hecho estudios de posgrado tienen menores promedios en Matemática y Comunicación Integral, en el resto de los casos sucede consistentemente lo contrario. Es decir, los alumnos o alumnas cuyos profesores han hecho estudios de posgrado tienen mejores promedios en Matemática y Comunicación Integral. Es interesante destacar que la diferencia más grande entre los promedios de alumnos cuyos profesores tienen posgrado y aquellos cuyos profesores no lo tienen, se da entre los promedios de Comunicación Integral de las escuelas privadas de sexto de primaria. No obstante, es necesario resaltar que las comparaciones se han limitado por el poco número de casos en algunas de las categorías.

## LAS DIFERENCIAS REGIONALES Y POR DEPARTAMENTO

Las diferencias por región natural (costa, sierra, selva) son consistentes a través de los diferentes grados y áreas. Los niños que estudian en escuelas de la costa tienen mejores resultados que los de la sierra y la selva. Al mismo tiempo, los colegios de la sierra tienen mejores resultados que aquellos de la selva. Esto sucede tanto para primaria y secundaria cuanto para Matemática y Comunicación Integral. Las tendencias son básicamente las mismas según escuelas públicas o privadas. Ello no obstante, destaca un pequeño cambio en el caso de las escuelas públicas y en los dos años de secundaria. La diferencia existente en Comunicación Integral entre la sierra y la selva desaparece o cambia su sentido (véase el cuadro 13). ¿Cuáles son los departamentos con los más altos y más bajos promedios en las pruebas? En el caso de Matemática de cuarto grado, Arequipa y Callao tienen los promedios más altos, mientras que los más bajos se presentan en Apurímac y Ucayali. En el caso del sexto grado, nuevamente Callao y Arequipa tienen los promedios más altos, mientras que los más bajos esta vez están en Apurímac

y Ucayali. En el caso de secundaria y siempre para Matemática, Callao y Lima tienen los mejores resultados, mientras que Huánuco y Madre de Dios, los más bajos. En quinto año de secundaria destacan Callao y Junín como los departamentos con los promedios más altos, mientras Tumbes y Ucayali son los más rezagados (véase el cuadro 14a). En el caso de cuarto de primaria, en Comunicación Integral los departamentos con los más altos promedios son Callao y Moquegua. Por otro lado, Apurímac y Loreto destacan como aquellos departamentos con los más bajos promedios. En el caso de sexto de primaria, Callao y Lima tienen los más altos promedios, mientras que los departamentos con los más bajos promedios son Apurímac y San Martín. En cuarto de secundaria destacan Callao y Lima como los departamentos con los mejores resultados, mientras que Huancavelica y Ucayali muestran los resultados relativos más bajos. Finalmente, Callao y Tacna son los departamentos con los promedios más altos, mientras que los menores promedios los tienen Apurímac y Huancavelica (véase el cuadro 14b).

## 4 RESULTADOS MULTIVARIADOS Y MULTINIVEL: EL CASO DE CUARTO GRADO EN MATEMÁTICA

En primer lugar, y tal como se puede observar en la columna perteneciente al modelo nulo (sin incluir variables independientes para ver la varianza total), en el cuadro final de resultados la varianza que ocurre por diferencias entre las escuelas es de 98,33. Esto quiere decir que 41%<sup>33</sup> de las diferencias en Matemática se debe a factores relacionados con el hecho de que los estudiantes estén agrupados en escuelas. Esa varianza es explicada por las variables de las escuelas, pero puede serlo también por variables individuales. De alguna manera, estas últimas pueden tener también efectos que son resultados de su agrupamiento en escuelas. Así, si se observan las nuevas varianzas resultantes del modelo 1, las

33 Este es un *ratio* de la varianza en el nivel 2 sobre la varianza total.

Cuadro 14a												
	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total	CE privados	CE estatales	Total
<b>Matemática</b>												
Amazonas	50,78	52,70	52,56	38,33	45,93	45,12	43,79	48,48	47,89	39,33	48,96	47,73
Ancash	64,72	53,10	55,22	63,78	50,18	52,85	53,86	42,30	44,74	58,96	44,11	47,56
Apurímac		42,71	42,71		34,90	34,90	47,31	40,44	41,24	47,67	38,54	39,45
Arequipa	66,01	63,29	63,96	58,96	53,04	54,39	55,22	48,94	50,62	59,27	48,13	51,20
Ayacucho	68,11	44,11	45,75	76,19	37,27	41,81	63,45	39,06	43,83	66,56	41,51	46,38
Cajamarca	56,35	53,74	54,40	61,87	49,25	51,59		45,33	45,33		45,19	45,19
Callao	65,99	58,00	60,29	63,03	54,99	57,47	55,34	50,73	52,29	62,37	47,74	53,44
Cusco		50,38	50,38	60,72	45,46	47,24	54,07	40,69	42,99	62,59	41,96	45,47
Huancavelica		53,08	53,08		44,49	44,49	53,46	39,59	41,14	56,39	42,47	43,74
Huánuco	58,30	52,54	53,20	60,16	37,99	40,04	57,47	36,76	39,24	57,59	41,09	43,09
Ica	60,22	53,53	53,84	53,93	47,82	48,35	64,29	42,77	45,06	63,16	43,37	45,34
Junín	65,86	54,53	55,26	66,89	53,02	54,40	63,23	46,37	50,13	64,96	48,78	52,40
La Libertad	62,46	55,62	56,92	64,73	45,58	49,88	61,47	43,95	47,05	58,51	45,40	47,67
Lambayeque	57,57	55,18	55,54	62,67	51,30	53,68	48,47	44,34	44,97	47,46	41,80	42,62
Lima	69,14	56,65	58,84	71,05	50,81	55,86	65,62	46,53	51,62	66,81	45,54	51,16
Loreto	53,44	46,44	47,54	56,91	35,25	37,99	46,38	39,55	40,29	47,35	40,41	41,17
Madre de Dios		46,86	46,86		37,36	37,36		39,60	39,60		37,48	37,48
Moquegua	65,78	58,31	59,25	65,62	51,00	52,70	68,57	44,37	47,46	64,88	47,65	49,37
Pasco	60,00	51,81	53,18	56,82	47,15	48,09	52,95	48,34	48,94	51,32	44,60	45,40
Piura	56,31	52,00	52,39	58,68	46,92	48,67	53,69	40,63	43,13	54,82	41,64	44,39
Puno	72,20	47,37	51,22	67,94	41,14	44,75	62,90	36,16	39,61	64,93	38,67	42,16
San Martín		46,83	46,83		37,80	37,80		40,89	40,89		38,21	38,21
Tacna	66,60	56,79	57,56	69,01	50,64	53,29	60,17	44,94	50,18	65,33	45,52	52,02
Tumbes		53,62	53,62	68,52	40,37	42,22	63,53	39,28	42,72	42,72	35,45	36,48
Ucayali		44,81	44,81		35,56	35,56		34,51	34,51		36,51	36,51
<b>Total</b>	64,73	53,18	54,54	65,28	46,03	48,90	59,99	43,37	46,38	61,44	43,43	46,68

variables individuales y familiares introducidas han logrado reducir 11%<sup>34</sup> de la variabilidad entre colegios (en adelante nivel 2).

De las variables de dicho modelo que están relacionadas con los antecedentes individuales, el género tiene un efecto negativo y significativo de 1,24 puntos porcentuales. Esto quiere decir que, en promedio, en todas las escuelas y controlando por el resto de variables del modelo, los hombres tienen 1,24 puntos más que las mujeres en la prueba de rendimiento en Matemática. De la misma forma, se observa que los estudiantes que caminan más de una hora para llegar a sus escuelas

tienen en promedio aproximadamente 3,59 puntos porcentuales menos en la prueba de Matemática. Otros efectos significativos interesantes son que rinden más los estudiantes a quienes les gusta la Matemática, quienes estudian en lugares cómodos y quienes entienden Matemática. Es importante destacar que en este nivel no todas las relaciones son lineales. Siempre controlando por las variables de dicho modelo, en el caso de repetición y número de días que hace tareas a la semana, se puede observar que mientras la primera tiene una relación significativa en forma de "U" (rinden significativamente más aquellos que no han repetido o han repetido mucho), la segunda la tendría en forma de "U" invertida (rinden notoriamente más aquellos que hacen tareas sólo de forma regular).

Mientras las variables demográficas no muestran resultados significativos interesantes,

34 Este porcentaje se calcula restando la varianza del nivel 2 en el modelo nulo de la varianza del nivel 2 en el nuevo modelo, y dividiendo luego ese resultado entre la varianza del modelo nulo.

Cuadro 14b

	Cuarto de primaria			Sexto de primaria			Cuarto de secundaria			Quinto de secundaria		
	CE	CE	Total	CE	CE	Total	CE	CE	Total	CE	CE	Total
	privados	estatales		privados	estatales		privados	estatales		privados	estatales	
<b>Lenguaje</b>												
Amazonas	48,78	52,13	51,89	47,33	49,83	49,56	40,02	45,37	44,69	43,63	52,25	51,16
Ancash	66,50	52,54	55,09	68,57	53,85	56,74	52,97	40,62	43,31	62,02	51,31	53,84
Apurímac		39,62	39,62		35,83	35,83	48,45	36,88	38,16	52,96	44,57	45,42
Arequipa	67,85	59,43	61,51	70,79	59,27	61,89	55,96	46,07	48,67	61,32	55,14	56,84
Ayacucho	78,00	42,92	45,32	80,88	42,19	46,70	55,18	33,36	37,69	68,99	47,07	51,20
Cajamarca	59,57	49,97	52,42	64,90	50,33	53,03		39,63	39,63		52,23	52,23
Callao	76,59	61,04	65,49	75,50	64,70	68,02	58,77	49,69	52,67	68,89	57,72	61,96
Cusco		44,08	44,08	66,24	47,39	49,58	52,39	37,37	39,71	65,83	47,65	50,56
Huancavelica		45,90	45,90		46,78	46,78	46,61	34,78	36,10	55,36	45,10	46,00
Huánuco	61,16	49,78	51,08	65,57	44,59	46,53	57,80	38,21	40,52	66,71	50,13	52,13
Ica	55,89	52,50	52,65	61,58	51,16	52,06	54,43	40,52	41,98	68,03	50,69	52,41
Junín	64,37	54,28	54,93	61,86	56,90	57,39	52,56	44,21	46,05	62,28	55,07	56,66
La Libertad	66,52	56,96	58,78	70,97	51,57	55,93	58,44	43,52	46,14	65,43	52,82	55,03
Lambayeque	55,71	56,45	56,33	63,98	54,35	56,36	51,80	42,09	43,55	58,03	54,16	54,73
Lima	74,44	56,77	59,87	75,90	58,59	62,91	62,56	45,28	49,87	70,93	54,74	59,00
Loreto	59,22	39,52	42,62	62,84	47,14	49,13	48,26	36,48	37,81	61,79	50,27	51,56
Madre de Dios		46,08	46,08		45,79	45,79		37,87	37,87		50,68	50,68
Moquegua	73,74	60,70	62,32	71,60	61,12	62,34	61,42	44,99	47,10	67,46	58,69	59,56
Pasco	59,78	47,06	49,20	56,89	47,60	48,51	44,40	42,57	42,81	51,02	49,93	50,06
Piura	61,28	52,07	52,89	58,48	54,08	54,73	53,13	37,93	40,83	61,59	51,84	53,83
Puno	67,34	42,51	46,35	77,63	41,32	46,20	55,44	34,83	37,53	65,56	44,28	47,12
San Martín		46,01	46,01		41,80	41,80		38,82	38,82		46,94	46,94
Tacna	72,14	56,18	57,42	72,77	59,42	61,34	55,12	45,59	48,81	67,99	56,71	60,40
Tumbes		51,21	51,21	74,91	45,62	47,55	55,13	40,45	42,50	55,20	44,50	46,04
Ucayali		45,16	45,16		45,69	45,69		35,69	35,69		47,92	47,92
<b>Total</b>	68,30	51,81	53,76	70,87	51,69	54,56	56,91	41,41	44,19	65,59	51,60	54,12

lo contrario ocurre con las variables culturales. Sin tomar en cuenta la variable socioeconómica, se observa que, en promedio para todas las escuelas, los hijos que no reciben ayuda de sus padres tienen en promedio 1 punto porcentual más que aquellos que sí la reciben. De la misma manera, se puede apreciar que un libro más en la casa produce en promedio un cambio significativo de 0,36 puntos porcentuales en la prueba de Matemática, controlando por el resto de variables. Finalmente, las diferencias entre los estudiantes cuyos padres aspiran a que terminen la universidad y aquellos que aspiran sólo a que terminen la primaria es significativa, al igual que la diferencia entre los primeros y aquellos cuyos padres aspiran a que se queden en superior no universitaria. Dado que la categoría de referencia es la primera, se tiene, por ejemplo, que, en promedio, los hijos cuyos

padres aspiran a que lleguen a la universidad en lugar de quedarse en superior no universitaria tienen 0,51 puntos porcentuales más en la prueba.

El modelo 2 introduce la variable socioeconómica. Se puede observar que mientras las variables de antecedentes individuales y demográficos no sufren cambios importantes, en el caso de las variables culturales y socioeconómicas sí se notan alteraciones. Lo más interesante es la neutralización de los efectos de la variable expectativas más no de la relacionada con el número de libros. Se observa, asimismo, que, en promedio, el cambio en una desviación estándar de la escala socioeconómica produce un cambio de 1,07 puntos porcentuales en la prueba de Matemática.

Por otro lado, es interesante observar que la nueva proporción de varianza del nivel 2

explicada es de 17,4%. Se aprecia, finalmente, que el promedio general en Matemática es de 54,44. Dicho promedio varía significativamente entre escuelas, por lo que se pueden introducir variables de dicho nivel para poder explicar estas diferencias.

El modelo 4 introduce variables a ese nivel. En primer lugar, la variable institucional de la gestión del colegio. Se puede observar que controlando por las variables del nivel 1, las escuelas privadas tienen aproximadamente 11 puntos porcentuales más en sus promedios de Matemática que las públicas. Al mismo tiempo, se puede notar cómo el efecto del nivel socioeconómico se reduce en la medida en que se controla parte de su impacto contextual. Este modelo tiene una varianza de 70,73 en el nivel 2 y, por lo tanto, explica 28% de dicha variabilidad.

El modelo 5 introduce variables relacionadas con características sociales de las escuelas. Se observa cómo el efecto de la gestión se reduce considerablemente al introducir las nuevas variables, lo que quiere decir que las diferencias entre escuelas públicas y privadas tienen que ver mucho con diferencias entre sus características sociales. Así, en promedio, un punto porcentual más en la escala de desnutrición de las escuelas reduce significativamente su promedio en 0,22. Asimismo, las escuelas de la costa y la sierra tienen significativamente mejores promedios en Matemática que las de la selva. Por ejemplo, controlando por el resto de variables de los niveles 1 y 2, la diferencia entre las escuelas de la costa y la selva es de aproximadamente 6,43 puntos porcentuales. Este modelo explica 41% de la variabilidad del nivel 2.

El modelo 6 introduce las características de las escuelas y profesores, y logra explicar 50% de las diferencias entre aquellas. Es interesante observar que las diferencias entre colegios públicos y privados se llegan a reducir en más de la mitad respecto del dato inicial. Otra diferencia que se reduce de manera importante es la que se da entre escuelas de la costa y la selva. Estos nuevos datos indicarían que dichas diferencias tienen que ver mucho con las características de las escuelas y profesores. De esas variables notamos las siguientes diferencias significativas (considerando en este caso 5% y 10% de probabilidad de error,

por el menor número de casos). En relación con las características personales y profesionales de los profesores y directores, y controlando por el resto de variables del modelo, se observa que las escuelas cuya profesora de Matemática era una mujer obtienen en promedio 1,96 puntos porcentuales más que aquellas cuyo profesor era un varón. Los años de servicio del profesor de Matemática muestran una relación significativa en forma de "U", lo que quiere decir que, controlando por el resto de variables, las escuelas con mejores rendimientos son aquellas en las que el profesor de Matemática tiene pocos o muchos años de servicio. Otras variables con resultados significativos son las horas que se dictan de Matemática y la titulación del director. Aquellas escuelas con directores titulados en institutos pedagógicos o universidades tienen mejores promedios (2,50 y 2,52 puntos adicionales respectivamente) que aquellas cuyos directores no se han titulado.

En el caso de las características pedagógicas y curriculares, se destaca que las escuelas en las que los profesores cuentan con el nuevo currículo tienen menor promedio que aquellas que no lo tienen.

En relación con los recursos de infraestructura y materiales, se destaca el hecho de que más importante para el rendimiento es tener buena infraestructura o biblioteca que contar con los materiales de aula. Controlando por las otras variables, las escuelas cuyo director dice tener problemas de infraestructura o de mobiliario rinden significativamente menos que aquellas cuyo director dice no enfrentar estas dificultades. De igual manera, las escuelas en las que, según el profesor, las aulas cuentan con biblioteca, rinden 1,70 puntos porcentuales más que aquellas que, según los profesores encuestados, no la tienen.

En relación con el clima institucional, sólo la variable que mide reuniones entre profesores ejerce un efecto significativo. Controlando siempre por el resto de variables del modelo, aquellas escuelas en las que, según el profesor, hay reuniones entre docentes, tienen 0,87 más puntaje promedio que las escuelas donde se reporta que no hay reuniones. Finalmente, la variable tamaño de la clase también tiene un efecto significativo y positivo sobre el rendimiento.

El modelo final incluye la variable promedio socioeconómico de la escuela. Esta variable permitirá distinguir cuánto del efecto socioeconómico ocurre fuera de las escuelas y cuánto es parte del contexto. Asimismo, hará posible controlar mejor los efectos de las escuelas y profesores. Se observa que este modelo llega a explicar 60% de las diferencias entre promedios de las escuelas. Adicionalmente, se nota que las diferencias entre promedios escolares en Matemática continúan siendo significativas en torno del promedio general de 54,30. Esto quiere decir que futuros análisis deberán considerar otras variables para poder explicar aún más esa variabilidad.

En primer lugar, se observa que el efecto neto de la variable socioeconómica (más allá de las diferencias individuales y familiares del nivel 1 y contextuales del nivel 2) es 0,63. Esto quiere decir que, en promedio y controlando por todas las variables del modelo, un cambio en una desviación estándar en el índice socioeconómico de las familias produce un cambio de 0,63 puntos porcentuales en el puntaje en Matemática. Por otro lado, se observan cambios importantes en las variables de escuela y profesores. En el caso de las variables sociales e institucionales, se aprecia la desaparición del efecto de la gestión y del efecto desnutrición, lo que nos indica que ambas diferencias se debían también en gran parte a diferencias en los promedios socioeconómicos de las escuelas. Las diferencias regionales también se reducen, pero continúan ejerciendo efectos significativos independientes. Finalmente, se observa que un cambio en una desviación estándar en el promedio socioeconómico de las escuelas produce un cambio significativo de aproximadamente 5 puntos porcentuales en el rendimiento promedio de Matemática.

En el caso de las variables relacionadas con características profesionales y personales de profesores y directores, hay cambios en las magnitudes de los coeficientes, pero no en sus poderes estadísticos, salvo la desaparición del efecto de la titulación universitaria del director.

Las variables de recursos, infraestructura y materiales sí pierden poder explicativo. Esto último nos indica la dificultad de dichas va-

riables para ejercer efectos independientes del socioeconómico. Sólo se sostiene el efecto de tener biblioteca, aunque se reduce su magnitud. Finalmente, desaparecen tanto los efectos relacionados con el clima de las escuelas como aquel relacionado con el tamaño del aula.

Antes de finalizar este análisis de resultados cabe señalar que, dado que la mayoría de las relaciones con el rendimiento sugeridas por los estudiantes y sus familias no variaban significativamente entre escuelas (es decir, estadísticamente se asume que dichas relaciones son las mismas en todos los colegios), se decidió mantener fijos sus coeficientes. Sólo la relación entre el rendimiento y el género, distancia de la escuela, gusto por la matemática y apoyo en las tareas variaba significativamente entre escuelas. Se hizo la prueba de considerar esas relaciones como aleatorias. En todos los casos, salvo el de género, los coeficientes de confiabilidad eran menores de 0,10, y esas diferencias no se podían explicar con las variables con las que se contaba. Para no introducir problemas de dependencia relacionados con el hecho de que no se puede controlar esa variabilidad, se decidió mantener fijos esos coeficientes. Sólo se pudo explicar la relación entre género y rendimiento, y se encontró que en aquellas escuelas donde los profesores son mujeres, la relación cambia de sentido y las mujeres logran mejores rendimientos. No obstante, no se reportan esos datos, dado que el índice de confiabilidad es igualmente bajo (menor de 0,15).

## 5. CONCLUSIONES

En el informe presentado por el Banco Mundial<sup>35</sup>, al analizar los datos de 1996 se llegaba a la conclusión de que 43% de las diferencias en Matemática se debía a diferencias relacionadas con el hecho de que los estudiantes estuvieran agrupados en las escuelas. Usando un modelo diferente, en este trabajo se ha llegado a una conclusión parecida. En este caso, 41% de dichas diferencias se debe a los contextos escolares. Ha sido interesante

35 The World Bank: *Peru Education...*, op. cit., 1999.

**Cuadro 15**  
**Resultados de rendimiento de Matemática usando modelos jerárquicos lineales: Modelos de rendimiento individual y promedio de escuelas**

Variables individuales y familiares	Modelo nulo	Modelos de rendimientos individuales		Modelos de rendimientos individuales y promedios de escuelas			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Antecedentes individuales del alumno</b>							
Agenero		-1,24 *	-1,23 *	-1,23 *	-1,22 *	-1,23 *	-1,23 *
Asiste		2,31 *	2,28 *	2,26 *	2,24 *	2,23 *	2,21 *
Prinido		0,37	0,09	0,02	-0,02	-0,06	-0,12
Pubnido		0,28	0,10	0,10	0,09	0,07	0,06
Repite		-1,15 *	-1,02 *	-1,01 *	-0,99 *	-0,98 *	-0,94 *
Repite2		0,27 *	0,23 *	0,23 *	0,24 *	0,23	0,22 *
Camino		-3,59 *	-3,51 *	-3,50 *	-3,50 *	-3,50 *	-3,49 *
Libmate		0,007	0,07	0,32	0,34	0,34	0,37
Tareama		336 *	338 *	338 *	331 *	326 *	329 *
Tareama2		-0,63 *	-0,63 *	-0,63 *	-0,62 *	-0,61 *	-0,61 *
Trambie		0,69 *	0,71 *	0,71 *	0,70 *	0,69 *	0,68 *
Gustmat		1,65 *	1,68 *	1,69 *	1,70 *	1,71 *	1,70 *
Entiemat		1,42 *	1,44 *	1,44 *	1,45 *	1,44 *	1,44 *
<b>Características demográficas de la familia del alumno</b>							
Madreviv		0,50	0,37	0,35	0,33	0,32	0,35
Avivcon		-0,26	-0,32**	-0,32	-0,33	-0,33	-0,34**
Totmiem		-0,01	-0,02	-0,02	-0,02	-0,02	-0,02
<b>Características socioeconómicas y culturales de los padres</b>							
Tayuda1		1,02 *	1,06 *	1,06 *	1,05 *	1,06**	1,07 *
Pexp1		-0,88 *	-0,76**	-0,76**	-0,73**	-0,74	-0,70
Pexp2		-0,28	-0,00	-0,02	-0,02	-0,01	0,00
Pexp3		-0,51 *	-0,35**	-0,34**	-0,35**	-0,34**	-0,33**
Hhlibros		0,36 *	0,20 *	0,19 *	0,19 *	0,19 *	0,17**
Innsse2			1,07 *	1,01 *	0,92 *	0,86 *	0,63 *
<b>Variables de la escuela y profesores</b>							
Intercepto o promedio general	54,41 *	54,43 *	54,44 *	54,44 *	54,39 *	54,33 *	54,30 *
<b>Características sociales e institucionales</b>							
Gestion				-10,62 *	-7,70 *	-3,15 *	0,06
Nutri					-0,22 *	-0,17 *	-0,02
Costa					6,43 *	4,61 *	3,74 *
Sierra					4,97 *	4,66 *	3,27**
Cinnsse2							517 *

Variables individuales y familiares	Modelo nulo	Modelos de rendimientos individuales		Modelos de rendimientos individuales y promedios de escuelas			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Características personales y profesionales de los docentes y directores</b>							
Ppmotiv						0,76	0,79
Ppgener						1,96*	1,82*
Otrabaj2						-0,46	-0,51
Pptutil3						0,71	0,36
Pptutil4						1,66	0,92
Ppservi						-0,33**	-0,40*
Ppservi2						0,01*	0,01*
Pphrpcl						0,33	0,24
Hrsemat						0,30**	0,47*
Dtitul3						2,50*	1,58**
Dtitul4						2,52*	1,77
<b>Características pedagógicas y curriculares</b>							
Ppncur						-1,82**	-1,85**
Ppntrad						-0,09	1,95
Gxppntr						-0,15	-1,18
<b>Recursos de materiales e infraestructura</b>							
Matmate						0,23	-0,04
Material						0,69	0,59
Dmobilia						-1,28**	-1,02
Dinfras						-1,41**	-0,84
Ctecho						1,69**	0,44
Biblio						1,70*	1,38*
<b>Clima institucional</b>							
Probce4						-0,88	-0,30
Probc12						0,97	0,40
Ppreunp						1,28*	0,47
Ppreund						-0,41	-0,14
Dapyfam						-0,71	-0,84
<b>Otras características</b>							
Totclas						0,11*	-0,01
Varianza Nivel 1	140,91	135,18	135,12	135,10	135,06	135,05	135
Varianza Nivel 2	98,33*	86,84*	81,20*	70,73*	58,06*	49,78*	40*

\* Significativa al 5%.

\*\* Significativa al 10%.

observar cómo gran parte de esas diferencias responde a factores individuales. Se ha visto cómo la introducción de estos últimos —en particular la variable socioeconómica— afecta

los dos niveles de varianza. Esto indicaría dos cosas. Primero, la importancia de la variable socioeconómica para estimar adecuadamente no sólo los efectos al interior de las escuela-

las, sino también aquellos relacionados con los contextos escolares. En segundo lugar, con este tipo de análisis se diluye la distinción entre lo alterable y lo no alterable. De alguna manera, se observa que parte de los efectos de la relación entre la variable socioeconómica y el rendimiento estudiantil opera en el nivel contextual, escolar en este caso, nivel que sí es posible modificar desde políticas educativas.

Por otro lado, es interesante observar la importancia de los efectos individuales. En primer lugar, el de la variable socioeconómica. Controlando incluso por la naturaleza socioeconómica de la propia escuela, esta variable sigue teniendo efectos. Es decir, tanto en colegios pobres como en colegios ricos, y más allá de cualquier otra diferencia presente en el modelo, existe un efecto neto en el rendimiento relacionado con las características socioeconómicas de las familias.

Asimismo, destaca que mientras la relación entre expectativas y rendimiento es absorbida casi por completo por la variable socioeconómica, no sucede lo mismo con las variables relacionadas con el número de libros y el apoyo en las tareas. Tal como las teorías del capital cultural y el capital social postulan, estas variables operan de forma independiente al nivel socioeconómico.

Se distinguen otros efectos importantes y que pueden ser corregidos desde las escuelas a través de programas especiales. Primero, la presencia de las diferencias significativas a favor de los hombres en el área de primaria y en cuarto grado. Este efecto es importante, e incluso se llegó a comprobar en modelos adicionales que dicho efecto podía ser alterado por el género del profesor. Es decir, que en colegios donde la profesora es una mujer, rinden más las mujeres.

Otras variables individuales cuyos efectos se han mantenido de forma sistemática y sobre los cuales hay posibilidades de intervención desde políticas educativas son las del gusto por la Matemática, la asistencia a la escuela, la distancia de la escuela y el número de días que hace tareas. En el caso de esta última, valdría la pena hacer un estudio más detallado para poder saber a ciencia cierta si su efecto es positivo o negativo en relación con el rendimiento. La naturaleza no longitu-

dinal de este estudio no nos permite saber si las tareas ayudan a mejorar el rendimiento de los niños. Únicamente sabemos que sólo quienes hacen tareas de forma regular tienen mayores rendimientos en esta prueba y en comparación con los otros estudiantes.

Por otro lado, en el caso de las variables de las escuelas, es interesante observar cómo el efecto de la gestión desaparece por completo una vez que se incluyen características sociales y variables relacionadas con las escuelas y los profesores. Esto señalaría que las diferencias entre escuelas públicas y privadas tienen poco que ver con la naturaleza administrativa y mucho más con diferencias en recursos sociales y de calidad educativa. Al mismo tiempo, vuelve a señalar el papel que pueden jugar las políticas educativas en la neutralización de efectos relacionados con características sociales. Si bien modificar el promedio socioeconómico de la escuela puede ser difícil —aunque algunos consideran que esto puede lograrse a través del mejoramiento del gasto por alumno—, mucho más manejable es la realización de campañas de nutrición. Se ha visto cómo esta variable —relacionada de forma importante con el promedio socioeconómico— puede llegar a reducir diferencias en promedios de escuelas.

Es sorprendente no haber encontrado efectos interesantes en las variables relacionadas con las escuelas y los profesores, una vez que son controladas las diferencias sociales de aquéllas. Estos resultados sólo destacan el género del profesor, las horas de dictado de Matemática y los años de servicio del profesor como variables importantes de los profesores. En el caso de las escuelas, sólo destaca el hecho de contar con biblioteca como variable importante. Si bien los resultados relacionados con la formación del profesor son consistentes con la literatura sobre el tema, es sorprendente que la variable relacionada con las prácticas pedagógicas no haya tenido resultados sobresalientes con la variable rendimiento en ninguno de los modelos. En todo caso, queda planteada la pregunta de si los nuevos métodos pedagógicos permiten mejores rendimientos en los alumnos, en comparación con la aplicación de métodos tradicionales. Asumiendo que las variables recogen adecuadamente lo que sucede dentro del

aula, una posible explicación y respuesta es que esas nuevas prácticas pedagógicas pueden haberse generalizado casi por completo, y en ese caso ya no habría diferencias que pudieran explicar el rendimiento.

A la luz de estos resultados, cabe preguntarse cuáles deberían ser los objetivos de la política educativa. Pensar que el MED sólo debe ocuparse de aquello que sucede dentro de las aulas sería un razonamiento equivocado. Políticas dirigidas a despertar el gusto por la Matemática, a construir escuelas de calidad que se ubiquen cerca de las residencias de los alumnos o a estimular la lectura en las familias, podrían ser bastante efectivas para mejorar el rendimiento de los niños. Lo mismo puede decirse de políticas que

se pueden realizar en las escuelas, pero que modifican características sociales. Un ejemplo de ello pueden ser las políticas dirigidas a reducir la desnutrición de los niños de las escuelas.

Finalmente, es necesario insistir en la importancia de diseñar pruebas longitudinales que permitan estimar mejor los efectos y, al mismo tiempo, responder con mayor certeza preguntas sobre cambios en rendimientos individuales. En este caso, los resultados están limitados por la naturaleza no longitudinal del diseño. Mientras los diseños longitudinales responden a inquietudes de cambio, este tipo de diseño se plantea la explicación de las diferencias entre estudiantes en un momento fijo en el tiempo.

# DISEÑO MUESTRAL EN LA APLICACIÓN NACIONAL CRECER 1998

ARTURO CALDERÓN  
CHOLLY FARRO  
JORGE BAZÁN

## 1 ANTECEDENTES

La Unidad de Medición de la Calidad Educativa (UMC) del Ministerio de Educación tenía programadas, como parte de las actividades del Proyecto CRECER 1998, diversas evaluaciones sobre determinadas áreas de desarrollo del currículo de educación primaria. El plan general de trabajo de la Unidad contemplaba que en 1998 estas evaluaciones debían ser aplicadas a alumnos de cuarto y sexto grado de primaria y cuarto y quinto de secundaria. Aun cuando en esta sección<sup>1</sup> nos dedicaremos a describir la metodología aplicada para primaria, cabe señalar que también fue seguida en el diseño y etapas de muestreo para secundaria. Lo que puede variar son los datos numéricos específicos por grado.

Así, las poblaciones objetivo del estudio para los diferentes grados fueron definidas como los alumnos de educación primaria escolarizada, estudiantes de centros educativos (CE) polidocentes completos de áreas urbanas, tanto estatales como no estatales. En cuarto grado esta población era de 407 243 alumnos y en sexto grado de 364 301, todos distribuidos en 6187 CE.

Dada la extensión y distribución de la población, la UMC suele usar un sistema de aplicaciones muestrales de las pruebas en lugar de un sistema censal. Por ello había que aplicar un diseño general que proporcionara muestras representativas y eficientes, con erro-

res de estimación dentro de márgenes razonables y alta confiabilidad.

Adicionalmente, el estudio planteaba recopilar datos relativos a los llamados “factores asociados al rendimiento”, y para ello se definieron tres poblaciones: la de padres de familia de los alumnos, la de profesores de los alumnos y la de directores de los CE donde estudiaban los alumnos de la población objetivo. Estas poblaciones no fueron objeto especial de diseño y estuvieron subordinadas a la población principal (la de los estudiantes que darían las pruebas). Sin embargo, como el procedimiento para obtener datos de estas poblaciones asociadas consistió en entrevistar a los padres, profesores y directores, los tamaños de muestra resultaron suficientemente grandes como para asegurar que en general y en el nivel agregado total las estimaciones que se derivaran de ellas tendrían un margen de error no mayor de 5%<sup>2</sup>. Así, se previó entrevistar a dos tercios de padres de familia, lo que permitiría errores me-

---

1 Actualización del informe preparado por Arturo Calderón y José Flores, quienes diseñaron las muestras de CRECER 1998. Miguel Acevedo, de la UMC, participó en la implementación de este diseño.

2 Este 5% alude a un error estimado mediante una muestra aleatoria simple, que no fue el empleado en CRECER 1998; no obstante, al haberse usado una muestra compleja y ser dicho valor bastante conservador, se esperaba que en la práctica fuera menor.

nores de 3%<sup>3</sup>. En lo que respecta a los profesores y directores ya no sería posible desagregar por estratos (como departamentos, por ejemplo), pues las muestras correspondientes resultarían muy pequeñas. Como en el caso de las poblaciones asociadas lo que importaba eran las correlaciones que se pudieran detectar entre sus características y el rendimiento a partir de las pruebas más que la estimación de parámetros específicos de las poblaciones, el tamaño de las muestras no representaba un problema.

## 2. OBJETIVO

El objetivo general del diseño era obtener una muestra representativa de alumnos de cuarto y sexto grado de educación primaria de menores, de modalidad escolarizada, estudiantes de CE polidocentes completos urbanos en 1998. Esta muestra debía permitir estimaciones del rendimiento en diversas materias, tanto globales como por región, con confiabilidad de 95% y errores de estimación no mayores de 5%. Además, las muestras debían posibilitar comparaciones entre algunos estratos predefinidos. En las líneas que siguen desarrollamos el esquema de elaboración de la muestra y presentamos los resultados.

La evaluación proporcionaría información agregada, es decir, no para cada alumno ni para cada escuela, docente o director. Para que esto hubiera sido posible se habría necesitado una muestra suficientemente grande por población objetivo, y la investigación no pretendía dar resultados para cada una.

La información que se esperaba recoger en CRECER 1998 podría haberse obtenido mediante una aplicación censal de pruebas y encuestas o a través de la aplicación de los mismos instrumentos a una muestra de la población. La magnitud de una aplicación censal suele implicar un altísimo costo, y muchas veces hasta puede resultar innecesaria. Así, se pensó que lo más aconsejable era tomar mues-

tras representativas de cada subpoblación y para cada grado de estudio, de manera que pudieran ser controladas y proporcionaran información de calidad.

En general, trabajar con muestras implica cierto riesgo de no representatividad o dificultades al recoger los datos. Estas dificultades se pueden clasificar en dos grandes grupos:

- a) *Dificultades debidas al muestreo*: derivadas de tratar de representar una población a partir del estudio de parte de ella (o sea de una muestra). Esto se puede controlar con un adecuado diseño muestral.
- b) *Dificultades en la toma de datos*: surgen en la aplicación de las pruebas y encuestas por errores al registrar los datos, reemplazo de casos, modificación de respuestas, etcétera. Estos problemas son sumamente serios y se controlan con una capacitación eficiente y un trabajo comprometido de examinadores, coordinadores y supervisores.

## 3. DISEÑO MUESTRAL

En diseños simples, las unidades de muestreo o unidades a seleccionar son las mismas que se van a analizar y, por lo tanto, constituyen el único objeto de estudio. En diseños complejos —los más empleados en la práctica— las unidades de muestreo no coinciden necesariamente con las unidades de análisis. Éste es el caso del muestreo en dos etapas o bietápico.

En CRECER 1998 se empleó conjuntamente muestreo estratificado y de conglomerados. Los conglomerados están formados por grupos de elementos que son unidades de análisis colectivas y no individuales. En nuestro caso cada escuela era un conglomerado, pues estaba integrada por un grupo de aulas y éstas, a su vez, por un conjunto de alumnos. Por otro lado, al ser también un muestreo estratificado, se combinaron los criterios de gestión (CE estatal o no estatal), región y departamento para formar los estratos requeridos. Así se obtuvieron 88 combinaciones, de las cuales 62 contaban con datos<sup>4</sup>, es decir, información a partir del cruce de la gestión y región para cada departamento. En cada uno

3 *Idem* nota 1.

4 Por ejemplo, un estrato vacío fue la combinación del departamento de Madre de Dios, región selva, escuelas no estatales.

de éstos se seleccionaron escuelas en forma proporcional a sus tamaños, medidos a partir del número de alumnos por grado.

Desde el punto de vista de la selección, el muestreo empleado fue bietápico. Las unidades primarias de muestreo (UPM) fueron las escuelas, y las unidades de muestreo de la segunda etapa (USM), los alumnos de dichas escuelas.

En suma:

- a) La muestra originalmente diseñada está dentro de un sistema de muestreo estratificado bietápico y de conglomerados, con las escuelas como unidades de primera etapa y los alumnos como unidades de segunda etapa.
- b) Se trata de una muestra con probabilidad de selección proporcional al tamaño del CE, que asegura igual oportunidad de formar parte de la muestra a todos los alumnos y requiere el uso de ponderaciones en las estimaciones que se hagan, dado que la distribución de la muestra no es proporcional a la distribución de la población por estratos.
- c) Se usó este esquema de muestreo (afijación no proporcional) porque se necesitaba garantizar comparabilidad y representatividad entre estratos, y una distribución proporcional de la muestra la concentraría en Lima y algunas otras ciudades, dejando sin casos a los demás estratos y aumentando el error de estimación en ellos. La única manera de mantener proporcionalidad y, a la vez, tamaños de muestra suficientemente grandes en los estratos, era usar una muestra global de unos 50 000 alumnos por grado, pero el costo de hacer esto no sólo era económico, sino sobre todo de pérdida de capacidad de administración eficiente de las muestras. La labor de la UMC en el proceso de 1996, cuando se manejó una muestra de 50 000 casos en cuarto grado, indicaba que había experiencia y capacidad suficiente para administrar muestras grandes, pero como en 1998 se tenía que hacer aplicaciones en dos grados de primaria y otros dos de secundaria, la cantidad de alumnos por evaluar subía a 200 000; como, además,

había de por medio entrevistas adicionales a padres de familia, profesores y directores, era claro que los costos y riesgos aumentaban sobremedida.

- d) Con el fin de restituir la representatividad de los estratos de la muestra en el universo, se usaron ponderaciones<sup>5</sup>, calculadas según los criterios básicos de estratificación empleada. El uso de ponderaciones no altera la calidad de las estimaciones; por el contrario, las hace más fiables y consistentes. En el anexo se explica en detalle la fórmula de ponderación utilizada.
- e) Aunque en primaria eran dos las poblaciones objetivo (cuarto y sexto grado), por tratarse de CE polidocentes completos se optó, para evitar errores de logística, por un diseño que contemplaba el muestreo simultáneo en ambas poblaciones. Así, las muestras de cuarto y sexto se tomaron en los mismos centros.

#### **4 ESTRUCTURA GENERAL DE LA MUESTRA**

##### **LOS ESTRATOS**

Tomando en cuenta la experiencia de años anteriores, se definieron estratos o subdivisiones de la población de alumnos cuya presencia en la muestra aseguraba representatividad y facilitaba la aplicación de las pruebas y encuestas. Los criterios para la formación de estratos fueron:

*Departamento:* veinticinco, incluyendo la Provincia Constitucional del Callao. Todos estarían presentes en la muestra global. Hubo una submuestra por cada departamento, con un mínimo de 450 alumnos en cada caso, para garantizar un margen de error menor de 5% en la estimación del rendimiento en el departamento.

*Región natural:* Participarían las tres regiones naturales (costa, sierra y selva), con el fin de asegurar un mínimo de 1500 casos para cada región, de modo que hubiera un margen de error menor de 3% en la estimación correspondiente.

5 Véase el anexo.

*Gestión*: tanto los CE estatales como los no estatales figuran en la muestra en proporción a su participación en la estructura general de la matrícula de alumnos. En este estrato el margen de error calculado es menor de 3%.

*Lima Metropolitana*: fue considerada como un estrato aparte, separado del estrato costa.

Adicionalmente, se consideró el desagregado por ubicación geográfica (norte, centro y sur), pero este criterio se mantuvo sólo en la costa, por la presencia de Lima, y en los demás casos se reservó a una post-estratificación.

Como ya se mencionó, la estratificación combinó criterios como la gestión (CE estatal o no estatal), región y departamento para formar los estratos requeridos. Así se obtuvieron 88 combinaciones, de las cuales 62 contaban con datos, es decir, información a partir del cruce de la gestión y región para cada departamento. Dentro de cada uno de éstos se seleccionaron escuelas en forma proporcional a sus tamaños, medidos con base en el número de alumnos por grado.

El tamaño de la muestra nacional se calculó de modo que hubiera un margen de error menor de 1% en la estimación general y que permitiera analizar otras estratificaciones (por ejemplo, región natural transversal norte, centro y sur, o género), que, sin ser parte determinante del diseño, brindarían información útil.

La representatividad de la muestra fue definida de modo que todos los alumnos de cada población tuvieran igual oportunidad de participar, tomando en cuenta la magnitud de cada estrato y del CE al que pertenecía el alumno, estimado a partir del número de éstos en los grados considerados. Este sistema de estratificación y asignación de tamaños mínimos de muestra por estrato se aplicó a todas y cada una de las cuatro poblaciones (cuarto y sexto grado de primaria y cuarto y quinto de secundaria).

## **LAS CUOTAS DE MUESTREO Y PRECISIONES**

Todos los estratos principales se cruzaron y se obtuvieron 88 combinaciones o “estratos

de sorteo”, de modo que el muestreo final permitía agregaciones y estimaciones 95% confiables tanto para el total nacional como por departamentos y por estratos, con márgenes de error no mayores de 5%, lo que facilitó las comparaciones entre estratos y departamentos. Las cuotas mínimas de alumnos fueron:

- a) En el ámbito nacional, la muestra sería no menor de 10 000 alumnos, cantidad suficiente como para tener un margen de error no mayor de un punto porcentual.
- b) Por departamento, la cuota mínima de casos sería de 500, con un margen de error no mayor de cuatro puntos porcentuales.
- c) Por cada estrato principal (que es agregado de sus estratos de sorteo), la cuota mínima sería de 1500 casos, lo que daba un margen de error (para cada nivel de estrato) menor de dos puntos porcentuales.
- d) La combinación de las diversas cuotas, cruzadas por estratos de sorteo, determinó un total de alrededor de 17 370 alumnos por grado. Por razones de diseño, para lograr que todos los alumnos de la población tuvieran igual oportunidad de integrar la muestra, se usó un sistema de sorteo de CE con probabilidad de selección proporcional al número de alumnos del grado correspondiente (que reflejaría el tamaño del CE), para lo cual se fijó una cuota de 30 alumnos por CE. Esto determinó una muestra de 579 CE en total.

## **SISTEMA DE SELECCIÓN**

En cada estrato se seleccionó una muestra probabilística de CE y luego en cada CE se escogió, mediante sorteo aleatorio simple, la sección y los alumnos que integrarían la muestra.

Los CE fueron elegidos mediante selección proporcional a su tamaño (a partir del número de alumnos), y luego en cada CE se escogió, también mediante sorteo, a los alumnos que integrarían la muestra.

En la primera etapa (selección de CE) los alumnos de cuarto y sexto grado de primaria

se seleccionaron en un mismo CE. Análogamente, los alumnos de cuarto y quinto de secundaria se escogieron también de un mismo CE. El tamaño del CE se tomó en cuenta en el diseño y fue usado en los análisis y estimaciones. El sorteo hace que la relación de CE sea representativa de su participación en el universo, pues aquellos colegios con más alumnos tendrían más probabilidad de ser seleccionados que aquellos más pequeños. Esto obliga a respetar la relación de CE seleccionados. Alterar la lista final del sorteo no mejora la muestra y le quita rigor al procedimiento efectuado; no obstante, en la práctica hubo que tomar algunas decisiones con carácter de fuerza mayor debido a problemas en campo —por ejemplo, reemplazando algunas escuelas—, pero respetando siempre que fueran de las mismas características que las originales.

La segunda etapa consistió en la selección de alumnos, previo sorteo de un aula, de haber más de una en el grado. El diseño pretendía retratar bien el rendimiento de un alumno “promedio” dentro de cada estrato, y su “representatividad” estuvo pensada en ese sentido. Por eso se estableció una misma cuota de alumnos por CE y por grado: 30 educandos.

Los 30 alumnos deberían ser seleccionados por sorteo, y no con base en su rendimiento, calificaciones o cualidades. Esto induciría a error y afectaría seriamente el estrato correspondiente y la muestra en general.

Por otro lado, fue necesario insistir en que, como se trataba de dos o más pruebas por alumno, además de las encuestas a los alumnos, padres de familia, profesores y directores, se debía asegurar la correspondencia entre todos los instrumentos, controlando que los datos de identificación coincidieran en todos los cuadernillos y hojas de respuestas. Esto estaba estrechamente vinculado al trabajo de organización y entrenamiento del personal que conformaría finalmente la red administrativa.

En suma:

a) En cada estrato de sorteo se determinó la cuota de CE a elegir (dividiendo entre

30 la cantidad de alumnos por seleccionar en el estrato y redondeando). Además, se fijó un mínimo de dos CE por estrato de sorteo para poder estimar el error de muestreo en todos los casos. Identificada la cantidad de CE por sortear, se los seleccionó mediante un programa computarizado que asignaba probabilidades de selección proporcionales al tamaño del CE (cantidad de alumnos en el grado).

b) La segunda etapa consistió en la selección de alumnos. Los 30 alumnos fueron seleccionados por sorteo aleatorio simple. En el caso de CE con más de una sección, se tuvo que hacer un sorteo simple de secciones para evitar complicaciones excesivas de logística. Esta etapa adicional del sorteo no sesga la muestra (sigue siendo representativa del CE), pero tampoco permite retirar del error de estimación el posible efecto de diferencias entre secciones del CE, si las hubiera, incrementando el margen de error. Pero es un riesgo calculado por el que se optó para controlar los errores “no de muestreo”.

## MUESTRA FINAL ESPERADA

Las consideraciones anteriores, aplicadas a los estratos, dieron una lista de 579 CE (después 578, porque en un estrato de muestreo de Madre de Dios no hubo suficiente cantidad de CE y se debió tomar más secciones en uno de ellos).

En las siguientes páginas presentamos las poblaciones objetivo y las muestras para los estratos empleados. Como ya se indicó, al tomar las muestras de cuarto y sexto en los mismos CE y con una tasa uniforme de 30 alumnos por CE, las muestras esperadas coinciden en número y estructura en ambos grados, aunque las muestras efectivas pueden haber diferido. La muestra esperada era de 17 370 alumnos en cuarto grado y 17 370 en sexto grado.

La representatividad de la muestra según criterios de estratificación fue la que aparece en los cuadros de la página siguiente.

<b>Población total (4° grado)</b>						
	<b>Estatal</b>		<b>No estatal</b>		<b>Total</b>	
	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>
Rural	22 222	262 426	357	4 197	22 579	266 623
Urbana	4 286	346 403	4 202	76 396	8 488	422 799
Nacional	26 508	608 829	4 559	80 593	31 067	689 422
<b>Polidocentes completos (4° grado)</b>						
	<b>Estatal</b>		<b>No estatal</b>		<b>Total</b>	
	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>
Rural	2 134	83 863	71	2 431	2 205	86 294
Urbana	3 646	337 347	2 485	67 858	6 131	405 205
Nacional	5 780	421 210	2 556	70 289	8 336	491 499
<b>Participación con respecto al total (%) (4° grado)</b>						
	<b>Total</b>					
	<b>Escuelas</b>			<b>Alumnos</b>		
Urbana	27,32			61,33		
Polidocentes completos	26,83			71,29		
Polidocentes completos urbanos	19,73			58,77		
<b>Población total (6° grado)</b>						
	<b>Estatal</b>		<b>No estatal</b>		<b>Total</b>	
	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>
Rural	22 222	176 433	357	3 290	22 579	179 723
Urbana	4 286	308 068	4 202	65 205	8 488	373 273
Nacional	26 508	484 501	4 559	68 495	31 067	552 996
<b>Polidocentes completos (6° grado)</b>						
	<b>Estatal</b>		<b>No estatal</b>		<b>Total</b>	
	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>
Rural	2 134	64 849	71	2 081	2 205	66 930
Urbana	3 646	301 583	2 485	61 218	6 131	362 801
Nacional	5 780	366 432	2 556	63 299	8 336	429 731
<b>Participación con respecto al total (%) (6° grado)</b>						
	<b>Total</b>					
	<b>Escuelas</b>			<b>Alumnos</b>		
Urbana	27,32			67,54		
Polidocentes completos	26,83			77,71		
Polidocentes completos urbanos	19,74			65,61		

Las coberturas esperadas de las muestras efectivas con respecto a las poblaciones respectivas son:

- 9,3% de los CE en general.
- 13,5% de los CE estatales.
- 3,4% de los CE no estatales.
- 4,3% de los alumnos de cuarto grado en general.
- 4,4% de los alumnos de cuarto grado de CE estatales.
- 3,7% de los alumnos de cuarto grado de CE no estatales.
- 4,8% de los alumnos de sexto grado en general.
- 4,9% de los alumnos de sexto grado de CE estatales.
- 4,1% de los alumnos de sexto grado de CE no estatales.

Distribución de la población de 4° grado de primaria por estratos						
Estratos	Estatales		No estatales		Total poblacional	
	CE	Alumnos	CE	Alumnos	CE	Alumnos
Costa Norte	592	58 153	339	12 667	931	70 820
Costa Central sin Lima	214	20 250	124	3 029	338	23 279
Costa Sur	101	7 917	40	1 253	141	9 170
Sierra Norte	326	19 001	33	1 064	359	20 065
Sierra Centro	458	35 392	113	2 772	571	38 164
Sierra Sur	658	49 396	288	8 771	946	58 167
Selva Norte	310	25 093	40	1 432	350	26 525
Selva Centro	111	13 452	22	548	133	14 000
Selva Sur	39	2 744	6	179	45	2 923
Lima Ciudad	848	106 741	1525	37 389	2373	144 130
<b>Total</b>	<b>3657</b>	<b>338 139</b>	<b>2530</b>	<b>69 104</b>	<b>6187</b>	<b>407 243</b>

Distribución de la muestra de 4° grado de primaria por estratos						
Estratos	Estatales		No estatales		Total poblacional	
	CE	Alumnos	CE	Alumnos	CE	Alumnos
Costa Norte	63	1 890	14	420	77	2 310
Costa Central sin Lima	44	1 320	6	180	50	500
Costa Sur	47	1 410	8	240	55	550
Sierra Norte	29	870	2	60	31	930
Sierra Centro	68	2 040	8	240	76	2 280
Sierra Sur	60	1 800	9	270	69	2 070
Selva Norte	46	1 380	6	180	52	1 560
Selva Centro	28	840	0	0	28	840
Selva Sur	19	570	0	0	19	570
Lima Ciudad	90	2 700	32	960	122	3 660
<b>Total</b>	<b>494</b>	<b>14 820</b>	<b>85</b>	<b>2550</b>	<b>579</b>	<b>17 370</b>

Distribución de la población de 6° grado de primaria por estratos						
Estratos	Estatales		No estatales		Total poblacional	
	CE	Alumnos	CE	Alumnos	CE	Alumnos
Costa Norte	592	50 212	339	11 203	931	61 415
Costa Central sin Lima	214	18 738	124	2 753	338	21 491
Costa Sur	101	7 122	40	1 221	141	8 343
Sierra Norte	326	16 716	33	942	359	17 658
Sierra Centro	458	31 991	113	2 499	571	34 490
Sierra Sur	658	44 520	288	8 062	946	52 582
Selva Norte	310	19 733	40	1 318	350	21 051
Selva Centro	111	11 122	22	502	133	11 624
Selva Sur	39	2 541	6	163	45	2 704
Lima Ciudad	848	99 797	1525	33 146	2373	132 943
<b>Total</b>	<b>3657</b>	<b>302 492</b>	<b>2530</b>	<b>61 809</b>	<b>6187</b>	<b>364 301</b>

Distribución de la muestra de 6° grado de primaria por estratos						
Estratos	Estatales		No estatales		Total poblacional	
	CE	Alumnos	CE	Alumnos	CE	Alumnos
Costa Norte	63	1 890	14	420	77	2 310
Costa Central sin Lima	44	1 320	6	180	50	500
Costa Sur	47	1 410	8	240	55	550
Sierra Norte	29	870	2	60	31	930
Sierra Centro	68	2 040	8	240	76	2 280
Sierra Sur	60	1 800	9	270	69	2 070
Selva Norte	46	1 380	6	180	52	1 560
Selva Centro	28	840	0	0	28	840
Selva Sur	19	570	0	0	19	570
Lima Ciudad	90	2 700	32	960	122	3 660
<b>Total</b>	494	14 820	85	2550	579	17 370

Poblaciones y muestras en primaria 1998 por departamento				
Departamentos	4° P. Población	4° P. Muestra	6° P. Población	6° P. Muestra
Amazonas	3 812	540	3 399	540
Ancash	15 731	540	13 951	540
Apurímac	7 195	510	5 756	510
Arequipa	19 528	870	18 622	870
Ayacucho	9 149	570	7 884	570
Cajamarca	9 491	510	8 343	510
Callao	14 695	510	13 531	510
Cusco	18136	540	16092	540
Huancavelica	3 713	510	3 239	510
Huánuco	9 003	510	8 511	510
Ica	12 889	840	12 087	840
Junín	18 087	510	16 484	510
La Libertad	22 405	540	19 501	540
Lambayeque	18 327	510	15 736	510
Lima	140 172	3 810	129 119	3 810
Loreto	10 913	510	7 837	510
Madre de Dios	1 565	510	1 411	510
Moquegua	2 091	510	2 190	510
Pasco	4 009	510	3 510	510
Piura	22 659	600	19 476	600
Puno	16 812	510	15 289	510
San Martín	10 463	510	8 624	510
Tacna	4 933	840	4 269	840
Tumbes	3 609	540	3 257	540
Ucayali	7 856	510	6 183	510
<b>Total</b>	407 243	17 370	364 301	17 370

**Población de alumnos y CE de 4° grado de educación primaria de CE  
polidocentes completos en el área urbana, 1998**

Departamentos	Estratos	Estatales urbanos		No estatales		Total	
		CE	Alumnos	CE	Alumnos	CE	Alumnos
Amazonas	Sierra	31	1 399	0	0	31	1 399
Amazonas	Selva	38	2 280	3	133	41	2 413
Ancash	Costa Norte	70	6 943	45	1 449	115	8 392
Ancash	Sierra	115	6 662	16	677	131	7 339
Apurímac	Sierra	98	7 069	5	126	103	7 195
Arequipa	Costa Sur	37	2 289	12	311	49	2 600
Arequipa	Sierra	180	11 631	184	5 297	364	16 928
Ayacucho	Sierra	117	8 386	17	434	134	8 820
Ayacucho	Selva	3	317	1	12	4	329
Cajamarca	Costa Norte	6	151	0	0	6	151
Cajamarca	Sierra	109	6 363	13	241	122	6 604
Cajamarca	Selva	30	2 436	6	300	36	2 736
Callao	Lima Ciudad (Callao)	89	10 870	162	3 825	251	14 695
Cusco	Sierra	166	14 830	54	2 035	220	16 865
Cusco	Selva	16	1 178	3	93	19	1 271
Huancavelica	Sierra	57	3 618	4	95	61	3 713
Huánuco	Sierra	85	6 706	15	498	100	7 204
Huánuco	Selva	20	1 723	3	76	23	1 799
Ica	Costa Central sin Lima	122	11 479	55	1 410	177	12 889
Ica	Sierra	0	0	0	0	0	0
Junín	Sierra	153	13 370	66	1 459	219	14 829
Junín	Selva	23	2 890	11	368	34	3 258
La Libertad	Costa Norte	135	13 861	101	4 487	236	18 348
La Libertad	Sierra	62	3 911	4	146	66	4 057
Lambayeque	Costa Norte	122	15 232	62	3 063	184	18 295
Lambayeque	Sierra	1	32	0	0	1	32
Lima	Costa Central sin Lima	92	8 771	69	1 619	161	10 390
Lima	Lima Ciudad	759	95 871	1363	33 564	2122	129 435
Lima	Sierra	13	338	1	9	14	347
Loreto	Selva	97	10 135	17	778	114	10 913
Madre de Dios	Selva	21	1 530	2	35	23	1 565
Moquegua	Costa Sur	21	1 646	9	251	30	1 897
Moquegua	Sierra	6	140	1	54	7	194
Pasco	Sierra	33	2 974	10	277	43	3 251
Pasco	Selva	9	758	0	0	9	758
Piura	Costa Norte	203	18 566	119	3 459	322	22 025
Piura	Sierra	8	634	0	0	8	634
Puno	Sierra	204	15 569	41	1 156	245	16 725
Puno	Selva	2	36	1	51	3	87
San Martín	Selva	145	10 242	14	221	159	10 463
Tacna	Costa Sur	43	3 982	19	691	62	4 673
Tacna	Sierra	4	157	3	103	7	260
Tumbes	Costa Norte	56	3 400	12	209	68	3 609
Ucayali	Selva	56	7 764	7	92	63	7 856
<b>Total</b>		<b>3657</b>	<b>338 139</b>	<b>2530</b>	<b>69 104</b>	<b>6187</b>	<b>407 243</b>

**Muestra de alumnos y CE de 4° grado de educación primaria de CE  
polidocentes completos en el área urbana, 1998**

Departamentos	Estratos	Estatales urbanos		No estatales		Total	
		CE	Alumnos	CE	Alumnos	CE	Alumnos
Amazonas	Sierra	6	180	0	0	6	180
Amazonas	Selva	10	300	2	60	12	360
Ancash	Costa Norte	7	210	2	60	9	270
Ancash	Sierra	7	210	2	60	9	270
Apurímac	Sierra	17	510	0	0	17	510
Arequipa	Costa Sur	12	360	2	60	14	420
Arequipa	Sierra	10	300	5	150	15	450
Ayacucho	Sierra	15	450	2	60	17	510
Ayacucho	Selva	2	60	0	0	2	60
Cajamarca	Costa Norte	0	0	0	0	0	0
Cajamarca	Sierra	11	330	0	0	11	330
Cajamarca	Selva	4	120	2	60	6	180
Callao	Lima Ciudad (Callao)	12	360	5	150	17	510
Cusco	Sierra	14	420	2	60	16	480
Cusco	Selva	2	60	0	0	2	60
Huancavelica	Sierra	17	510	0	0	17	510
Huánuco	Sierra	12	360	2	60	14	420
Huánuco	Selva	3	90	0	0	3	90
Ica	Costa Central sin Lima	25	750	3	90	28	840
Ica	Sierra	0	0	0	0	0	0
Junín	Sierra	12	360	2	60	14	420
Junín	Selva	3	90	0	0	3	90
La Libertad	Costa Norte	11	330	4	120	15	450
La Libertad	Sierra	3	90	0	0	3	90
Lambayeque	Costa Norte	14	420	3	90	17	510
Lambayeque	Sierra	0	0	0	0	0	0
Lima	Costa Central sin Lima	19	570	3	90	22	660
Lima	Lima Ciudad	78	2 340	27	810	105	3 150
Lima	Sierra	0	0	0	0	0	0
Loreto	Selva	15	450	2	60	17	510
Madre de Dios	Selva	17	510	0	0	17	510
Moquegua	Costa Sur	13	390	2	60	15	450
Moquegua	Sierra	2	60	0	0	2	60
Pasco	Sierra	12	360	2	60	14	420
Pasco	Selva	3	90	0	0	3	90
Piura	Costa Norte	15	450	3	90	18	540
Piura	Sierra	2	60	0	0	2	60
Puno	Sierra	15	450	2	60	17	510
Puno	Selva	0	0	0	0	0	0
San Martín	Selva	17	510	0	0	17	510
Tacna	Costa Sur	22	660	4	120	26	780
Tacna	Sierra	2	60	0	0	2	60
Tumbes	Costa Norte	16	480	2	60	18	540
Ucayali	Selva	17	510	0	0	17	510
<b>Total</b>		<b>494</b>	<b>14 820</b>	<b>85</b>	<b>2550</b>	<b>579</b>	<b>17 370</b>

**Población de alumnos y CE de 6° grado de educación primaria de CE  
polidocentes completos en el área urbana, 1998**

Departamentos	Estratos	Estatales urbanos		No estatales		Total	
		CE	Alumnos	CE	Alumnos	CE	Alumnos
Amazonas	Sierra	31	1 226	0	0	31	1 226
Amazonas	Selva	38	2 051	3	122	41	2 173
Ancash	Costa Norte	70	6 011	45	1 260	115	7 271
Ancash	Sierra	115	6 064	16	616	131	6 680
Apurímac	Sierra	98	5 667	5	89	103	5 756
Arequipa	Costa Sur	37	2 044	12	295	49	2 339
Arequipa	Sierra	180	11 424	184	4 859	364	16 283
Ayacucho	Sierra	117	7 206	17	406	134	7 612
Ayacucho	Selva	3	260	1	12	4	272
Cajamarca	Costa Norte	6	136	0	0	6	136
Cajamarca	Sierra	109	5 612	13	178	122	5 790
Cajamarca	Selva	30	2 066	6	351	36	2 417
Callao	Lima Ciudad (Callao)	89	10 205	162	3 326	251	13 531
Cusco	Sierra	166	12 936	54	1 931	220	14 867
Cusco	Selva	16	1 133	3	92	19	1 225
Huancavelica	Sierra	57	3 154	4	85	61	3 239
Huánuco	Sierra	85	6 294	15	483	100	6 777
Huánuco	Selva	20	1 659	3	75	23	1 734
Ica	Costa Central sin Lima	122	10 726	55	1 361	177	12 087
Ica	Sierra	0	0	0	0	0	0
Junín	Sierra	153	12 393	66	1 291	219	13 684
Junín	Selva	23	2 480	11	320	34	2 800
La Libertad	Costa Norte	135	12 216	101	3 918	236	16 134
La Libertad	Sierra	62	3 219	4	148	66	3 367
Lambayeque	Costa Norte	122	13 021	62	2 696	184	15 717
Lambayeque	Sierra	1	19	0	0	1	19
Lima	Costa Central sin Lima	92	8 012	69	1 392	161	9 404
Lima	Lima Ciudad	759	89 592	1363	29 820	2122	119 412
Lima	Sierra	13	296	1	7	14	303
Loreto	Selva	97	7 197	17	640	114	7 837
Madre de Dios	Selva	21	1 381	2	30	23	1 411
Moquegua	Costa Sur	21	1 699	9	297	30	1 996
Moquegua	Sierra	6	130	1	64	7	194
Pasco	Sierra	33	2 648	10	227	43	2 875
Pasco	Selva	9	635	0	0	9	635
Piura	Costa Norte	203	15 757	119	3 143	322	18 900
Piura	Sierra	8	576	0	0	8	576
Puno	Sierra	204	14 214	41	1 007	245	15 221
Puno	Selva	2	27	1	41	3	68
San Martín	Selva	145	8 419	14	205	159	8 624
Tacna	Costa Sur	43	3 379	19	629	62	4 008
Tacna	Sierra	4	149	3	112	7	261
Tumbes	Costa Norte	56	3 071	12	186	68	3 257
Ucayali	Selva	56	6 088	7	95	63	6 183
<b>Total</b>		<b>3657</b>	<b>302 492</b>	<b>2530</b>	<b>61 809</b>	<b>6187</b>	<b>364 301</b>

**Muestra de alumnos y CE de 6° grado de educación primaria de CE  
polidocentes completos en el área urbana, 1998**

Departamentos	Estratos	Estatales urbanos		No estatales		Total	
		CE	Alumnos	CE	Alumnos	CE	Alumnos
Amazonas	Sierra	6	180	0	0	6	180
Amazonas	Selva	10	300	2	60	12	360
Ancash	Costa Norte	7	210	2	60	9	270
Ancash	Sierra	7	210	2	60	9	270
Apurímac	Sierra	17	510	0	0	17	510
Arequipa	Costa Sur	12	360	2	60	14	420
Arequipa	Sierra	10	300	5	150	15	450
Ayacucho	Sierra	15	450	2	60	17	510
Ayacucho	Selva	2	60	0	0	2	60
Cajamarca	Costa Norte	0	0	0	0	0	0
Cajamarca	Sierra	11	330	0	0	11	330
Cajamarca	Selva	4	120	2	60	6	180
Callao	Lima Ciudad (Callao)	12	360	5	150	17	510
Cusco	Sierra	14	420	2	60	16	480
Cusco	Selva	2	60	0	0	2	60
Huancavelica	Sierra	17	510	0	0	17	510
Huánuco	Sierra	12	360	2	60	14	420
Huánuco	Selva	3	90	0	0	3	90
Ica	Costa Central sin Lima	25	750	3	90	28	840
Ica	Sierra	0	0	0	0	0	0
Junín	Sierra	12	360	2	60	14	420
Junín	Selva	3	90	0	0	3	90
La Libertad	Costa Norte	11	330	4	120	15	450
La Libertad	Sierra	3	90	0	0	3	90
Lambayeque	Costa Norte	14	420	3	90	17	510
Lambayeque	Sierra	0	0	0	0	0	0
Lima	Costa Central sin Lima	19	570	3	90	22	660
Lima	Lima Ciudad	78	2 340	27	810	105	3 150
Lima	Sierra	0	0	0	0	0	0
Loreto	Selva	15	450	2	60	17	510
Madre de Dios	Selva	17	510	0	0	17	510
Moquegua	Costa Sur	13	390	2	60	15	450
Moquegua	Sierra	2	60	0	0	2	60
Pasco	Sierra	12	360	2	60	14	420
Pasco	Selva	3	90	0	0	3	90
Piura	Costa Norte	15	450	3	90	18	540
Piura	Sierra	2	60	0	0	2	60
Puno	Sierra	15	450	2	60	17	510
Puno	Selva	0	0	0	0	0	0
San Martín	Selva	17	510	0	0	17	510
Tacna	Costa Sur	22	660	4	120	26	780
Tacna	Sierra	2	60	0	0	2	60
Tumbes	Costa Norte	16	480	2	60	18	540
Ucayali	Selva	17	510	0	0	17	510
<b>Total</b>		<b>494</b>	<b>14 820</b>	<b>85</b>	<b>2550</b>	<b>579</b>	<b>17 370</b>

## MUESTRA FINAL EFECTIVA

Como es usual en investigaciones por muestreo, aun con un cuidadoso trabajo de campo, siempre habrá diferencias entre las muestras diseñadas o esperadas y las muestras realmente logradas o efectivas. En el caso de primaria, una vez aplicadas las pruebas, las muestras efectivas resultaron distintas por grado. En total se lograron cubrir 578 CE (ya se mencionó que en un estrato de muestreo de Madre de Dios no hubo suficiente cantidad de CE y se debieron tomar más secciones en un CE).

En el caso de cuarto grado, la muestra efectiva fue de 16 997 alumnos, lo que significa una cobertura de 98% de la muestra diseñada, que es un resultado excelente para este tipo de trabajos. Esta muestra equivale a 4,2% de la población objetivo y permite estimaciones dentro de los márgenes de error pre-determinados en el diseño.

En sexto grado la muestra efectiva fue de 16 833 alumnos, que equivale a una cobertura de 97%, un resultado también excelente. Cabe notar que como la población de alumnos de sexto es menor (por la inevitable deserción), la muestra efectiva equivale a 4,6% de la población objetivo y, como en el caso

de cuarto grado, las estimaciones se mantienen dentro de lo diseñado al planificar la muestra.

La cobertura de las muestras efectivas obtenidas fue:

- 9,3% de los CE en general.
- 13,5% de los CE estatales.
- 3,4% de los CE no estatales.
- 4,2% de los alumnos de cuarto grado en general.
- 4,3% de los alumnos de cuarto grado de CE estatales.
- 3,6% de los alumnos de cuarto grado de CE no estatales.
- 4,6% de los alumnos de sexto grado en general.
- 4,8% de los alumnos de sexto grado de CE estatales.
- 3,9% de los alumnos de sexto grado de CE no estatales.

Cabe notar que las ponderaciones para las estimaciones se calculan sobre las muestras efectivas.

De aquí en adelante se presentan las poblaciones objetivo y las muestras efectivas por departamentos y estratos principales.

Distribución de población de 4° grado de primaria por estratos						
Estratos	Estatales		No estatales		Total poblacional	
	CE	Alumnos	CE	Alumnos	CE	Alumnos
Costa Norte	592	58 153	339	12 667	931	70 820
Costa Central sin Lima	214	20 250	124	3 029	338	23 279
Costa Sur	101	7 917	40	1 253	141	9 170
Sierra Norte	326	19 001	33	1 064	359	20 065
Sierra Central	458	35 392	113	2 772	571	38 164
Sierra Sur	658	49 396	288	8 771	946	58 167
Selva Norte	310	25 093	40	1 432	350	26 525
Selva Central	111	13 452	22	548	133	14 000
Selva Sur	39	2 744	6	179	45	2 923
Lima Ciudad	848	106 741	1525	37 389	2373	144 130
<b>Total</b>	<b>3657</b>	<b>338 139</b>	<b>2530</b>	<b>69 104</b>	<b>6187</b>	<b>407 243</b>

Distribución de la muestra efectiva de 4° grado de primaria por estratos						
Estratos	Estatales		No estatales		Total poblacional	
	CE	Alumnos	CE	Alumnos	CE	Alumnos
Costa Norte	64	1 870	14	384	78	2 254
Costa Central sin Lima	44	1 309	6	170	50	500
Costa Sur	47	1 399	8	230	55	550
Sierra Norte	29	849	2	60	31	909
Sierra Central	68	2 015	8	218	76	2 233
Sierra Sur	59	1 727	9	259	68	1 986
Selva Norte	44	1 282	7	209	51	1 491
Selva Central	28	837	0	0	28	837
Selva Sur	20	576	0	0	20	576
Lima Ciudad	89	2 649	32	954	121	3 603
<b>Total</b>	<b>492</b>	<b>14 513</b>	<b>86</b>	<b>2484</b>	<b>578</b>	<b>16 997</b>

Distribución de población de 6° grado de primaria por estratos						
Estratos	Estatales		No estatales		Total poblacional	
	CE	Alumnos	CE	Alumnos	CE	Alumnos
Costa Norte	592	50 212	339	11 203	931	61 415
Costa Central sin Lima	214	18 738	124	2 753	338	21 491
Costa Sur	101	7 122	40	1 221	141	8 343
Sierra Norte	326	16 716	33	942	359	17 658
Sierra Central	458	31 991	113	2 499	571	34 490
Sierra Sur	658	44 520	288	8 062	946	52 582
Selva Norte	310	19 733	40	1 318	350	21 051
Selva Central	111	11 122	22	502	133	11 624
Selva Sur	39	2 541	6	163	45	2 704
Lima Ciudad	848	99 797	1525	33 146	2 373	132 943
<b>Total</b>	<b>3657</b>	<b>302 492</b>	<b>2530</b>	<b>61 809</b>	<b>6187</b>	<b>364 301</b>

Distribución de la muestra efectiva de 6° grado de primaria por estratos						
Estratos	Estatales		No estatales		Total poblacional	
	CE	Alumnos	CE	Alumnos	CE	Alumnos
Costa Norte	64	1 860	14	375	78	2 235
Costa Central sin Lima	44	1 306	6	163	50	500
Costa Sur	47	1 389	8	221	55	550
Sierra Norte	29	852	2	47	31	899
Sierra Central	68	1 957	8	198	76	2 155
Sierra Sur	60	1 712	9	258	69	1 970
Selva Norte	44	1 287	7	203	51	1 490
Selva Central	28	838	0	0	28	838
Selva Sur	19	559	0	0	19	559
Lima Ciudad	89	2 658	32	950	121	3 608
<b>Total</b>	<b>492</b>	<b>14 418</b>	<b>86</b>	<b>2415</b>	<b>578</b>	<b>16 833</b>

<b>Poblaciones y muestras efectivas en primaria (1998) por departamento</b>				
<b>Departamentos</b>	<b>4° P. Población</b>	<b>4° P. Muestra</b>	<b>6° P. Población</b>	<b>6° P. Muestra</b>
Amazonas	3 812	518	3 399	512
Ancash	15 731	528	13 951	508
Apurímac	7 195	496	5 756	492
Arequipa	19 528	845	18 622	826
Ayacucho	9 149	554	7 884	543
Cajamarca	9 491	473	8 343	477
Callao	14 695	510	13 531	506
Cusco	18 136	536	16 092	526
Huancavelica	3 713	504	3 239	486
Huánuco	9 003	498	8 511	482
Ica	12 889	828	12 087	823
Junín	18 087	497	16 484	482
La Libertad	22 405	568	19 501	569
Lambayeque	18 327	481	15 736	479
Lima	140 172	3 744	129 119	3 748
Loreto	10 913	496	7 837	496
Madre de Dios	1 565	490	1 411	500
Moquegua	2 091	485	2 190	476
Pasco	4 009	510	3 510	491
Piura	22 659	582	19 476	590
Puno	16 812	501	15 289	486
San Martín	10 463	502	8 624	503
Tacna	4 933	838	4 269	833
Tumbes	3 609	506	3 257	490
Ucayali	7 856	507	6 183	509
<b>Total</b>	<b>407 243</b>	<b>16 997</b>	<b>364 301</b>	<b>16 833</b>

**Muestra efectiva de alumnos y CE de 4° grado de educación primaria de CE polidocentes completos en el área urbana, 1998**

Departamentos	Estratos	Estatales urbanos		No estatales		Total	
		CE	Alumnos	CE	Alumnos	CE	Alumnos
Amazonas	Sierra	7	204	0	0	7	204
Amazonas	Selva	9	254	2	60	11	314
Ancash	Costa Norte	7	210	2	57	9	267
Ancash	Sierra	7	201	2	60	9	261
Apurímac	Sierra	17	496	0	0	17	496
Arequipa	Costa Sur	12	351	2	58	14	409
Arequipa	Sierra	10	297	5	139	15	436
Ayacucho	Sierra	15	434	2	60	17	494
Ayacucho	Selva	2	60	0	0	2	60
Cajamarca	Costa Norte	0	0	0	0	0	0
Cajamarca	Sierra	10	294	0	0	10	294
Cajamarca	Selva	3	90	3	89	6	179
Callao	Lima Ciudad (Callao)	12	360	5	150	17	510
Cusco	Sierra	13	390	2	60	15	450
Cusco	Selva	3	86	0	0	3	86
Huancavelica	Sierra	17	504	0	0	17	504
Huánuco	Sierra	12	358	2	50	14	408
Huánuco	Selva	3	90	0	0	3	90
Ica	Costa Central sin Lima	25	748	3	80	28	828
Ica	Sierra	0	0	0	0	0	0
Junín	Sierra	12	359	2	48	14	407
Junín	Selva	3	90	0	0	3	90
La Libertad	Costa Norte	12	359	4	119	16	478
La Libertad	Sierra	3	90	0	0	3	90
Lambayeque	Costa Norte	14	396	3	85	17	481
Lambayeque	Sierra	0	0	0	0	0	0
Lima	Costa Central sin Lima	19	561	3	90	22	651
Lima	Lima Ciudad	77	2 289	27	804	104	3 093
Lima	Sierra	0	0	0	0	0	0
Loreto	Selva	15	436	2	60	17	496
Madre de Dios	Selva	17	490	0	0	17	490
Moquegua	Costa Sur	13	389	2	53	15	442
Moquegua	Sierra	2	43	0	0	2	43
Pasco	Sierra	12	360	2	60	14	420
Pasco	Selva	3	90	0	0	3	90
Piura	Costa Norte	15	444	3	78	18	522
Piura	Sierra	2	60	0	0	2	60
Puno	Sierra	15	441	2	60	17	501
Puno	Selva	0	0	0	0	0	0
San Martín	Selva	17	502	0	0	17	502
Tacna	Costa Sur	22	659	4	119	26	778
Tacna	Sierra	2	60	0	0	2	60
Tumbes	Costa Norte	16	461	2	45	18	506
Ucayali	Selva	17	507	0	0	17	507
<b>Total</b>		<b>492</b>	<b>14 513</b>	<b>86</b>	<b>2 484</b>	<b>578</b>	<b>16 997</b>

**Muestra efectiva de alumnos y CE de 6° grado de educación primaria de CE  
polidocentes completos en el área urbana, 1998**

Departamentos	Estratos	Estatales urbanos		No estatales		Total	
		CE	Alumnos	CE	Alumnos	CE	Alumnos
Amazonas	Sierra	7	201	0	0	7	201
Amazonas	Selva	9	258	2	53	11	311
Ancash	Costa Norte	7	206	2	51	9	257
Ancash	Sierra	7	204	2	47	9	251
Apurímac	Sierra	17	492	0	0	17	492
Arequipa	Costa Sur	12	348	2	51	14	399
Arequipa	Sierra	10	289	5	138	15	427
Ayacucho	Sierra	15	423	2	60	17	483
Ayacucho	Selva	2	60	0	0	2	60
Cajamarca	Costa Norte	0	0	0	0	0	0
Cajamarca	Sierra	10	297	0	0	10	297
Cajamarca	Selva	3	90	3	90	6	180
Callao	Lima Ciudad (Callao)	12	360	5	146	17	506
Cusco	Sierra	14	407	2	60	16	467
Cusco	Selva	2	59	0	0	2	59
Huancavelica	Sierra	17	486	0	0	17	486
Huánuco	Sierra	12	350	2	43	14	393
Huánuco	Selva	3	89	0	0	3	89
Ica	Costa Central sin Lima	25	750	3	73	28	823
Ica	Sierra	0	0	0	0	0	0
Junín	Sierra	12	347	2	45	14	392
Junín	Selva	3	90	0	0	3	90
La Libertad	Costa Norte	12	359	4	120	16	479
La Libertad	Sierra	3	90	0	0	3	90
Lambayeque	Costa Norte	14	389	3	90	17	479
Lambayeque	Sierra	0	0	0	0	0	0
Lima	Costa Central sin Lima	19	556	3	90	22	646
Lima	Lima Ciudad	77	2 298	27	804	104	3 102
Lima	Sierra	0	0	0	0	0	0
Loreto	Selva	15	436	2	60	17	496
Madre de Dios	Selva	17	500	0	0	17	500
Moquegua	Costa Sur	13	385	2	53	15	438
Moquegua	Sierra	2	38	0	0	2	38
Pasco	Sierra	12	351	2	50	14	401
Pasco	Selva	3	90	0	0	3	90
Piura	Costa Norte	15	443	3	87	18	530
Piura	Sierra	2	60	0	0	2	60
Puno	Sierra	15	426	2	60	17	486
Puno	Selva	0	0	0	0	0	0
San Martín	Selva	17	503	0	0	17	503
Tacna	Costa Sur	22	656	4	117	26	773
Tacna	Sierra	2	60	0	0	2	60
Tumbes	Costa Norte	16	463	2	27	18	490
Ucayali	Selva	17	509	0	0	17	509
<b>Total</b>		<b>492</b>	<b>14 418</b>	<b>86</b>	<b>2415</b>	<b>578</b>	<b>16 833</b>

**Secundaria: Resumen de resultados obtenidos.  
Muestra efectiva de alumnos y CE de 4° grado de educación  
secundaria de CE polidocentes completos en el área urbana, 1998**

Departamentos	Estratos	Estatales urbanos		No estatales		Total	
		CE	Alumnos	CE	Alumnos	CE	Alumnos
Amazonas	Sierra	30	904	0	0	30	904
Amazonas	Selva	26	987	3	107	29	1 094
Ancash	Costa Norte	36	3 336	37	939	73	4 275
Ancash	Sierra	111	3 989	13	570	124	4 559
Apurímac	Sierra	66	3 208	3	103	69	3 311
Arequipa	Costa Sur	26	1 640	11	238	37	1 878
Arequipa	Sierra	112	8 423	140	4 686	252	13 109
Ayacucho	Sierra	55	2 916	20	482	75	3 398
Ayacucho	Selva	2	74	2	22	4	96
Cajamarca	Costa Norte	8	180	0	0	8	180
Cajamarca	Sierra	87	4 526	14	178	101	4 704
Cajamarca	Selva	31	1 713	8	181	39	1 894
Callao	Lima Ciudad (Callao)	34	5 536	135	2 917	169	8 453
Cusco	Sierra	94	7 427	53	1 706	147	9 133
Cusco	Selva	11	687	5	116	16	803
Huancavelica	Sierra	73	2 452	4	179	77	2 631
Huánuco	Sierra	60	3 852	18	536	78	4 388
Huánuco	Selva	10	1 085	5	81	15	1 166
Ica	Costa Central sin Lima	58	7 240	37	959	95	8 199
Ica	Sierra	2	15	0	0	2	15
Junín	Sierra	62	5 940	69	1 806	131	7 746
Junín	Selva	10	850	10	286	20	1 136
La Libertad	Costa Norte	70	7 446	62	1 936	132	9 382
La Libertad	Sierra	58	1 891	4	18	62	1 909
Lambayeque	Costa Norte	60	7 447	30	1 042	90	8 489
Lambayeque	Sierra	1	14	0	0	1	14
Lima	Costa Central sin Lima	37	5 206	48	1 190	85	6 396
Lima	Lima Ciudad	429	53 196	984	23 031	1413	76 227
Lima	Sierra	20	400	0	0	20	400
Loreto	Selva	25	3 627	14	506	39	4 133
Madre de Dios	Selva	8	403	2	25	10	428
Moquegua	Costa Sur	13	1 344	7	257	20	1 601
Moquegua	Sierra	6	120	1	46	7	166
Pasco	Sierra	19	1 438	10	179	29	1 617
Pasco	Selva	5	312	0	0	5	312
Piura	Costa Norte	96	7 652	94	2 370	190	10 022
Piura	Sierra	11	461	0	0	11	461
Puno	Sierra	84	7 353	36	1 119	120	8 472
Puno	Selva	0	0	0	0	0	0
San Martín	Selva	71	4 517	8	11	79	4 528
Tacna	Costa Sur	23	1 962	18	677	41	2 639
Tacna	Sierra	3	50	3	117	6	167
Tumbes	Costa Norte	18	1 552	15	272	33	1 824
Ucayali	Selva	29	3 428	9	95	38	3 523
<b>Total</b>		<b>2090</b>	<b>176 799</b>	<b>1 932</b>	<b>49 083</b>	<b>4022</b>	<b>225 882</b>

**Muestra efectiva de alumnos y CE de 5° grado de educación secundaria de CE  
polidocentes completos en el área urbana, 1998**

Departamentos	Estratos	Estatales urbanos		No estatales		Total	
		CE	Alumnos	CE	Alumnos	CE	Alumnos
Amazonas	Sierra	30	860	0	0	30	860
Amazonas	Selva	26	884	3	101	29	985
Ancash	Costa Norte	36	3 044	37	865	73	3 909
Ancash	Sierra	111	3 430	13	567	124	3 997
Apurímac	Sierra	66	2 739	3	111	69	2 850
Arequipa	Costa Sur	26	1 448	11	233	37	1 681
Arequipa	Sierra	112	7 739	140	4 587	252	12 326
Ayacucho	Sierra	55	2 620	20	414	75	3 034
Ayacucho	Selva	2	85	2	11	4	96
Cajamarca	Costa Norte	8	180	0	0	8	180
Cajamarca	Sierra	87	4 254	14	211	101	4 465
Cajamarca	Selva	31	1 465	8	153	39	1 618
Callao	Lima Ciudad (Callao)	34	5 006	135	2 660	169	7 666
Cusco	Sierra	94	6 180	53	1 514	147	7 694
Cusco	Selva	11	657	5	86	16	743
Huancavelica	Sierra	73	2 099	4	123	77	2 222
Huánuco	Sierra	60	3 360	18	509	78	3 869
Huánuco	Selva	10	1 001	5	78	15	1 079
Ica	Costa Central sin Lima	58	6 058	37	905	95	6 963
Ica	Sierra	2	16	0	0	2	16
Junín	Sierra	62	5 217	69	1 693	131	6 910
Junín	Selva	10	733	10	231	20	964
La Libertad	Costa Norte	70	6 923	62	1 745	132	8 668
La Libertad	Sierra	58	1 463	4	19	62	1 482
Lambayeque	Costa Norte	60	6 481	30	1 031	90	7 512
Lambayeque	Sierra	1	12	0	0	1	12
Lima	Costa Central sin Lima	37	4 806	48	1 152	85	5 958
Lima	Lima Ciudad	429	48 340	984	22 046	1413	70 386
Lima	Sierra	20	324	0	0	20	324
Loreto	Selva	25	3 335	14	505	39	3 840
Madre de Dios	Selva	8	361	2	14	10	375
Moquegua	Costa Sur	13	1 150	7	249	20	1 399
Moquegua	Sierra	6	120	1	58	7	178
Pasco	Sierra	19	1 350	10	167	29	1 517
Pasco	Selva	5	251	0	0	5	251
Piura	Costa Norte	96	7 173	94	2 283	190	9 456
Piura	Sierra	11	417	0	0	11	417
Puno	Sierra	84	6 235	36	1 065	120	7 300
Puno	Selva	0	0	0	0	0	0
San Martín	Selva	71	3 777	8	97	79	3 874
Tacna	Costa Sur	23	1 852	18	706	41	2 558
Tacna	Sierra	3	39	3	85	6	124
Tumbes	Costa Norte	18	1 382	15	290	33	1 672
Ucayali	Selva	29	2 847	9	88	38	2 935
<b>Total</b>		<b>2090</b>	<b>157 713</b>	<b>1932</b>	<b>46 652</b>	<b>4022</b>	<b>204 365</b>

La representatividad de la muestra según criterios de estratificación fue la siguiente:

<b>Población total (4° de secundaria)</b>						
	<b>Estatal</b>		<b>No estatal</b>		<b>Total</b>	
	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>
Rural	2620	27 360	139	1 504	2759	28 864
Urbana	2698	68 181	2028	32 398	4726	100 579
Nacional	5318	95 541	2167	33 902	7485	129 443
<b>Participación con respecto al total (%) (4° de secundaria)</b>						
<b>Total</b>						
		<b>Escuelas</b>		<b>Alumnos</b>		
Urbana		63,14		77,70		
Polidocentes completos		100,00		100,00		
Polidocentes completos urbanos		100,00		100,00		
<b>Población total (5° de secundaria)</b>						
	<b>Estatal</b>		<b>No estatal</b>		<b>Total</b>	
	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>	<b>Escuelas</b>	<b>Alumnos</b>
Rural	2620	24 015	139	1 090	2759	25 105
Urbana	2698	63 689	2028	30 426	4726	94 115
Nacional	5318	87 704	2167	31 516	7485	119 220
<b>Participación con respecto al total (%) (5° de secundaria)</b>						
<b>Total</b>						
		<b>Escuelas</b>		<b>Alumnos</b>		
Urbana		63,14		78,94		
Polidocentes completos		100,00		100,00		
Polidocentes completos urbanos		100,00		100,00		

# ANEXO

## FÓRMULA DE PONDERACIÓN

Las ponderaciones se emplean en situaciones en las que se ha elegido muestras no proporcionales y se requiere “devolver” la proporcionalidad que la distribución de sus estratos tiene en el universo.

La fórmula empleada en CRECER 1998 para tal fin debió contemplar las distintas etapas del diseño, y quedó de la siguiente manera:

$$W = (n/N) (M/(j*30)) (30/m)$$

cuyos términos se interpretan como se explica a continuación:

- $n/N$  = Total de alumnos en la muestra/Total de alumnos en el universo.  
Ésta es la fracción de muestreo, es decir, cuánto representa la muestra del universo.
- $(M/(j*30))$  = Total de alumnos en el estrato en el universo/(CE en el estrato en la muestra \* Alumnos esperados en el aula); siendo  $j*30$  = Total de alumnos en la muestra.  
Ésta es la cantidad de veces que el universo contiene a la muestra en el estrato.
- $(30/m)$  = Alumnos esperados en el aula/Alumnos hallados en el aula.

Veamos un ejemplo aplicando la fórmula para un estrato del departamento de Amazonas:

	Centros educativos		Alumnos	
	Población	Muestra	Población	Muestra
Universo nacional	6092	576	404 964	16 827
Estrato 1 de Amazonas	31	7	1 399	206
1 CE del estrato 1 de Amazonas			30*	26

\* Valor esperado por aula.

$$W = (n/N) (M/(j*30)) (30/m) = (16\ 827/404\ 964) (1\ 399/(7*30)) (30/26)$$

Ésta es la fórmula aplicada en las bases de datos por CE para cada alumno.

# ESTIMACIÓN DEL ERROR ESTÁNDAR EN LAS PRUEBAS CRECER 1998

CHOLLY FARRO  
JORGE BAZÁN  
ALBERTO TORREBLANCA  
CESAR MILLONES

Esta sección del documento tiene por finalidad dar cuenta de métodos alternativos para la estimación del error estándar<sup>1</sup> de los promedios de la variable rendimiento de las pruebas CRECER 1998.

La estimación de los errores estándar calculados por los programas estadísticos convencionales supone que la información recogida proviene de estudios con diseños no complejos (generalmente muestreo aleatorio simple). Como se verá luego, el diseño muestral de las pruebas CRECER 1998 es complejo<sup>2</sup>. Por ello, se requiere un tratamiento especial para el cálculo de los errores estándar de los promedios de rendimiento en las pruebas.

Esta sección presenta la metodología y los cálculos de los errores estándar desde dos enfoques: i) el que se refiere a las fórmulas derivadas del diseño muestral de las pruebas CRECER 1998; y, ii) el generado a partir de la metodología basada en el remuestreo sobre la misma muestra.

## **1 METODOLOGÍA BASADA EN EL DISEÑO DE LA MUESTRA**

### **EL DISEÑO MUESTRAL DE LAS PRUEBAS CRECER 1998**

En diseños simples las unidades de muestreo o unidades a seleccionar son las mismas que se van a analizar y, por lo tanto, constituyen el único objeto de estudio. En diseños complejos

—los más empleados en la práctica—, las unidades de muestreo no coinciden necesariamente con las unidades de análisis. Éste es el caso del muestreo en dos etapas o bietápico.

En CRECER 1998 se empleó conjuntamente muestreo estratificado y de conglomerados<sup>3</sup>. Asimismo, se combinaron los criterios de gestión (escuela estatal o no estatal), región y departamento para formar estratos<sup>4</sup>. Dentro de éstos se seleccionaron conglomerados (escuelas) en forma proporcional a sus tamaños<sup>5</sup>.

Desde el punto de vista de la selección, el muestreo empleado fue bietápico. Las uni-

1 En esta sección la denominación "error estándar" se refiere a la desviación del estimador de interés (la media de los rendimientos) respecto del promedio de ese estimador. No se emplea el término "desviación estándar", pues éste es aplicado más bien a las desviaciones de los datos originales respecto de su media. Por otro lado, la desviación estándar está asociada a un diseño de muestreo aleatorio simple, que no es el caso de las pruebas CRECER 1998.

2 Véase también la sección anterior, Diseño muestral en la aplicación nacional CRECER 1998.

3 Las unidades de análisis son conjuntos de elementos y no cada uno de ellos. En nuestro caso, una escuela es un conglomerado, pues está formado por un grupo de aulas y éstas de alumnos.

4 62 estratos contienen información y se forman del cruce de la gestión y región para cada departamento.

5 Véase Calderón. A.: "Informe sobre la muestra para evaluación en cuarto y sexto grados de primaria del Proyecto CRECER 1998". Lima, 1999.

dades primarias de muestreo (UPM) fueron las escuelas, y las unidades de muestreo de la segunda etapa (USM) los alumnos de dichas escuelas.

En la primera etapa, y en cada estrato  $h$ , se seleccionaron  $n_h$  UPM o escuelas. Por ejemplo, en el departamento de Amazonas, en cada uno de sus tres estratos, se eligieron siete, nueve y dos escuelas respectivamente.

En la segunda etapa, en cada estrato  $h$  y en cada escuela  $j$  se seleccionaron  $m_{jh}$  alumnos. Así, en el primer estrato de Amazonas, con siete escuelas, se eligieron aproximadamente 30 alumnos en cada una, es decir, un total aproximado de 210 alumnos<sup>6</sup> en el primer estrato de este departamento. El muestreo en esta etapa es también identificado como muestreo por conglomerados, en el que cada conglomerado está constituido por la escuela, la que, a su vez, contiene las unidades últimas de análisis (los alumnos).

En síntesis, el muestreo estratificado usado en CRECER 1998 es complejo porque combina características de las unidades (por ejemplo, estatal o no estatal) con criterios geográficos o de ubicación (costa, sierra, selva y departamentos). La combinación resultante de ambos da un total de 62 estratos relevantes, los que aparecen en el cuadro 1.

## DESCRIPCIÓN DE LA FÓRMULA DE ESTIMACIÓN DEL ERROR ESTÁNDAR

Para obtener la fórmula del error estándar de la media total en un muestreo bietápico con estratos en la primera etapa y conglomerados en la segunda etapa es necesario calcular la varianza total<sup>7</sup> (ecuación 4), que agre-

6 30 es el valor esperado de alumnos por escuela. En los casos en que dicha cantidad no se obtuvo, se empleó una ponderación (variable peso\_fin), que además recoge las cuotas fijadas para los criterios de estratificación

7 La varianza total puede ser la nacional, departamental, regional o de gestión.

8 Azorín, F. y J.L. Sánchez-Crespo: *Métodos y aplicaciones del muestreo*. Madrid: Alianza Editorial, 1986.

9 El subíndice  $bh$  hace alusión a *between* (entre) y al estrato  $h$ .

ga las varianzas de los estratos de la muestra (ecuación 3). A su vez, la varianza de los estratos agrega las varianzas de las escuelas. En este último caso la varianza del estrato recoge una varianza de primera etapa en la que está presente la varianza *entre* escuelas (ecuación 1) y una varianza de segunda etapa que incluye la varianza *dentro* de escuelas (ecuación 2)<sup>8</sup>.

La nomenclatura empleada es la siguiente:

- $X_{ijh}$ : Puntaje del alumno  $i$  de la escuela  $j$  del estrato  $h$ .
- $\bar{X}_{\bullet jh}$ : Media de la escuela  $j$  en el estrato  $h$  sobre la base del puntaje de los alumnos.
- $\bar{X}_{\bullet\bullet h}$ : Media del estrato  $h$  con base en las medias de las escuelas.
- $\bar{X}_{\dots}$ : Media nacional a partir de las medias de los estratos.
- $N_h$ : Escuelas en la población en el estrato  $h$ .
- $n_h$ : Escuelas en la muestra en el estrato  $h$ .
- $M$ : Alumnos en la población.
- $N$ : Escuelas en la población.
- $\bar{M}$ : Promedio de alumnos en la población por escuela.
- $M_h$ : Alumnos en la población en el estrato  $h$ .
- $\bar{M}_{jh}$ : Promedio de alumnos en la población por escuela  $j$  en el estrato  $h$ .
- $m_h$ : Alumnos en la muestra en el estrato  $h$ .
- $\bar{m}_{jh}$ : Promedio de alumnos en la muestra por escuela  $j$  en el estrato  $h$ .
- $f_{1h}$ : Fracción de muestreo de escuelas.
- $f_{2h}$ : Fracción de muestreo de alumnos.
- $L$ : Estratos.

## Varianza del estrato

Para obtener la varianza de un estrato se calculan las varianzas debidas a las etapas dentro de ese estrato.

### Primera etapa: Cálculo de la varianza entre las escuelas<sup>9</sup>

$$S_{bh}^2 = \bar{m}_h \frac{\sum_j^{n_h} (\bar{X}_{\bullet jh} - \bar{X}_{\bullet\bullet h})^2}{n_h - 1} \quad (1)$$

**Cuadro 1**  
**Estratificación de la población y muestra de Matemática en 4° grado de primaria**

Código	Estratos de sorteo			Población		Muestra	
	Departamentos	Regiones	Gestión	Escuelas (N)	Alumnos (M)	Escuelas (n)	Alumnos (m)
0117	Amazonas	Sierra	Público	31	1 399	7	206
0121	Amazonas	Selva	Público	38	2 280	9	249
0123	Amazonas	Selva	Privado	3	133	2	60
0201	Ancash	Costa Norte	Público	70	6 943	7	205
0203	Ancash	Costa Norte	Privado	45	1 449	2	58
0217	Ancash	Sierra	Público	115	6 662	7	201
0219	Ancash	Sierra	Privado				
0317	Apurímac	Sierra	Público	98	7 069	17	490
0413	Arequipa	Costa Sur	Público	37	289	12	351
0415	Arequipa	Costa Sur	Privado	12	311	2	58
0417	Arequipa	Sierra	Público	180	11 631	10	299
0419	Arequipa	Sierra	Privado	84	5 297	5	137
0517	Ayacucho	Sierra	Público	117	8 386	15	426
0519	Ayacucho	Sierra	Privado	17	434	2	60
0521	Ayacucho	Selva	Público	3	317	2	60
0617	Cajamarca	Sierra	Público	109	6 363	10	295
0621	Cajamarca	Selva	Público	30	2 436	3	88
0623	Cajamarca	Selva	Privado	6	300	3	89
0709	Callao	Lima Ciudad (Callao)	Público	89	10 870	12	306
0711	Callao	Lima Ciudad (Callao)	Privado	162	3 825	5	148
0817	Cusco	Sierra	Público	166	14 830	14	411
0819	Cusco	Sierra	Privado	54	2 035	2	59
0821	Cusco	Selva	Público	16	1 178	2	59
0917	Huancavelica	Sierra	Público	57	3 618	17	499
1017	Huánuco	Selva	Público	20	1 723	3	90
1019	Huánuco	Sierra	Privado	15	498	2	50
1021	Huánuco	Selva	Público	20	1 723	3	90
1105	Ica	Costa Central sin Lima	Público	122	1 149	25	736
1107	Ica	Costa Central sin Lima	Privado	55	1 410	3	81
1217	Junín	Sierra	Público	153	13 370	12	359
1219	Junín	Sierra	Privado	66	1 459	2	47
1221	Junín	Selva	Público	23	2 890	3	90
1301	La Libertad	Costa Norte	Público	135	13 861	12	353
1303	La Libertad	Costa Norte	Privado	101	4 487	4	120
1317	La Libertad	Sierra	Público	62	3 911	3	89
1401	Lambayeque	Costa Norte	Público	122	15 232	14	393
1403	Lambayeque	Costa Norte	Privado	62	3 063	3	85
1505	Lima	Costa Central sin Lima	Público	92	8 771	19	560
1507	Lima	Costa Central sin Lima	Privado	69	1 619	3	89
1509	Lima	Lima Ciudad	Público	759	95 871	77	2 264
1511	Lima	Lima Ciudad	Privado	1363	33 564	27	792
1621	Loreto	Selva	Público	97	10 135	15	430

Código	Estratos de sorteo			Población		Muestra	
	Departamentos	Regiones	Gestión	Escuelas (N)	Alumnos (M)	Escuelas (n)	Alumnos (m)
1623	Loreto	Selva	Privado	17	778	2	59
1721	Madre de Dios	Selva	Público	21	1 530	15	485
1813	Moquegua	Costa Sur	Público	21	1 530	15	485
1815	Moquegua	Costa sur	Privado	9	251	2	54
1817	Moquegua	Sierra	Público	6	140	2	43
1917	Pasco	Sierra	Público	33	2 974	12	358
1919	Pasco	Sierra	Privado	10	277	2	60
1921	Pasco	Selva	Público	9	758	3	90
2001	Piura	Costa Norte	Público	203	18 566	15	448
2003	Piura	Costa Norte	Privado	119	3 459	3	79
2017	Piura	Sierra	Público	8	634	2	60
2117	Puno	Sierra	Público	204	15 569	15	441
2119	Puno	Sierra	Privado	41	1 156	2	59
2221	San Martín	Selva	Público	145	10 242	17	499
2313	Tacna	Costa Sur	Público	43	3 982	22	656
2315	Tacna	Costa Sur	Privado	19	691	4	120
2317	Tacna	Sierra	Público	4	157	2	60
2401	Tumbes	Costa Norte	Público	56	3 400	16	458
2403	Tumbes	Costa Norte	Privado	12	209	2	44
2521	Ucayali	Selva	Público	56	7 764	17	508
<b>Total</b>				6092	404 964	576	16 827

Segunda etapa: Cálculo de la varianza dentro de las escuelas<sup>10</sup>

$$S^2_{wh} = \frac{\sum_j n_{jh} \sum_i (\bar{X}_{ijh} - \bar{X}_{\bullet jh})^2}{n_{jh}(\bar{m}_{jh} - 1)} \quad (2)$$

Luego, a partir de estos dos resultados parciales se calcula la varianza del estrato (que incluye entonces las varianzas “dentro” y “entre”). La raíz cuadrada de esta varianza es el error estándar de la media del estrato.

$$S^2_{\bullet\bullet h} = V(\bar{X}_{\bullet\bullet h}) = (1 - f_{1h}) \frac{S^2_{bh}}{n_h \bar{M}_{jh}} + (1 - f_{2h}) \frac{S^2_{wh}}{n_h \bar{m}_{jh}} \quad (3)$$

donde las fracciones de muestreo son:

$$f_{1h} = \frac{n_h}{N_h} \quad f_{2h} = \frac{\bar{m}_{jh}}{\bar{M}_{jh}}$$

10 El subíndice wh hace alusión a *within* (dentro) y al estrato h.

y los promedios de alumnos por escuelas en la población y en la muestra para un estrato son:

$$\bar{M}_{jh} = \frac{\bar{M}_h}{N_h} \quad \bar{m}_{jh} = \frac{m_h}{n_h}$$

Por lo tanto, la fórmula se compone fundamentalmente de dos términos: la varianza entre escuelas y la varianza dentro de las escuelas. Estas varianzas se suman luego de haber multiplicado cada una por el complemento de la fracción de muestreo correspondiente y dividido por el número de unidades de muestreo respectivo, como se explica a continuación.

### Varianza total

Finalmente, para calcular la varianza total, y sobre la base de la ecuación (3), se utiliza la siguiente expresión:

$$S^2_{\dots} = V(\bar{X}_{\dots}) = \sum_i W^2_h V(\bar{X}_{\bullet\bullet h}) \quad (4)$$

$$\text{donde } W_h = \frac{N_h \bar{M}_{jh}}{N \bar{M}} \text{ y } \bar{M} = \frac{M}{N} \text{ dan las}$$

ponderaciones de los estratos.

### Error estándar

$$EE(\bar{X}...) = S...$$

### APLICACIÓN DE LA FÓRMULA

A manera de ejemplo, en el anexo 1 se presenta el uso de la fórmula para el cálculo del error estándar de dos departamentos. El proceso consiste en identificar los estratos que los componen y, dentro de ellos, las estadísticas de las escuelas (media y varianza). A partir de esto se llega a una estimación del error estándar del departamento. Para la ilustración, se han elegido los departamentos de Arequipa y Amazonas. Los valores finales son los que comparamos —como se verá luego— con los obtenidos a través de los métodos de remuestreo.

## 2. METODOLOGÍA BASADA EN REMUESTREO

Si observamos la fórmula de dos etapas descrita en el primer acápite, notaremos la dificultad para su uso y aplicación en un programa computacional. En general, la complejidad de los diseños muestrales, así como las diferentes probabilidades de muestreo, hacen que la derivación de la fórmula del cálculo del error y su aplicación sea un proceso difícil y en algunos diseños casi imposible<sup>11</sup>.

Avances recientes en el desarrollo de métodos de estimación basados en remuestreo y de los programas que los implementan ayudan a reducir estas dificultades introduciendo métodos de simulación sobre la muestra que, sin apartarse de los lineamientos teóricos de las definiciones de los errores estándar, proveen métodos para calcularlos de manera eficiente y con precisión suficiente.

### MÉTODOS ALTERNATIVOS DE REMUESTREO

Dos de estos métodos de remuestreo son el Bootstrap y el Jackknife<sup>12</sup>. Ambos son métodos empíricos de estimación de los errores estándar de la media (y de cualquier otro estimador). El procedimiento que siguen se basa en información tomada de la muestra original (que actúa como población) de la cual se generan muestras (llamadas réplicas) sobre las que se calcula el error.

En el método de Bootstrap, las muestras generadas reemplazan a la muestra original. En el método de Jackknife, las muestras generadas se basan en la eliminación sucesiva de una parte (pequeña) de la muestra de acuerdo con un criterio establecido, que puede variar dependiendo de la muestra. En las pruebas CRECER 1998 se usó el método de Jackknife.

### EL JACKKNIFE Y LOS CRITERIOS PARA EL REMUESTREO. USO DEL WESVAR

La idea básica de este método es seleccionar repetidamente submuestras de la muestra original mediante un criterio de “corte”. Luego se calcula el estadístico de interés (en nuestro caso la media) en cada una de estas submuestras, lo que forma una distribución del estimador, y finalmente se estima el error estándar de esta distribución de réplicas.

WesVar PC es el *software* que hemos usado para la estimación de las estadísticas de interés basado en el remuestreo<sup>13</sup>. Se dispone de dos métodos de Jackknife: 1 (JK1) y 2 (JK2). Nosotros hemos seleccionado JK1 suponiendo

11 Westat, Inc.: A User's Guide to WesVarPC. Versión 2.1. Rockville: Ewestat Inc., 1997.

12 Véase, por ejemplo, Efron, B. y J. Tibshirani: “An Introduction to the Bootstrap”. *Monographs on Statistics and Applied Probability*, 57. New York: Chapman & Hall, 1993; y Mooney, C.Z. y R.D. Duval: *Bootstrapping: a Non Parametric Approach to Statistical Inference*. Newbury Park, CA: Sage, 1993.

13 El *software* soporta diversos métodos de formación de réplicas, además de variantes dentro del mismo método de Jackknife. Más información sobre el *software* en: <http://www.westat.com/statsoft.html>.

un muestreo aleatorio. En este caso sabemos que no responde al diseño de la muestra<sup>14</sup>. Sin embargo, deseamos conocer si este método proporciona estimadores confiables que puedan ser aproximados a los verdaderos.

Para JK1 establecimos tres criterios de "corte", es decir, tres maneras de segmentar iterativamente la muestra conformando réplicas de la muestra original. Estos criterios fueron el rendimiento, estratos y escuelas.

La fórmula para hallar la varianza total es la que sigue:

$$V(\hat{\theta}) = \frac{G-1}{G} \sum_{k=1}^G (\hat{\theta}_k - \hat{\theta})^2$$

donde:

$\theta$  es un estimador de interés (por ejemplo, la media).

$\hat{\theta}$  es el estimador de  $\theta$  de la muestra original.

$\hat{\theta}_k$  es el estimador k-ésimo de  $\theta$  basado en la réplica k.

G es el total de réplicas.

$V(\hat{\theta})$  es la varianza estimada de  $\theta$ .

En la versión Jackknife 1 (JK1) son identificados G subconjuntos de UPM. Estos subconjuntos son definidos de acuerdo con criterios de corte. Por ejemplo, G escuelas, G estratos, etcétera. Se recomienda que los subconjuntos formados sean grupos aleatorios de casi igual tamaño y que cada uno de ellos refleje de algún modo la muestra original.

### PROGRAMA DESARROLLADO POR LAS UMC QUE IMPLEMENTAN JK1

Una de las limitaciones que enfrentamos para aplicar el método JK1 es que la versión de WesVar (2.1) que tiene la UMC sólo soporta

hasta 256 UPM ( $G < 256$ ), restricción que determinó la necesidad de elaborar un programa propio que siga el mismo algoritmo de réplicas<sup>15</sup>. Este programa es flexible como para poder ser aplicado a otros criterios para un G cualquiera.

### 3 CONCLUSIONES

Una vez calculados los errores estándar con los dos enfoques descritos, fue posible comparar los métodos de remuestreo con la referencia ideal que desarrolla la fórmula según el diseño muestral. Es importante notar que para todos los casos, y con fines de consistencia en la comparabilidad, se emplearon los rendimientos en su escala original sin ponderar.

Las siguientes son las principales conclusiones:

1. Cuando se comparan los diferentes criterios para formar las réplicas en JK1, del cuadro 2 se concluye que:

- Usando el criterio "rendimiento" se obtienen sobreestimaciones altas respecto del diseño.
- Usando el criterio "estratos" se obtienen subestimaciones respecto del diseño.
- Usando "escuelas" como criterio se obtienen sobreestimaciones moderadas respecto del diseño.

En primer lugar, descartamos el criterio (b) de "estratos" porque subestima los errores y además porque los estratos no se acomodan a todas las desagregaciones planificadas. Esto surge de la mirada de las estimaciones para Arequipa y Amazonas.

Consideramos que las sobreestimaciones de errores son ventajosas en el sentido de que son más conservadoras para certificar diferencias significativas entre grupos. Para tener un mejor margen de decisión calculamos los errores para más departamentos. En este sentido optamos por el tercer criterio (c), pues presenta la mejor aproximación<sup>16</sup>.

Es importante anotar que en informes anteriores de la UMC se estimaron errores estándar basados en el criterio (a). Esta solución, dentro de los límites del WesVar, fue bastante

14 Un diseño apropiado sería Jkn, donde n denota más de 2 unidades primarias por estrato y recoge eventualmente la idea de un muestreo estratificado. Esto está implementado en la última versión de WesVar.

15 Véase el anexo 2.

16 Esto se logró mediante el programa diseñado por la UMC.

conservadora para certificar las diferencias entre las desagregaciones (departamentos, gestión, etcétera).

2. Cuando se compara el programa JK1 de WesVar y de la UMC no se encuentran diferencias en las estimaciones obtenidas.

3. Una estimación bastante aproximada pero conservadora de los errores estándar de las medias de rendimiento está dada por el método JK1 con criterio de escuelas desarrollado por la UMC.

4. Cuando se comparan las variaciones en las relaciones de orden de los errores estándar de departamentos según escuelas (EEE), según rendimiento (EER), por muestreo alea-

torio simple (S) y a partir del diseño (EEDIS) a través de un análisis correlacional, se concluye que el error estándar basado en el criterio de escuelas se correlaciona significativamente con el error estándar basado en el diseño.

Finalmente, sería recomendable que la UMC contara con la última versión de WesVarPC, ya que debido a su mayor flexibilidad podría adaptarse a los diseños muestrales futuros y lograr así el análisis apropiado de las muestras complejas que desarrolla la Unidad. Por lo pronto, el error estándar basado en JK1 según escuelas es una solución aproximada que reserva las relaciones de orden del error verdadero debido al diseño.

<b>Cuadro 2</b> <b>Comparación de la estimación del error estándar de la media de rendimiento en Arequipa y Amazonas basado en el método JK1 para diferentes criterios y utilizando el diseño de la muestra</b>						
	JK1 (Programa WesVar)			JK1 (Programa UMC)	Diseño	Muestreo aleatorio simple
	Rendimiento	Estrato	Escuela			
Lima	2,878		0,860	0,860	0,730	0,252
Callao	2,961		2,377	2,377	1,187	0,710
Apurímac	2,686		2,002	2,002	1,216	0,635
Ayacucho	2,969		2,539	2,539	1,312	0,697
Amazonas	2,961	1,163	1,959	1,959	1,436	0,653
Cajamarca	2,917		2,451	2,451	1,492	0,692
Arequipa	2,852	0,748	1,586	1,586	1,495	0,477
Ancash	3,044		2,680	2,680	1,609	0,735

<b>Cuadro 3</b> <b>Correlaciones de los errores estándar según diferentes métodos</b>					
Correlaciones					
		EER	EEE	EEDIS	S
EER	Correlación de Pearson	1,000	,415	,218	,350
	Sig. (2-tailed)		,307	,604	,395
	N	8	8	8	8
EEE	Correlación de Pearson	,415	1,000	,742*	,933**
	Sig. (2-tailed)	,307		,035	,001
	N	8	8	8	8
EEDIS	Correlación de Pearson	,218	,742*	1,000	,619
	Sig. (2-tailed)	,604	,035		,102
	N	8	8	8	8
S	Correlación de Pearson	,350	,933**	,619	1,000
	Sig. (2-tailed)	,395	,001	,102	
	N	8	8	8	8

\* Correlación significativa al 5% (2-tailed).

\*\* Correlación significativa al 1% (2-tailed).



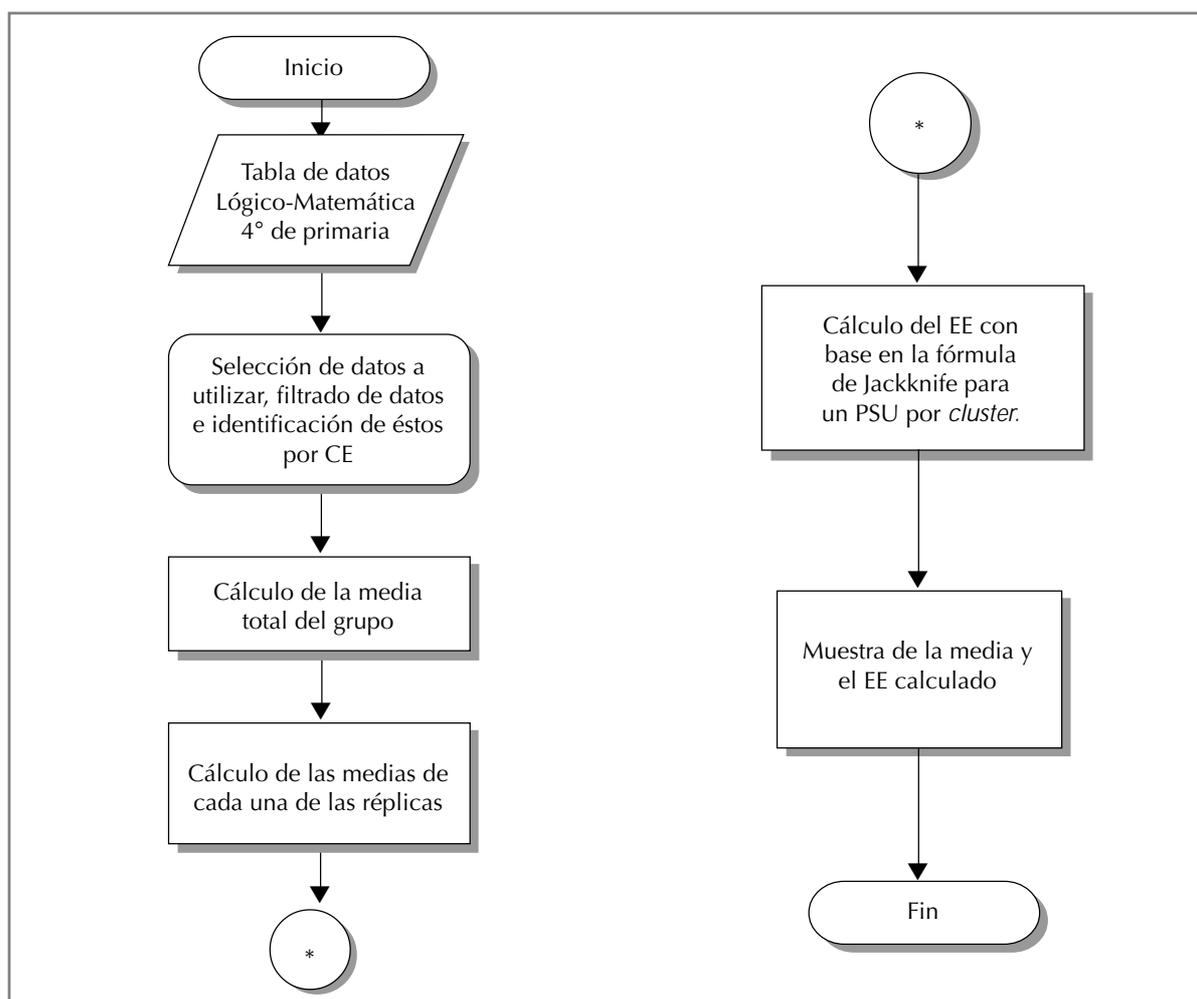


## ANEXO 2

### ALGORITMO DEL PROGRAMA QUE CALCULA EE UTILIZANDO LA TÉCNICA DE REMUESTREO JACKKNIFE

El único insumo que se utilizó fue la tabla de datos de cuarto de primaria para Lógico-Matemática, en la que debían estar los campos que identificaban cada una de las características de los CE, además de las variables de disgregación.

El algoritmo se representa en el siguiente esquema:



# EVALUACIÓN PSICOMÉTRICA DE LAS PREGUNTAS DE LAS PRUEBAS CRECER 1998

JORGE BAZÁN  
CESAR MILLONES

Esta sección del documento tiene por finalidad reportar la evaluación de las preguntas de las pruebas CRECER 1998 desde el punto de vista psicométrico, es decir, revisar los criterios utilizados para determinar si estas preguntas son óptimas para su inclusión en el cálculo del puntaje total de las pruebas; y complementa un informe<sup>1</sup> anterior que enfoca las características de cada prueba.

La evaluación de las preguntas se basa en la metodología descrita en el primer acápite. Allí se expone el conjunto de criterios que nos permitirán presentar los resultados del segundo acápite. En éste se muestran las preguntas que arrojaron mayor problema, y se ilustra en cada caso los criterios y las decisiones tomadas. El tercer acápite resume las conclusiones derivadas de los resultados encontrados. El anexo 1 detalla aún más la evaluación de las preguntas observadas. El anexo 2 es un breve glosario de términos para los lectores que quieran familiarizarse o recordar los términos técnicos empleados en este documento.

## 1 METODOLOGÍA PARA EL ANÁLISIS DE LAS PREGUNTAS

El objetivo del análisis de las preguntas es eliminar aquellas que no satisfacen un conjunto de propiedades que hacen de la prueba un instrumento eficaz y no sesgado en la estimación del rendimiento. Estas propiedades están supeditadas al marco referencial teórico

(modelo) que se adopte para el análisis de las pruebas. Como se sabe, los modelos más usados son los de la Teoría Clásica de los Tests y los modelos de la familia de la Teoría de Respuestas a Ítems o TRI<sup>2</sup>.

Lo común de estos enfoques es que se parte del supuesto de que existe un verdadero puntaje subyacente en los examinados que es necesario estimar a través de los datos observados o puntajes obtenidos en las pruebas. Las diferencias entre los puntajes verdaderos y los puntajes observados constituyen los errores de medición. Estos errores tienen diversas fuentes de origen que es de interés explicar antes de hacer uso de los puntajes obtenidos.

Una dimensión de la evaluación de las preguntas es su precisión o su consistencia. Otras dimensiones, menos cuantitativas, tienen que ver con la forma como se ha construido la pregunta y cómo se ha administrado<sup>3</sup>.

- 1 Millones, O. y J. Bazán: Evaluación psicométrica de las pruebas CRECER 1998. Lima: Ministerio de Educación-Unidad de Medición de Calidad Educativa, 2000. Documento de trabajo incluido en la última sección de este volumen.
- 2 Para una revisión de estos conceptos, véase por ejemplo Muñoz Fernández, J.: *Teoría de respuesta a los ítems. Un nuevo enfoque en la evolución psicológica y educativa*. Madrid: Ediciones Pirámide S.A.; y, del mismo autor: *Teoría Clásica de los Tests*.
- 3 Estos aspectos son presentados en Millones y Bazán: *Evaluación psicométrica...*, *op. cit.*

Los factores determinantes de los errores de medición provienen de la construcción del ítem, de factores “intra-examinado” y de errores del examinador<sup>4</sup>. Las fuentes de error “intra-examinado” pueden incluir los cambios diarios del examinado cuando se le toma la prueba; por ejemplo, cambios en su salud, sus estados de ánimo, eficiencia, motivación y comprensión de las instrucciones. La posibilidad de estimar errores de estas diversas fuentes para todas las preguntas de la prueba mejora el poder de inferencia del puntaje total de los examinados en la población.

A partir del enfoque de la TRI se ha propuesto una variedad de modelos matemáticos para caracterizar el puntaje de una pregunta mediante la presentación de formas o funciones llamadas “curvas características”. El interés de este enfoque es mejorar la predicción (probabilidad) de acierto de un individuo ante una pregunta, una vez que se han controlado factores de dificultad, discriminación y habilidad de los examinados<sup>5</sup>.

En el análisis de las pruebas CRECER 1998 se han aplicado los criterios que han sido definidos por el enfoque clásico, como comparabilidad y variabilidad de la prueba, y que determinan un conjunto de indicadores referenciales que se presentan a continuación.

## CRITERIOS PARA EL ANÁLISIS DE LAS PREGUNTAS

El análisis de las preguntas se hace desde dos dimensiones complementarias: cualitativa y cuantitativa.

4 Véase AERA, APA, NCME: *Standars for Educational and Psychological Testing*, preparado por un comité conjunto de la American Educational Research Association, American Psychological Association y el National Council on Measurement Education. Washington: AERA, 1999.

5 Véase, por ejemplo, Muñiz: *Teoría de respuesta...*, op. cit., 1990.

6 Por ejemplo, Lazarte, A.: “Análisis de preguntas”. Separata del curso PSB234. Lima: Pontificia Universidad Católica del Perú-Facultad de Psicología, 1995.

## Dimensión cualitativa

Desde el punto de vista cualitativo, el análisis de las preguntas incluye la revisión de su forma (enunciado y alternativas de respuesta), atendiendo a la característica que presentan los distractores de las preguntas y la calidad de la clave o respuesta, es decir, si ésta está bien o mal especificada. Se incluye también el análisis de la distribución de las opciones en términos de porcentajes tanto para la respuesta como para los distractores.

Este tipo de análisis es parte del proceso de construcción de la prueba y está a cargo de los especialistas o grupo de expertos responsables del desarrollo de las pruebas.

## Dimensión cuantitativa

Desde el punto de vista cuantitativo, el análisis de las preguntas incluye el cálculo de los índices que corresponden a las propiedades psicométricas de las preguntas. Estas propiedades se definen de acuerdo con el modelo de análisis adoptado. Los criterios aquí presentados se refieren a la Teoría Clásica de los Tests: validez (correlación pregunta-prueba), dificultad, discriminación e índice de no respuesta.

El examen de estos indicadores nos permite concluir si las preguntas han sido elaboradas en forma óptima o no. En el caso favorable, los puntajes obtenidos serán válidos, confiables y diferenciables para los usos que se les desea dar. En lo que sigue se presentan los indicadores de la dimensión cuantitativa del análisis de las preguntas.

Es importante notar que las categorías presentadas para cada criterio (véase filas de los cuadros 1 al 4) son niveles referenciales propuestos por la UMC y basados en las apreciaciones encontradas en la literatura<sup>6</sup>, así como en las decisiones tomadas por la propia Unidad. En cada caso se definen y comentan estas categorías y se discuten sus implicaciones para el análisis de las preguntas.

En lo que sigue, las respuestas a las preguntas se consideran dicotómicas en tanto aciertan o no con la respuesta.

## CONFIABILIDAD DE LAS PREGUNTAS

El concepto clásico de validez está relacionado con la confiabilidad de una pregunta. El término validez de un instrumento de medida indica cuán bien mide lo que pretende medir<sup>7</sup>. La validez, si nos referimos a las preguntas, evalúa la capacidad de una pregunta para medir el constructo propuesto, esto es, si la pregunta mide lo que la prueba se propone medir. Un indicador de validez es la correlación pregunta-prueba. Éste es un indicador de la consistencia interna de la pregunta, pues toma como referencia el puntaje total de la prueba, asumiendo que este puntaje es un indicador del constructo a medir.

Otras medidas de confiabilidad de las preguntas podrían obtenerse con los resultados de probar las preguntas varias veces en el tiempo al mismo examinado, lo que debiera dar resultados consistentes. El error estándar de la pregunta es también otro indicador de su confiabilidad.

El criterio de validez interna o confiabilidad adoptado en CRECER 1998 es medido por la correlación (biserial) pregunta-prueba, tal como se presenta en los cuadros que siguen. El análisis de las preguntas CRECER 1998 se basó en el cálculo de la correlación pregunta-prueba. Este coeficiente tiene un rango teórico de -1 a 1 y, para los efectos de la evaluación, se establecieron en este rango categorías que sirvieron para clasificar las preguntas.

El cuadro 1 presenta las categorías que se usaron para evaluar la validez de las preguntas.

<b>Cuadro 1</b> <b>Clasificación de la validez según correlación pregunta-prueba</b>	
<b>Clasificación</b>	<b>Índice de validez</b>
Muy buena	0,20 - 1
Aceptable	0 - 0,19
Eliminable	< 0

## DIFICULTAD DE LAS PREGUNTAS

La dificultad de las preguntas se mide con base en el porcentaje de aciertos en las respuestas. El indicador varía entre 0 y 1. Una pregunta con una dificultad de 0,3 es más difícil que una pregunta con una dificultad de 0,8. En el primer caso sólo 30% acertó la respuesta; en el segundo, 80%. Es decir, y como se ha observado en la literatura, el indicador se mide en forma inversa a la dificultad de la pregunta.

Aplicado a las pruebas CRECER, el índice de dificultad nos indica el grado de dificultad de cada pregunta en las áreas seleccionadas. Si una pregunta tiene un índice de dificultad cercano a 0 ó 1, la pregunta debe ser alterada o descartada, pues no está dando información acerca de las diferencias entre las habilidades de los examinados.

Una pregunta tiene una dificultad media si fue respondida de manera correcta por aproximadamente 50% de los examinados, o sea, si su índice de dificultad es cercano a 0,5. Por lo general, los índices de dificultad entre 0,3 y 0,7 maximizan la información que el *test* provee sobre la diferencia entre los examinados.

El cuadro 2 presenta las categorías para determinar el grado de dificultad de las preguntas de las pruebas CRECER.

<b>Cuadro 2</b> <b>Clasificación del nivel de dificultad de las preguntas</b>	
<b>Clasificación</b>	<b>Índice de dificultad</b>
Muy fácil	0,75 - 1
Fácil	0,55 - 0,74
Intermedio	0,45 - 0,54
Difícil	0,25 - 0,44
Muy difícil	0,00 - 0,24

7 Véase Nunnally, J. e I. Berstein: *Teoría psicométrica*. México: McGraw-Hill, 1995 (3ª edición).

Para la selección de las preguntas al momento de elaborar las pruebas definitivas, la UMC dio prioridad a preguntas con índices de dificultad intermedia (de 0,45 a 0,54). Las preguntas fáciles y difíciles fueron consideradas en menor medida. No se tomaron en cuenta preguntas muy fáciles o muy difíciles (índice de dificultad mayor de 0,74 o menor de 0,25 respectivamente).

Además de la consideración del porcentaje de examinados que eligieron la opción correcta, también es importante el análisis del porcentaje de examinados que eligieron cada opción en cada pregunta. Las opciones incorrectas se llaman distractores.

### DISCRIMINACIÓN DE LAS PREGUNTAS

La discriminación de una pregunta se mide por el grado en que ayuda a ampliar las diferencias estimadas entre los que obtuvieron un puntaje total de la prueba relativamente alto y los que alcanzaron un puntaje relativamente bajo. El índice de este indicador (véase el glosario de términos, anexo 2) varía entre -1 y 1. Los valores positivos indican que la pregunta discrimina a favor del grupo superior, mientras los negativos lo hacen a favor del grupo inferior.

El cuadro 3 presenta las categorías del índice de discriminación usadas en el análisis de las preguntas.

La discriminación de un ítem es muy alta si su valor está entre 0,4 y 1, y muy baja si su valor está entre 0 y 0,19. Si los valores son negativos, la pregunta debe eliminarse.

Cuadro 3 Clasificación de la discriminación de las preguntas	
Clasificación	Índice de discriminación
Muy alta	0,40 - 1
Alta	0,30 - 0,39
Moderada	0,20 - 0,29
Muy baja	0 - 0,19
Eliminable	< 0

### NO RESPUESTA DE LAS PREGUNTAS

El índice de no respuesta de una pregunta se mide por la proporción de personas que no contestan la pregunta (no señalan ninguna opción como la correcta) o por la proporción de personas que, al indicar una opción como la correcta, la invalidan. El programa de lectura óptica de CRECER 1998 invalida una respuesta a una opción si hay además otras opciones elegidas, o cuando esta opción tiene un trazo que no es detectado por el programa.

El índice de no respuesta ( $Nr_i$ ) se define por:

$$Nr_i = 1 - p_i - q^*_i$$

donde  $p_i$  es la dificultad de la pregunta y  $q^*_i$  la proporción de personas que marcaron opciones erróneas.

### Categorías de no respuestas

El cuadro 4 presenta una clasificación de niveles de tasa de no respuesta de las preguntas de acuerdo con los índices de no respuesta.

Este indicador informa si el evaluado ha desarrollado o no contenidos relacionados con lo que la pregunta mide. Así, si el indicador es alto es muy probable que el evaluado no conteste la pregunta por desconocimiento del contenido evaluado por ella.

El que relacionáramos este indicador con el orden de la pregunta puede ser indicativo del tamaño de la prueba. La posibilidad de que el alumno no haya tenido suficiente tiempo para terminar la prueba, especialmente cuando el índice es más alto en las últimas

Cuadro 4 Clasificación de la no respuesta de las preguntas	
Clasificación	Índice de no respuesta
Adecuada	0 - 0,15
Aceptable	0,16 - 0,20
Tolerable	0,21 - 0,29
Eliminable	0,30 - 1

preguntas, justificaría un tamaño menor de la prueba. Si las últimas preguntas presentan una tendencia a mayores índices de no respuesta, podemos suponer que el alumno no revisó estas preguntas. En este caso la dificultad de las preguntas está subestimada, porque no se sabe si los que trabajan más despacio acertarían al responder a esta pregunta.

Para la selección de preguntas al momento de elaborar las pruebas definitivas, la UMC dio prioridad a preguntas con índices de no respuesta que estuvieran en la categoría “adecuada”. En menor medida consideró las preguntas en las categorías “tolerable” y “aceptable”. No se consideraron preguntas con índices de no respuesta por encima de 0,30.

Finalmente, para los análisis de las pruebas la respuesta de no acierto considera además la tasa de no respuesta. Si bien en el análisis de preguntas es necesario distinguir la no respuesta en los resultados finales, la no respuesta ha sido tomada como no acierto.

## PROCESO DE ANÁLISIS DE LAS PREGUNTAS

En el análisis de las preguntas de las pruebas CRECER 1998 se empleó (corrió) el programa ITEMAN (modelo clásico). Para cada prueba se obtuvo un reporte ITEMAN de las preguntas basado en el procesamiento de aproximadamente 16 700 registros (estudiantes) por prueba.

Se formaron equipos de especialistas para revisar el *output* (salidas) generado por estos programas. Los especialistas verificaron las propiedades óptimas descritas en los acápitales anteriores para cada pregunta de cada prueba. Con el ITEMAN se obtuvo la información de la dimensión cuantitativa, es decir, sobre los índices de discriminación, dificultad, no respuesta y validez (correlación biserial) de las preguntas. Un programa adicional permitió clasificar los índices en las categorías ya señaladas.

En la dimensión cualitativa se analizó información sobre distractores competitivos y respuestas correctas mal especificadas. Todos estos aspectos guiaron la decisión de los especialistas respecto de la inclusión o eliminación de las preguntas para el puntaje fi-

nal. El anexo 1 detalla el conjunto de problemas presentados en las preguntas de las distintas pruebas.

Se utilizó también el programa RASCAL (modelo Rasch de un parámetro). Con este programa se evaluó la posibilidad de extender el análisis para considerar las habilidades de los alumnos en futuras aplicaciones<sup>8</sup>. Los resultados generados por ambos programas se encuentran disponibles en los documentos “Evaluación del rendimiento estudiantil CRECER: Análisis de ítemes” para cada grado de educación primaria y secundaria<sup>9</sup>.

## 2 RESULTADOS

### Nota para identificar las columnas

Primaria 4°		6°	
(CI)	= Comunicación Integral	(L)	= Lenguaje
(LM)	= Lógico-Matemática	(M)	= Matemática
(PS)	= Personal Social	(CHS)	= Ciencias Histórico- Sociales
(CA)	= Ciencia y Ambiente	(CN)	= Ciencias Naturales
Secundaria 4°		5°	
(LL4)	= Lenguaje y Literatura	(LL5)	= Lenguaje y Literatura
(M4F1)	= Matemática Forma 1	(M5F1)	= Matemática Forma 1
(M4F2)	= Matemática Forma 2	(M5F2)	= Matemática Forma 2

## ÍNDICE DE VALIDEZ

Como se observa en los cuadros 5a y 5b, la validez de las preguntas medida por el coeficiente de correlación pregunta-prueba (corre-

8 Para la distinción teórica entre puntaje y habilidad, véase Muñiz: *Teoría de respuesta...*, op. cit., 1990.

9 Véase MED: “Evaluación del rendimiento estudiantil CRECER, 4° grado de educación secundaria: Análisis de ítemes”. Informe técnico de la UMC, junio de 1999.

Cuadro 5a Número de preguntas que caen dentro de la clasificación de validez (Primaria)								
Validez	4° de primaria				6° de primaria			
	CI	LM	PS	CA	L	M	CHS	CN
Muy buena	30	31	27	30	32	30	28	29
Aceptable		1				1	2	1
Eliminable						1		
<b>Total</b>	30	32	27	30	32	32	30	30

Cuadro 5b Número de preguntas que caen dentro de la clasificación de validez (Secundaria)						
Validez	4° de secundaria			5° de secundaria		
	LL4	MF1	MF2	LL5	MF1	MF2
Muy buena	37	25	24	37	24	25
Aceptable	3		1	2	1	
Eliminable						
<b>Total</b>	40	25	25	39	25	25

lación biserial) es “muy buena”. El rango de porcentajes de preguntas que caen dentro de este calificativo es alto para cada prueba y para cada grado. Así, en la tabla 5b se observa que para Lenguaje y Literatura de cuarto de secundaria hay 37 preguntas, de un total de 40 —esto es, el 92,5%—, que tienen el calificativo de “muy buena”. Por otro lado, el calificativo es de “muy buena” para el 100% de las preguntas en cuatro pruebas en primaria y dos en secundaria.

## ÍNDICE DE DIFICULTAD

Como se menciona en la metodología, en las pruebas piloto no se consideraron preguntas que fueron “muy fáciles” o “muy difíciles” y que son los rangos extremos de dificultad (índice de dificultad mayor de 0,74 o menor de 0,25 respectivamente). Con las pruebas finales se observó un número relativamente bajo

de preguntas que caen en estos rangos extremos. Como puede verse en los cuadros 6a y 6b, el porcentaje varía de 37,5% (12 preguntas sobre un total de 32) en Lógico-Matemática de cuarto de primaria a 4% (1 pregunta de 25) en Matemática Forma 1 de quinto de secundaria.

El caso de Lógico-Matemática de cuarto de primaria (37,5%) es, sin embargo, un caso aislado, pues, como se observa en el cuadro 6a, el número de preguntas en estos extremos es bajo para el resto de las pruebas.

## ÍNDICE DE DISCRIMINACIÓN

La discriminación de una pregunta es muy alta si su valor está entre 0,4 y 1; y es alta si el índice está entre 0,30 y 0,40 (véase metodología). Como se puede apreciar en los cuadros 7a y 7b, en estas dos categorías caen la mayoría de las preguntas. El porcentaje de

Cuadro 6a Número de preguntas que caen dentro de la clasificación de dificultad (Primaria)								
Dificultad	4° de primaria				6° de primaria			
	CI	LM	PS	CA	L	M	CHS	CN
Muy fácil	3	8	3	3	3	1	2	7
Fácil	12	5	15	11	13	10	16	15
Intermedio	8	8	4	9	11	7	8	1
Difícil	6	7	4	6	5	10	4	7
Muy difícil	1	4	1	1		4		
<b>Total</b>	<b>30</b>	<b>32</b>	<b>27</b>	<b>30</b>	<b>32</b>	<b>32</b>	<b>30</b>	<b>30</b>

Cuadro 6b Número de preguntas que caen dentro de la clasificación de dificultad (Secundaria)						
Dificultad	4° de secundaria			5° de secundaria		
	LL4	M4F1	M4F2	LL5	M5F1	M5F2
Muy fácil			1	2		1
Fácil	11	7	6	19	8	7
Intermedio	6	7	5	6	5	4
Difícil	19	9	12	9	11	12
Muy difícil	4	2	1	3	1	1
<b>Total</b>	<b>40</b>	<b>25</b>	<b>25</b>	<b>39</b>	<b>25</b>	<b>25</b>

Cuadro 7a Número de preguntas que caen dentro de los calificativos de discriminación (Primaria)								
Discriminación	4° de primaria				6° de primaria			
	CI	LM	PS	CA	L	M	CHS	CN
Muy alta	20	14	22	16	25	21	16	16
Alta	8	4	4	7	5	4	8	7
Moderada	2	8	1	6	2	3	5	6
Muy baja		6		1		3	1	1
Eliminable						1		
<b>Total</b>	<b>30</b>	<b>32</b>	<b>27</b>	<b>30</b>	<b>32</b>	<b>32</b>	<b>30</b>	<b>30</b>

<b>Cuadro 7b</b> Número de preguntas que caen dentro de los calificativos de discriminación (Secundaria)						
Discriminación	4° de secundaria			5° de secundaria		
	LL4	M4F1	M4F2	LL5	M5F1	MF2
Muy alta	15	13	16	20	17	16
Alta	13	7	5	11	6	8
Moderada	9	4	3	6	1	1
Muy baja	3	1	1	2	1	
Eliminable						
<b>Total</b>	40	25	25	39	25	25

preguntas que caen en el rango de “alta” y “muy alta” varía de 56,2% en Lógico-Matemática de cuarto de primaria a 96,3% para Personal Social en el mismo grado. Igualmente, el caso de Lógico-Matemática es un caso aislado de un nivel de discriminación bajo o moderado.

### ÍNDICE DE NO RESPUESTA

Como se observa en los cuadros 8a y 8b, los índices de no respuesta son adecuados para todos los niveles de la primaria, y aceptables para los niveles de la secundaria. En quinto de secundaria se presentan, sin embargo, dos

preguntas en el rango de “tolerable” para Matemática (forma 2).

### PREGUNTAS CON PROBLEMAS (PREGUNTAS OBSERVADAS)

En los cuadros 9 y 10 se identifican cuáles han sido las preguntas, para cada prueba y cada grado (de secundaria y de primaria respectivamente), que han sido observadas por las razones ya mencionadas. En el anexo 1 se detallan los problemas presentados para las preguntas observadas, así como las decisiones tomadas en cada caso. Cabe notar que para la primaria no se presenta el problema de no respuesta (éste no fue un problema para dicho nivel).

<b>Cuadro 8a</b> Número de preguntas que caen dentro de los calificativos de no respuesta (Primaria)								
No respuesta	4° de primaria				6° de primaria			
	CI	LM	PS	CA	L	M	CHS	CN
Adecuada	30	32	27	30	32	32	30	30
Aceptable								
Tolerable								
Eliminable								
<b>Total</b>	30	32	27	30	32	32	30	30

<b>Cuadro 8b</b> Número de preguntas que caen dentro de los calificativos de no respuesta (Secundaria)						
No respuesta	4° de secundaria			5° de secundaria		
	LL4	M4F1	M4F2	LL5	M5F1	M5F2
Adecuada	40	21	23	39	23	19
Aceptable		4	2		2	4
Tolerable						2
Eliminable						
<b>Total</b>	40	25	25	39	25	25

<b>Cuadro 9</b> Identificación de las preguntas que presentaron problemas (Secundaria)						
Problema	4° de secundaria			5° de secundaria		
	LL4	M4F1	M4F2	LL5	M5F1	M5F2
<b>Validez</b>						
- Aceptable	10*, 18*, 24*		18*	2*, 28*, 30*	24*	
<b>Dificultad</b>						
- Pregunta muy fácil		23	4	2*, 4		15
- Pregunta muy difícil	6, 10*, 19, 25	16*, 25	16	17, 22, 28*, 30	24*	5
<b>Discriminación</b>						
- Muy baja	10*, 18*, 24*	16		2*, 28*, 30*	24*	
<b>No respuesta</b>						
- Aceptable		8, 20, 25			17, 25	8, 24
- Tolerable						25
<b>Distractores competitivos</b>	3, 4, 12, 13, 21, 24*, 27, 33, 34, 40	5, 8, 9, 20, 22, 24*	1, 7, 9, 11, 12, 22	12, 14, 35	1, 2, 3, 6, 7, 13, 14, 17	3, 6, 7, 8, 12, 16, 17, 19, 24
<b>Pregunta mal especificada</b>	2, 6, 7, 10*, 11, 18, 19, 23, 25, 29	4, 6, 16, 17, 25	2, 11, 16, 7*, 18*, 19, 23	1, 8, 15, 17, 22, 28*, 30*	9, 10, 11*, 12, 16, 23, 24*	4, 5, 9
<b>Conclusión</b>						
Preguntas observadas	10*, 18, 24*	16, 24*	17*, 18*	28*, 30*	11*, 24*	
Preguntas eliminadas	10*, 24*			28*, 30*		

\* Véase el anexo 1.

**Cuadro 10**  
**Identificación de las preguntas que presentaron problemas**  
**(Primaria)**

Problema	4° de primaria					6° de primaria				
	CI	LM	PS	CA	L	M	CHS	CN		
<b>Validez</b>										
- Aceptable		32*				20	4*,13		1*	
- Eliminable						7				
<b>Dificultad</b>										
- Pregunta muy fácil	9, 14, 21	2, 4, 5, 7, 8, 11, 17, 24	6, 9, 12	1, 5, 16	10, 30, 32	6	29, 30		7, 11, 17, 19, 23, 26, 29	
- Pregunta muy difícil	7	12*, 21, 26, 29	4	19		9, 13*, 21, 30				
<b>Discriminación</b>										
- Muy baja		5, 7, 12*, 21, 22, 32*		18*		13*, 20*, 30	4*		1*	
- Eliminable							7*			
<b>Distractores competitivos</b>	4, 12, 19, 24, 30	13, 14, 18	8, 16	7, 15, 19, 25, 30	13, 20	7*, 15, 22, 23, 29	4*, 17		12, 24, 26, 30	
<b>Mal especificada</b>	29	2, 15, 21, 22, 26, 29	4*, 11	17, 18*, 21	3*	13*, 14, 20*, 21, 30	24		29	
<b>Conclusión</b>										
- Preguntas observadas		6*, 32*	4*	18*	3*	7*, 13*, 20*	4*		1*	
- Preguntas eliminadas		6*, 32*			3*	7*, 13*				

\* Véase el anexo 1.

### 3 CONCLUSIONES

En general, el análisis de las preguntas basado en la metodología presentada arrojó buenos resultados tanto desde la perspectiva de los indicadores cualitativos como de los cuantitativos. En primaria y en secundaria, algunos resultados que inciden en el cálculo del puntaje final pueden sintetizarse como sigue:

#### En primaria:

En el cuarto grado, con excepción del área de Lógico-Matemática, todas las preguntas de las pruebas deben ser consideradas para el puntaje final. En la prueba de Lógico-Matemática deben excluirse dos preguntas.

En las pruebas de sexto grado se presentaron problemas con tres preguntas, una en Lenguaje y dos en Matemática. Estas preguntas deben excluirse en el puntaje final. En las otras pruebas todas las preguntas deben considerarse para el puntaje final.

#### En secundaria:

Dos preguntas de Lenguaje y Literatura de cuarto, y dos de Lenguaje y Literatura de quinto, deben ser excluidas para el puntaje final. En el resto de las pruebas todas las preguntas deben incluirse.

Los motivos de exclusión de las preguntas mencionadas se exponen en el anexo 1.

Si se quisiera replicar la construcción de pruebas para futuras aplicaciones, es preci-

so tomar en cuenta las siguientes recomendaciones:

- Los cambios sustantivos de las versiones originalmente planteadas y piloteadas tienen efectos importantes en los valores de las propiedades de las preguntas. En este sentido, es necesario que cuando se haga una modificación pospiloto, las versiones corregidas estén igualmente sujetas a pruebas de campo para su validación final. Éste es el caso de las preguntas eliminadas de las pruebas de primaria (Lógico-Matemática de cuarto, Matemática y Lenguaje de sexto), y de las de secundaria (Lenguaje y Literatura de cuarto y quinto).
- La misma recomendación se sugiere para otros cambios “menores” en la estructura de la pregunta (enunciado y distractores). Estas modificaciones no garantizan la idoneidad de la pregunta en una versión definitiva. Se trata de una pregunta prácticamente nueva. Es el caso específico de la pregunta eliminada en la prueba de Lenguaje en sexto grado de primaria.
- En el trabajo de las pruebas piloto los equipos de revisión de las preguntas observadas deben considerar tanto los indicadores psicométricos como los propios criterios pedagógicos para decidir la conveniencia de la eliminación de una pregunta. En ese sentido, el equipo debe ser integral y estar conformado por el especialista de psicometría, el constructor de la prueba y especialistas en el área evaluada.

# ANEXO 1

## ANÁLISIS DE LAS PREGUNTAS OBSERVADAS EN LAS PRUEBAS CRECER 1998

En los cuadros 9 y 10 de esta sección del documento de trabajo se han identificado las preguntas observadas de las pruebas CRECER 1998. Éstas presentan algunos de los siguientes problemas: índices de validez (correlación pregunta-prueba), discriminación, no respuesta y dificultad que no satisfacen las categorías de selección priorizadas por la UMC<sup>1</sup>; y porcentajes de respuesta en los distractores mayores a la respuesta correcta (posible respuesta correcta mal especificada) o competitiva (porcentaje cercano)<sup>2</sup>.

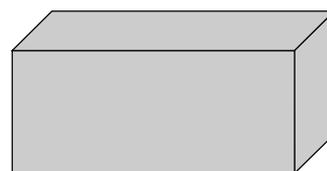
Las preguntas observadas fueron revisadas por un equipo compuesto por el responsable de la prueba, integrantes de su equipo pedagógico, el especialista en psicometría y el jefe de la Unidad. Esta evaluación implicaba explicar el comportamiento no esperado de la pregunta observada y tomar una decisión final sobre la pertinencia de conservar o no la pregunta para el puntaje total de la prueba respectiva.

### CUARTO GRADO DE PRIMARIA

#### Lógico-Matemática

6) ¿Cuántas caras tiene el siguiente cuerpo geométrico?

- A) 6
- B) 4
- C) 3
- D) 2



Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing		Point Biser.
							Low	High	
6	0-6	,45	,27	,23	A	,45	,34	,60	,23*
					B	,09	,14	,05	-,13
					C	,42	,46	,32	-,12
					D	,04	,06	,02	-,08
					Other	,01	,00	,00	-,08

1 Revisese las definiciones, los criterios de selección y las categorías de la dimensión cuantitativa en el anexo 2.

2 Revisese las definiciones de la dimensión cualitativa en el anexo 2.

*Dimensión cuantitativa:* El índice de discriminación es moderado.

*Dimensión cualitativa:* Se observa que el distractor C es competitivo con la respuesta correcta A. Esto se explica porque la pregunta tiene dos alternativas correctas. Esta pregunta incide en aspectos de percepción antes que en contenidos. La figura del cuerpo geométrico no es idónea, porque genera confusión; su percepción depende de la relación figura-fondo. En tales condiciones no se puede penar la respuesta C como incorrecta.

**Decisión:** No conservar la pregunta para el puntaje de la prueba de Lógico-Matemática.

- 12) Pepe tiene 2,40 metros de sogas, Alicia 2,4 metros y María dos metros con 40 centímetros. ¿Cuál de las respuestas es cierta?
- A) Pepe tiene la sogas más larga.
  - B) Alicia tiene la sogas más larga.
  - C) María tiene la sogas más larga.
  - D) Todas las sogas tienen la misma longitud.

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	High	Point Biser.
12	0-12	,17	,14	,21	A	,68	,66	,61	-,07
					B	,04	,07	,02	-,10
					C	,10	,11	,10	-,03
					D	,17	,12	,26	,21*
					Other	,02	,00	,00	-,12

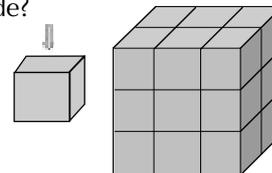
*Dimensión cuantitativa:* Índice de discriminación muy bajo.

*Dimensión cualitativa:* Se observa que el distractor A presenta mayor porcentaje de respuesta que la respuesta correcta D. Sin embargo, la pregunta mide la capacidad especificada. Se trata de una pregunta adecuadamente formulada, y el porcentaje de respuesta en el distractor A sólo refleja una respuesta incorrecta.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Lógico-Matemática.

- 32) ¿Cuántos cubos como el que señala la flecha hay en la figura grande?

- A) 9
- B) 12
- C) 15
- D) 18



Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	High	Point Biser.
32	0-32	,38	,18	,18	A	,25	,25	,22	-,04
					B	,11	,12	,09	-,03
					C	,18	,16	,19	,02
					D	,38	,30	,48	,18*
					Other	,08	,00	,00	-,25

*Dimensión cuantitativa:* La correlación pregunta-prueba es aceptable, y el índice de discriminación de la pregunta es muy bajo.

*Dimensión cualitativa:* El distractor A es competitivo con la respuesta correcta, lo que se explica porque ciertos estudiantes no identifican la segunda hilera de cubos. La pregunta no mide la capacidad especificada, pues incide en aspectos de percepción antes que en contenidos. Por ello, resulta inapropiada para los propósitos de la prueba de Matemática.

**Decisión:** No conservar la pregunta para el puntaje de la prueba de Lógico-Matemática.

## Personal Social

- 4) El Perú pasó por diferentes épocas a lo largo de su historia. Observa la línea de tiempo y contesta: ¿en qué época vivió Túpac Amaru II?  
(Se incluye una línea de tiempo que contiene las épocas de la historia del Perú.)
- A) Preincaica.  
B) Incaica.  
C) Virreinal.  
D) Republicana.

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing		Point Biser.
							Low	High	
4	0-4	,22	,24	,25	A	,16	,17	,12	-,05
					B	,48	,43	,46	,03
					C	,22	,13	,37	,25*
					D	,14	,26	,05	-,26
					Other	,01	,00	,00	-,05

*Dimensión cuantitativa:* El índice de discriminación es moderado.

*Dimensión cualitativa:* El distractor B presenta mayor porcentaje de respuesta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Personal Social.

## Ciencia y Ambiente

- 18) Cuando se tiende la ropa mojada al sol y luego se recoge seca, el agua ha cambiado del estado:
- A) líquido al estado gaseoso.  
B) líquido al estado sólido.  
C) gaseoso al estado líquido.  
D) sólido al estado líquido.

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing		Point Biser.
							Low	High	
18	0-18	,27	,18	,20	A	,27	,21	,39	,20*
					B	,44	,39	,45	,03
					C	,12	,19	,05	-,17
					D	,17	,20	,10	-,11
					Other	,01	,00	,00	-,05

*Dimensión cuantitativa:* Índice de discriminación muy bajo.

*Dimensión cualitativa:* El distractor B presenta mayor porcentaje de respuesta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Ciencia y Ambiente.

## SEXTO GRADO DE PRIMARIA

### Lenguaje

3) Los lenguados, las anchovetas y los pulpos forman parte del:

- A) Plancton.
- B) Océano.
- C) Necton.
- D) Bentos.

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.
3	0-3	,28	,25	,25	A	,16	,23	,08	-,17
					B	,44	,45	,41	-,05
					C	,28	,18	,42	,25*
					D	,09	,10	,07	-,05
					Other	,03	,00	,00	-,07

*Dimensión cuantitativa:* El índice de discriminación es moderado.

*Dimensión cualitativa:* El distractor B muestra mayor porcentaje de respuesta. Esto se explica porque la pregunta presenta dos alternativas correctas: la respuesta océano del distractor B también es correcta. Muchos alumnos pueden llegar a esta respuesta a partir de una lectura básica de la pregunta sin relacionarla con el texto. Para llegar a la respuesta C es necesario que los alumnos reconozcan al lenguado y a la anchoveta en la categoría de peces, y al pulpo en la categoría de seres marinos. Esto introduce un posible sesgo, que depende de si se conoce a estos seres marinos por cercanía geográfica.

**Decisión:** No conservar la pregunta para el puntaje total de la prueba de Lenguaje.

### Matemática

7) Sonia tiene 15 soles y Daniela tiene 6 soles más que Sonia, pero 8 soles menos que Juana. ¿Cuántos soles tiene Juana?

- A) 29
- B) 17
- C) 14
- D) 13

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.
7	0-7	,49	-,04	,00	A	,49	,53	,49	,00*
					B	,04	,07	,02	-,09
					C	,07	,09	,04	-,10
					D	,37	,27	,44	,11?
					Other	,02	,00	,00	-,09

*Dimensión cuantitativa:* La correlación pregunta-prueba es nula, y la pregunta tiene discriminación negativa. Dados estos dos criterios, la pregunta es eliminable.

*Dimensión cualitativa:* El distractor D es competitivo con la respuesta correcta, pero sólo refleja una respuesta incorrecta. Sin embargo, la pregunta no mide la capacidad especificada, pues el razonamiento resulta más determinante que el contenido evaluado para llegar a la respuesta correcta. Esta pregunta no discrimina en absoluto entre los que rinden más y los que rinden menos. Se trata de una pregunta modificada de la versión piloto cuyos indicadores no coincidieron con lo esperado.

**Decisión:** No conservar la pregunta para el puntaje total de la prueba de Matemática.

- 13) El precio de una blusa es S/. 30. Si Ana la compró con el 20% de descuento, ¿cuánto pagó por la blusa?
- A) S/. 50  
 B) S/. 24  
 C) S/. 20  
 D) S/. 6

Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.
13	0-13	,14	,18	,26	A	,30	,49	,10	-,34
					B	,14	,06	,25	,26*
					C	,18	,19	,14	-,06
					D	,34	,20	,50	,23
					Other	,04	,00	,00	-,08

*Dimensión cuantitativa:* Muy baja discriminación.

*Dimensión cualitativa:* Los distractores A, C y D presentan mayor porcentaje de respuesta. Esto se explica porque la expresión “descuento” del enunciado de la pregunta induce a la respuesta al azar y porque se trata de una pregunta no directa. Ciertos alumnos suman las cantidades presentadas porque no entienden de qué se está hablando; otros sólo calculan el porcentaje, sin entender tampoco la noción de descuento.

Se trata de una pregunta modificada de la versión piloto cuyos indicadores no coincidieron con lo esperado. La pregunta no es pertinente pues no corresponde exactamente a la tabla de especificaciones en la que se medía la resolución de problemas con el concepto de porcentaje y no sobre descuento con porcentaje.

**Decisión:** No conservar la pregunta para el puntaje total de la prueba de Matemática.

- 20) El 10% de los alumnos de un salón llegó tarde a clase. Si hay 40 alumnos, ¿cuántos llegaron temprano?
- A) 4 alumnos.  
 B) 36 alumnos.  
 C) 50 alumnos.  
 D) 90 alumnos.

Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.
20	0-20	,32	,17	,17	A	,44	,31	,53	,15
					B	,32	,24	,40	,17*
					C	,17	,34	,04	-,31
					D	,02	,03	,01	-,07
					Other	,05	,00	,00	-,14

*Dimensión cuantitativa:* La correlación pregunta-prueba es aceptable, y la pregunta presenta muy baja discriminación.

*Dimensión cualitativa:* El distractor A presenta mayor porcentaje de respuesta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Matemática.

## Ciencias Histórico-Sociales

- 4) La migración del campo a la ciudad ha sido intensa en los últimos años en el Perú. Entre sus muchas causas está:
- A) El crecimiento de la población total.
  - B) El terrorismo.
  - C) El comercio ambulatorio.
  - D) La sobrepoblación de las ciudades.

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low High		Point Biser.
4	0-4	,37	,04	,05	A	,32	,32	,30	-,02
					B	,37	,34	,38	,05*
					C	,12	,14	,12	-,04
					D	,17	,17	,19	,02
					Other	,01	,00	,00	-,06

*Dimensión cuantitativa:* La correlación pregunta-prueba es casi nula pero aceptable, y la pregunta tiene muy baja discriminación.

*Dimensión cualitativa:* El distractor A es competitivo con la respuesta correcta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Ciencias Histórico-Sociales.

## Ciencias Naturales

- 1) Si a un vaso que contiene agua le agregamos un poco de aceite y movemos el contenido con una cuchara, obtendremos:
- A) Una mezcla.
  - B) Una solución.
  - C) Un compuesto.
  - D) Una combinación.

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low High		Point Biser.
1	0-1	,56	,04	,04	A	,56	,54	,58	,04*
					B	,03	,03	,05	,05
					C	,06	,04	,09	,08?
					D	,34	,39	,28	-,10
					Other	,00	,00	,00	-,03

CHECK THE KEY  
A was specified, C works better

*Dimensión cuantitativa:* La correlación pregunta-prueba es casi nula pero aceptable, y la pregunta presenta muy baja discriminación.

*Dimensión cualitativa:* El distractor D es competitivo con la respuesta correcta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Ciencias Naturales.

## CUARTO DE SECUNDARIA

### Lenguaje y Literatura

10) ¿Qué sugiere el párrafo final del texto sobre los amigos de Cuéllar?

- A) Que se han adaptado al sistema.
- B) Que siguen siendo rebeldes a su modo.
- C) Que se han vuelto insensibles.
- D) Que ellos sí han madurado plenamente.

Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing		Point Biser.
							Low	High	
'10	0-10	,09	,05	,09	A	,09	,07	,12	,09*
					B	,05	,09	,02	-,12
					C	,06	,10	,02	-,14
					D	,77	,67	,83	,15?
					Other	,03	,00	,00	-,16
	CHECK THE KEY A was specified, D works better								

*Dimensión cuantitativa:* La correlación pregunta-prueba es casi nula pero aceptable, y la pregunta presenta muy baja discriminación.

*Dimensión cualitativa:* El distractor D muestra un mayor porcentaje de respuesta. Esto se explica porque la pregunta presenta dos alternativas correctas: el distractor D corresponde a una lectura literal y básica del texto presentado, más bien ligada a las vivencias personales de los alumnos (era la respuesta correcta en la etapa piloto); por otro lado, la alternativa A corresponde a una lectura interpretativa y más general del texto. Se trata de una pregunta modificada de la versión piloto cuyos indicadores no coincidieron con lo esperado.

**Decisión:** No conservar la pregunta para el puntaje total de la prueba de Lenguaje.

18) Marque la serie que defina mejor el poema:

- A) sensualidad–sonoridad–ritmo–colorido.
- B) sonoridad–crítica–ritmo–intimidad.
- C) sensualidad–monotonía–colorido–crítica.
- D) densidad–rigidez–sequedad–colorido.

Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing		Point Biser.
							Low	High	
18	0-18	,30	,12	,13	A	,30	,25	,38	,13*
					B	,37	,36	,37	,00
					C	,14	,13	,12	-,03
					D	,12	,11	,11	-,00
					Other	,07	,00	,00	-,20

*Dimensión cuantitativa:* La correlación pregunta-prueba es aceptable, y la pregunta presenta un índice de discriminación muy bajo.

*Dimensión cualitativa:* El distractor B muestra un mayor porcentaje de respuesta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Lenguaje.

24) ¿En qué tiempo y modo verbal se encuentra el verbo subrayado?

*Al salir de mi casa me encontré con un ciruelo cargado de frutos.*

- A) Pretérito imperfecto, modo indicativo.
- B) Pretérito indefinido, modo indicativo.
- C) Futuro imperfecto, modo subjuntivo.
- D) Pretérito imperfecto, modo subjuntivo.

Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.
24	0-24	,35	,11	,09	A	,28	,18	,42	,23?
					B	,35	,26	,37	,09*
					C	,15	,24	,07	-,18
					D	,11	,11	,10	-,02
					Other	,11	,00	,00	-,24
	CHECK THE KEY B was specified, A works better								

*Dimensión cuantitativa:* La correlación pregunta-prueba es aceptable, y la pregunta presenta muy baja discriminación.

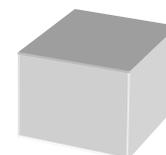
*Dimensión cualitativa:* El distractor A presenta un porcentaje de respuesta competitivo con la respuesta correcta. La categoría "indefinido" en la respuesta B parece extraña entre las alternativas ("perfecto simple" en la versión piloto), y ciertos alumnos, por estrategia de descarte, pueden haberla considerado correcta, mientras otros, con la misma estrategia, pueden haberla considerado incorrecta. Esto desvirtúa el propósito de la pregunta, la que presenta muy baja discriminación y poca correlación con la prueba. Se trata de una pregunta modificada de la versión piloto cuyos indicadores no coincidieron con lo esperado en la versión definitiva.

**Decisión:** No conservar la pregunta para el puntaje total de la prueba de Lenguaje.

## Matemática Forma 1

16) Se desea forrar totalmente la figura cuyas medidas son: 20 cm de largo, 10 cm de ancho y 10 cm de altura. ¿Cuántos centímetros cuadrados de papel se necesitan?

- A) 600 cm<sup>2</sup>
- B) 800 cm<sup>2</sup>
- C) 1000 cm<sup>2</sup>
- D) 2000 cm<sup>2</sup>



Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.
16	0-16	,11	,15	,25	A	,21	,23	,19	-,04
					B	,27	,30	,20	-,10
					C	,11	,05	,20	,25*
					D	,29	,24	,34	,07
					Other	,11	,00	,00	-,18

*Dimensión cuantitativa:* Muy baja discriminación.

*Dimensión cualitativa:* Los distractores A, B y D presentan mayor porcentaje de respuesta, pero esto sólo refleja respuestas incorrectas pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Matemática.

- 24) El área de un cuadrado es  $81 \text{ cm}^2$ . El área de otro cuadrado cuyo lado es igual a la diagonal del primero es:
- A)  $162 \text{ cm}^2$
  - B)  $81 \text{ cm}^2$
  - C)  $114,21 \text{ cm}^2$
  - D)  $161 \text{ cm}^2$

Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing		Point Biser.
							Low	High	
24	0-24	,42	,25	,25	A	,42	,30	,55	,25*
					B	,31	,30	,27	-,04
					C	,05	,06	,05	-,03
					D	,06	,07	,05	-,04
					Other	,16	,00	,00	-,24

*Dimensión cuantitativa:* El porcentaje de no respuesta (otros) es aceptable.

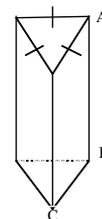
*Dimensión cualitativa:* El distractor B presenta un porcentaje competitivo con la respuesta correcta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada. El porcentaje de no respuesta de la pregunta se debe a su ubicación y dificultad (posiblemente ciertos estudiantes no alcanzaron a leer la pregunta o la obviaron); de esta manera se considera una respuesta incorrecta.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Matemática.

## Matemática Forma 2

- 17) El gráfico mostrado es un prisma regular cuya base es un triángulo equilátero.  $\overline{AB} = 12 \text{ cm}$  y  $\overline{BC} = 4 \text{ cm}$ . Hallar el volumen del prisma.

- A)  $24\sqrt{3} \text{ cm}^3$
- B)  $48\sqrt{3} \text{ cm}^3$
- C)  $16\sqrt{3} \text{ cm}^3$
- D)  $64\sqrt{3} \text{ cm}^3$



Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing		Point Biser.
							Low	High	
17	0-17	,32	,29	,27	A	,12	,12	,11	-,02
					B	,32	,19	,47	,27*
					C	,32	,39	,21	-,16
					D	,07	,06	,08	,04
					Other	,17	,00	,00	-,14

*Dimensión cuantitativa:* El porcentaje de no respuesta es aceptable.

*Dimensión cualitativa:* El distractor C presenta mayor porcentaje de respuesta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada. El porcentaje de no respuesta aceptable en la pregunta se debe a la dificultad de la pregunta y no a su ubicación (posiblemente ciertos estudiantes prefirieron no contestarla); así, se considera una respuesta incorrecta.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Matemática.

18) Si  $|9 - x| = 6$ , entonces  $|x - 9|$  será igual a:

- A) 6
- B) -6
- C) 12
- D) -24

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low High		Point Biser.
18	0-18	,27	,17	,18	A	,27	,18	,35	,18*
					B	,57	,52	,59	,02
					C	,07	,10	,05	-,08
					D	,04	,07	,01	-,11
					Other	,05	,00	,00	-,22

*Dimensión cuantitativa:* La correlación pregunta-prueba es aceptable.

*Dimensión cualitativa:* El distractor B presenta mayor porcentaje de respuesta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Matemática.

## QUINTO DE SECUNDARIA

### Lenguaje y Literatura

2) ¿Cuál es la principal virtud de Job?

- A) La templanza.
- B) La fe.
- C) La sabiduría.
- D) La paciencia.

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low High		Point Biser.
2	0-2	,88	,12	,16	A	,03	,04	,02	-,04
					B	,88	,81	,93	,16*
					C	,04	,07	,01	-,13
					D	,05	,06	,03	-,05
					Other	,01	,00	,00	-,12

*Dimensión cuantitativa:* La correlación pregunta-prueba es aceptable, y presenta muy baja discriminación.

*Dimensión cualitativa:* Ningún problema. La pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Lenguaje.

28) ¿Cuál de los elementos del afiche transmite el principal contenido del mensaje?

- A) Las manos enguantadas.
- B) La sombra de todo.
- C) La jeringa completa.
- D) La gota de sangre.

Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low High		Point Biser.
28	0-28	,21	,10	,11	A	,21	,18	,27	,11*
					B	,08	,11	,06	-,07
					C	,13	,17	,10	-,08
					D	,54	,44	,56	,10
					Other	,04	,00	,00	-,25

*Dimensión cuantitativa:* La correlación pregunta-prueba es aceptable, y la discriminación muy baja.

*Dimensión cualitativa:* El distractor D muestra un mayor porcentaje de respuesta. Esto se explica porque la pregunta presenta dos alternativas correctas: en el enunciado se indaga sobre el elemento del afiche que “transmite el principal contenido del mensaje”. Esta última frase no sólo se refiere al mensaje del afiche sino también al *contenido de ese mensaje*; y a no cualquier contenido, sino el principal. Es decir, se indaga por una “transmisión” de un aspecto muy elaborado y no evidente. No sabemos cuál es, para los alumnos, el contenido principal del afiche. Por ello, las respuestas A y D son correctas. La alternativa A corresponde al elemento principal del afiche, pero no necesariamente al contenido principal. La alternativa D corresponde al contenido principal, pero no necesariamente es el elemento principal.

Además, parece haber interferencia de la pregunta anterior, pues la respuesta correcta es “Recomendar precaución en el manejo de sangre”<sup>3</sup>.

Se trata de una pregunta modificada de la versión piloto cuyos indicadores no coincidieron con lo esperado en la versión definitiva.

**Decisión:** No conservar la pregunta para el puntaje total de la prueba de Lenguaje.

30) La imagen parece reflejar:

- A) El comportamiento típico de los adultos.
- B) La decepción de los derrotados.
- C) La destrucción que siempre causan los fanáticos.
- D) El entusiasmo de un grupo humano.

Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low High		Point Biser.
30	0-30	,36	,11	,10	A	,05	,11	,01	-,16
					B	,17	,17	,13	-,05
					C	,36	,28	,39	,10*
					D	,38	,32	,46	,12?
					Other	,04	,00	,00	-,27

CHECK THE KEY  
C was specified, D works better

3 La pregunta anterior es: “Este afiche se encuentra en el laboratorio de un hospital. ¿Cuál crees que es la intención de tenerlo allí?”.

*Dimensión cuantitativa:* La correlación pregunta-prueba es aceptable, y la pregunta presenta muy baja discriminación.

*Dimensión cualitativa:* El distractor D muestra un mejor porcentaje de respuesta. Esto se explica porque la pregunta presenta dos alternativas correctas. La alternativa D corresponde a una lectura comprensiva de la imagen mediada posiblemente por la vivencia del alumno. La alternativa C exige una lectura crítica, una postura, ante la imagen. Se trata de una pregunta modificada de la versión piloto cuyos indicadores no coincidieron con lo esperado.

**Decisión:** No conservar la pregunta para el puntaje total de la prueba de Lenguaje.

## Matemática Forma 1

11) La expresión trigonométrica  $\sin(90^\circ - A) - \cos(90^\circ - A)$  es igual a:

- A)  $\sin A - \cos A$
- B)  $-\sin A + \cos A$
- C)  $-\sin A - \cos A$
- D)  $\sin A + \cos A$

Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low High		Point Biser.
11	0-11	,32	,22	,23	A	,34	,35	,31	-,04
					B	,32	,21	,43	,23*
					C	,12	,13	,10	-,05
					D	,16	,21	,11	-,12
					Other	,06	,00	,00	-,11

*Dimensión cuantitativa:* Ningún problema.

*Dimensión cualitativa:* El distractor A presenta mayor porcentaje de respuesta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Matemática.

24) Calcular el área de la siguiente figura, si su longitud de arco mide 6 metros y su radio 3 metros.

- A)  $6 \text{ m}^2$
- B)  $18 \text{ m}^2$
- C)  $12 \text{ m}^2$
- D)  $9 \text{ m}^2$



Seq. No. Key	Item Statistics				Alternative Statistics				
	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low High		Point Biser.
24	0-24	,22	,12	,16	A	,06	,06	,06	-,00
					B	,47	,46	,46	-,02
					C	,11	,10	,12	,01
					D	,22	,16	,28	,16*
					Other	,13	,00	,00	-,19

*Dimensión cuantitativa:* La correlación pregunta-prueba es aceptable, y la pregunta presenta muy baja discriminación.

*Dimensión cualitativa:* El distractor B presenta mayor porcentaje de respuesta, pero esto sólo refleja una respuesta incorrecta pues la pregunta está adecuadamente formulada.

**Decisión:** Conservar la pregunta para el puntaje total de la prueba de Matemática.

## ANEXO 2

### GLOSARIO DE TÉRMINOS PSICOMÉTRICOS

El glosario complementa esta sección y la siguiente, Evaluación psicométrica de las pruebas CRECER 1998. Contiene los términos más usados en los análisis de preguntas desde el enfoque que corresponde a la Teoría Clásica de los Tests, y en el análisis de las propiedades psicométricas de las pruebas a partir de ese enfoque.

**Alfa de Cronbach:** Coeficiente que estima la confiabilidad de una prueba. Es apropiado para pruebas compuestas de preguntas dicotómicas o de selección múltiple. Se trata de un coeficiente de consistencia interna y es una cota inferior de la confiabilidad. Su fórmula es:

$$\alpha = \frac{k}{k-1} \left[ 1 - \frac{\sum_{i=1}^k \sigma_i^2}{\sigma_x^2} \right]$$

donde  $\sigma_i^2$  son las varianzas de las preguntas  $i$  y  $\sigma_x^2$  es la varianza de la prueba.

**Análisis factorial:** Procedimiento matemático de análisis de una matriz de correlaciones entre variables para determinar si un número limitado de factores (constructos o variables no observables) es suficiente para explicar las correlaciones.

**Análisis psicométrico de las preguntas:** Término general para los procedimientos destinados a evaluar o determinar la utilidad de una pregunta de una prueba (véase también índice de dificultad, índice de validez). El análisis usa las técnicas de la psicometría para el estudio de la dimensión cuantitativa y cualitativa de las preguntas de las pruebas. Un objetivo importante es eliminar las preguntas que no satisfacen los requerimientos deseados.

**Confiabilidad:** El grado en que una prueba mide algo consistentemente. Propiedad de una prueba que permite establecer si ésta proporciona la misma medida (puntaje) al volver a medir al mismo individuo bajo "condiciones similares" (invarianza para la propiedad medida). Los métodos para determinar la confiabilidad incluyen formas paralelas, prueba-reprueba, consistencia interna.

**Consistencia interna:** El grado en que las preguntas que contiene una prueba miden lo mismo (un mismo constructo). La confiabilidad de una prueba calculada por el coeficiente alfa de Cronbach, por ejemplo, es una medida de la consistencia interna.

**Constructo:** Definición operacional de una propiedad (variable) o característica de los evaluados. Por ejemplo, rendimiento en Matemática.

**Correlación:** Grado de relación o asociación entre dos variables, tales como una prueba y una medida de criterio, o una pregunta y la prueba a la que pertenece.

**Coefficiente de correlación:** Índice numérico del grado de relación entre dos variables. El coeficiente usualmente toma el rango de  $-1$  (relación perfecta negativa), pasa por  $0$  (ausencia total de una relación) y va hasta  $+1$  (relación perfecta positiva). Los dos tipos de coeficientes de correlación más comunes son la correlación producto momento de Pearson y el coeficiente punto-biserial.

**Coefficiente de correlación producto momento de Pearson:** Es una medida de la asociación lineal de dos variables continuas, y se define como:

$$\rho_{iX} = \frac{\sigma_{iX}}{\sigma_i \sigma_X}$$

donde:

$\rho_{iX}$ : Correlación pregunta-prueba.

$\sigma_X$ : Desviación estándar de los puntajes en la prueba de los evaluados.

$\sigma_i$ : Desviación estándar de los puntajes de la pregunta  $i$ .

$\sigma_{iX}$ : Covarianza entre la pregunta  $i$  y el puntaje de la prueba.

**Coefficiente de correlación punto-biserial:** Coeficiente de correlación calculado entre una variable dicotómica y una variable continua derivada del coeficiente de correlación producto-momento. Una fórmula computacional para la correlación punto biserial está dada por:

$$V_{i(1)} = \rho_{pbis} = \frac{\mu_{i+} - \mu_X}{\sigma_X} \sqrt{\rho_i(1-\rho_i)}$$

donde:

$V_{i(1)}$ : Índice de validez interna o simplemente índice de validez.

$\rho_{pbis}$ : Correlación biserial.

$\mu_{i+}$ : Media de los puntajes en la prueba para los evaluados que acertaron la pregunta  $i$ .

$\mu_X$ : Media de los puntajes en la prueba para todos los evaluados.

$\sigma_X$ : Desviación estándar de los puntajes en la prueba de los evaluados.

**Criterio:** Estándar o variable con el cual son comparados los puntajes en un instrumento psicométrico o por el cual ellos son evaluados. La validez de una prueba u otro procedimiento psicométrico usado en seleccionar o clasificar personas es determinado por su habilidad para predecir un criterio específico de conducta en la situación por la que las personas son seleccionadas o clasificadas.

**Diferenciabilidad:** Propiedad de una prueba que permite establecer su sensibilidad para diferenciar a los individuos a través de los puntajes (invarianza de las relaciones de orden entre los puntajes de los individuos sometidos a la prueba con respecto a la posesión del constructo).

**Dimensión cualitativa del análisis de preguntas:** En el caso de la UMC, se refiere a la selección de preguntas basada en el análisis de los distractores competitivos, y la respuesta correcta mal especificada. Las preguntas son revisadas por un grupo de expertos responsables de las pruebas, considerando la distribución de los porcentajes de respuesta correcta y de los distractores. El objeto de este análisis es asegurar la pertinencia de la pregunta para los propósitos de la evaluación teniendo en cuenta el posible mal funcionamiento de los distractores o del enunciado.

**Dimensión cuantitativa del análisis de preguntas:** En el caso de la UMC, se refiere a la selección de preguntas basada en el análisis de los índices psicométricos clásicos: dificultad, validez, discriminación y no respuesta. Se utilizan como criterio las categorizaciones de la literatura. El propósito es asegurar que la prueba presente la mayor diferenciabilidad y sea coherente con una evaluación basada en normas.

**Distractor:** Una de las opciones incorrectas en las preguntas con opciones múltiples.

**Distractor competitivo:** Es el caso de uno o más distractores de una pregunta de opción múltiple, cuyo porcentaje de respuesta es cercano al porcentaje de la respuesta correcta o dificultad de la pregunta. Es indicativo del funcionamiento inadecuado de una pregunta, pues la alternativa correcta no discrimina adecuadamente respecto de las otras alternativas. Por lo general, esto sucede cuando no se han seleccionado distractores adecuados.

**Estadísticas finales de la prueba:** Comprende el reporte del número de preguntas finales de las pruebas, el número de evaluados, las medias de los índices de dificultad, validez, discriminación y no respuesta de todas las preguntas finales de una prueba y el alfa de Cronbach final.

**Formas paralelas de pruebas:** Dos pruebas que tienen propiedades equivalentes en el sentido de que contienen el mismo número de preguntas de igual dificultad y están altamente correlacionadas. Los puntajes obtenidos por los examinados en una forma de la prueba son muy cercanos a los que obtendrían con la otra forma de la prueba.

**Índice de confiabilidad:** También conocido como coeficiente de confiabilidad de una prueba, es un índice entre 0 y 1 de la confiabilidad de una prueba (véase “Confiabilidad”). Los métodos para determinar la confiabilidad de una prueba incluyen la prueba-reprueba, formas paralelas y consistencia interna.

**Índice de dificultad:** Es el indicador más común en las pruebas de rendimiento. En el caso de las preguntas de elección múltiple, el promedio de la pregunta corresponde a la proporción de evaluados que contestan la pregunta correctamente.

Para la pregunta  $i$ , su dificultad o *índice de dificultad* está definido por la proporción  $p_i$ :

$$p_i = \frac{R}{N}$$

donde:

$P_i$ : Índice de dificultad.

R: Número de evaluados que respondieron correctamente la pregunta.

N: Número de evaluados.

#### *Categorías de dificultad*

El cuadro 1 presenta una clasificación de niveles de dificultad de acuerdo con los índices de dificultad.

Cuadro 1 Clasificación del nivel de dificultad de las preguntas dicotómicas <sup>1</sup>	
Clasificación	Índice de dificultad
Muy fácil	0,75 - 1
Fácil	0,55 - 0,74
Intermedio	0,45 - 0,54
Difícil	0,25 - 0,44
Muy difícil	0,00 - 0,24

#### *Justificación*

Una vez clasificada la dificultad, es posible jerarquizar las preguntas desde las más fáciles hasta las más difíciles, como es el caso de las pruebas de dificultad creciente.

Las dificultades de las preguntas están relacionadas con el estudio de las diferencias individuales. Para las preguntas dicotómicas, la varianza muestral de la pregunta tiene que ser considerada porque provee información acerca de las diferencias entre los evaluados. Una pregunta ofrece la mayor cantidad de información acerca de las diferencias entre evaluados cuando  $p_i = 0,5$  y, por lo tanto, la varianza es maximizada<sup>2</sup>. Por esto se recomienda seleccionar preguntas en un rango de índices de dificultad alrededor de 0,5 (algunos autores sugieren entre 0,3 y 0,7).

1 Tomado de Ecurra, L.: Análisis estadístico de preguntas. Separata del Seminario de Construcción de Pruebas I. Lima: UNMSM-Facultad de Psicología, 1995.

2 La varianza de la pregunta es máxima para dicho valor e igual a 0,25.

### Crterios de seleccin

Para la seleccin de preguntas al momento de elaborar las pruebas definitivas, la UMC dio prioridad a preguntas con ndices de dificultad que estuvieran en la categora intermedia (véase el cuadro 2 de este anexo). En menor medida tom en cuenta las preguntas en las categoras fciles y difciles. No se consideraron preguntas muy fciles o muy difciles.

**ndice de discriminacin:** Es una medida de con cuanta efectividad una pregunta discrimina o diferencia entre los examinados que tienen alto puntaje de aquellos que obtienen un bajo puntaje.

Este ndice se determina con la distribucin de puntajes de rendimiento en uno o dos puntajes de corte, y se clasifica a los evaluados en grupos que acumulan puntajes por debajo y por encima de esos puntajes de corte. Por ejemplo, dividir en dos mitades y clasificar evaluados en la mitad inferior y superior, dividir en el tercio superior y el tercio inferior, etcétera. Podra ser en los siguientes grupos:

- Grupo superior, que contiene el 27% de los casos con puntajes totales ms altos.
- Grupo intermedio, que contiene el 46% de los casos con puntajes intermedios.
- Grupo inferior, que contiene el 27% de los casos con puntajes totales ms bajos.

De ellos separan los grupos extremos.

Una vez que los dos grupos han sido identificados, el *ndice de discriminacin*  $\sigma_{ix}$  de la pregunta  $i$  se define como:

$$D_i = P_{i(S)} - P_{i(I)}$$

donde:

$D_i$ : Proporcin de evaluados en el grupo superior que contestó la pregunta  $i$  correctamente.

$P_{i(I)}$ : Proporcin de evaluados en el grupo inferior que contestó la pregunta  $i$  correctamente.

De otra manera, en el grupo superior y el grupo inferior se calcula separadamente para cada pregunta el porcentaje de evaluados que la responden correctamente; ambos datos se restan, y el resultado final es la discriminacin porcentual de la pregunta.

Puede variar entre  $-1$  y  $1$ . Los valores positivos indican que la pregunta discrimina en favor del grupo superior, mientras que los valores negativos indican que la pregunta discrimina a favor del grupo inferior.

### Categoras de discriminacin

El cuadro 2 presenta una clasificacin de niveles de discriminacin de las preguntas de acuerdo con los ndices de discriminacin.

Cuadro 2 Clasificacin de la discriminacin de las preguntas dicotmicas <sup>3</sup>	
Clasificacin	ndice de discriminacin
Muy alta	0,40 - 1
Alta	0,30 - 0,39
Moderada	0,20 - 0,29
Muy baja	0 - 0,19
Eliminable	< 0

### Justificacin

Mide el grado con el cual la pregunta permite establecer diferencias entre los evaluados que tengan niveles altos y bajos de acierto (grupos de alto y bajo rendimiento) en las pruebas.

3 Adaptado de Ecurra, L.: Anlisis estadstico., *op. cit.*

Desde el punto de vista de la Teoría Clásica de las Pruebas (evaluación de normas), uno de los propósitos de toda prueba es brindar información acerca de las diferencias en el rendimiento medido por la prueba con relación a estratos considerados en la muestra. Por lo tanto, interesa obtener índices que muestren cuán efectivamente una pregunta discrimina entre evaluados que tienen alto puntaje en el rendimiento y los que tienen un puntaje bajo. A falta de criterio externo, el puntaje total de la misma prueba es utilizado para separar grupos de alto y bajo rendimiento. Así, el propósito es identificar preguntas que los evaluados que obtienen altos puntajes contestan bien con una alta probabilidad, mientras que los evaluados con bajos puntajes contestan mal. Una pregunta que es contestada igualmente bien por evaluados con alto y bajo puntaje no discrimina bien entre estos dos grupos, y no sería de utilidad. Una pregunta que es bien contestada por los evaluados de puntaje bajo y mal por los de puntaje alto, es una pregunta con *discriminación negativa*, y tampoco es deseable.

#### *Criterios de selección*

Para la selección de preguntas al momento de elaborar las pruebas definitivas, la UMC dio prioridad a preguntas con índices de discriminación que estuvieran en la categoría muy alta y alta. En menor medida consideró las preguntas en la categoría moderada. No se consideraron preguntas con índices de discriminación negativos y en la categoría de muy baja discriminación.

Lazarte<sup>4</sup> menciona que las preguntas con discriminación muy alta funcionan bien; las de alta discriminación necesitan poca o ninguna revisión; las de discriminación en estudio necesitan revisión; las de muy baja son poco apropiadas y deben ser reformadas; y, por último, nada puede hacerse con las eliminables.

**Índice de no respuesta:** La tasa de no respuesta en una pregunta de las pruebas de rendimiento está formada por la proporción de personas que no contestan la pregunta (no señalan ninguna opción como la correcta) y la proporción de personas que, aun cuando indican una opción como la correcta, la invalidan por doble marca de opciones o por mal trazo de la opción correcta, de manera que no es leída apropiadamente. El índice de no respuesta se define como:

$$Nr_i = 1 - p_i - q_{i(l)}$$

En este caso,  $q_{i(l)}$  es la proporción de personas que no aciertan pues marcaron opciones erróneas.

#### *Categorías de no respuestas*

En el cuadro 3 se presenta una clasificación de niveles de tasa de no respuesta de las preguntas de acuerdo con los índices de no respuesta.

Cuadro 3 Clasificación de la no respuesta de las preguntas dicotómicas	
Clasificación	Índice de no respuesta
Adecuada	0 - 0,15
Aceptable	0,16 - 0,20
Tolerable	0,21 - 0,29
Eliminable	0,30 - 1

#### *Justificación*

La importancia de este indicador es que resulta informativo en el caso de que el evaluado no haya desarrollado contenidos relacionados con lo que la pregunta mide. También en el caso de que relacionáramos este indicador con el orden de la pregunta, puede ser indicativo del tamaño de la prue-

4 Lazarte, A.: "Análisis de preguntas", *op. cit.*

ba. Si las últimas preguntas presentan una tendencia a mayores índices de no respuesta, podemos suponer que ellas se explican porque el alumno no revisó estas preguntas y por tanto la dificultad de las preguntas está subestimada porque no se conoce si acertarían en esta pregunta los que trabajan más despacio.

#### *Criterios de selección*

Para la selección de preguntas al momento de elaborar las pruebas definitivas, la UMC dio prioridad a preguntas con índices de no respuesta que estuvieran en la categoría “adecuadas”. En menor medida consideró las preguntas en las categorías tolerable y aceptable. No se consideraron preguntas con índices de no respuesta por encima de 0,30.

Para los propósitos de la UMC, la respuesta de no acierto está formada por la tasa de no respuesta más  $q_{i(t)}$ , que es la proporción de respuesta a las opciones erróneas:

$$q_i = Nr + q_{i(t)} = 1 - p_i$$

**Índice de validez:** Índice que muestra la correlación del puntaje en la pregunta con el puntaje total obtenido en la prueba. En general este índice se denomina *correlación pregunta-prueba* y es un índice de validez interna debido a que el puntaje en la prueba es usado como criterio. La correlación entre la pregunta y el total es la correlación de Pearson. En el caso de las preguntas dicotómicas se denomina correlación biserial.

#### *Categorías de validez*

El cuadro 4 presenta una clasificación de niveles de validez de las preguntas de acuerdo con los índices de validez.

Cuadro 4 Clasificación de la validez de las preguntas dicotómicas	
Clasificación	Índice de validez
Muy buena	0,20 –1
Aceptable	0 - 0,19
Eliminable	< 0

#### *Justificación*

La *correlación pregunta-prueba* es un índice de validez interna debido a que el puntaje en la prueba es usado como criterio interno. Calcula el grado con el cual una pregunta mide válidamente aquella habilidad que se quiere medir. Si la pregunta se correlaciona significativamente con la prueba, entonces la pregunta mide un aspecto específico de lo que mide en general la prueba.

Mientras más se acerca el coeficiente de correlación al valor 1, el índice de validez de la pregunta será mejor, y se acepta como criterio empírico que éste debe ser por lo menos mayor o igual que 0,20 para tomar en cuenta la pregunta. Se quiere asegurar la buena correspondencia entre lo que mide la pregunta y lo que mide la prueba.

La correlación punto-biserial se utiliza en las situaciones en las cuales una de las variables que se correlaciona es continua y la otra dicotómica. Éste es el caso de las correlaciones pregunta-prueba en las pruebas de opción múltiple. Es la correlación producto-momento de Pearson entre la pregunta dicotómica (0 ó 1) y la prueba continua.

#### *Criterios de selección para la validez*

Para la selección de preguntas al momento de elaborar las pruebas definitivas, la UMC privilegió preguntas con índices de validez interna que estuvieran en la categoría de muy buena. En menor

medida consideró las preguntas en la categoría aceptable. No se consideraron preguntas con índices de validez interna negativos.

**Muestra representativa:** Subconjunto de mediciones seleccionadas de una población en la que la forma de la muestra es semejante a la población en todas las formas significativas.

**Preguntas de opción múltiple:** Es el tipo de pregunta que puede tener más de dos alternativas de respuesta pero sólo una es correcta. Consta, por lo general, de tres partes: *enunciado*, la *respuesta correcta* y los *distractores*. Son las más utilizadas en pruebas de rendimiento.

**Prueba referida a criterio:** Una prueba diseñada con especificaciones de contenido restringidas que sirven a un limitado rango de altos propósitos específicos. La prueba ayuda a determinar en qué medida los evaluados mantienen un cierto objetivo educacional.

**Prueba referida a normas:** Una prueba en la cual los puntajes son interpretados con respecto a las normas obtenidas de una muestra simple de examinados.

**Psicometría:** Teoría e investigación que comprende la medición de las características psicológicas (cognitivas y afectivas). Es un área temática de la Psicología que modela la medición de variables psicológicas y educacionales como el rendimiento a través de diversos modelos de escalamiento. Trata aspectos como la metodología de construcción de pruebas, el estudio de las propiedades de las pruebas y preguntas, y la aplicación e interpretación de sus resultados a través de modelos y técnicas estadísticas y matemáticas.

**Respuesta correcta mal especificada:** Es el caso de uno o más distractores de una pregunta de opción múltiple cuyo porcentaje (o porcentajes) de respuesta es mayor (son mayores) que el porcentaje de la respuesta correcta o dificultad de la pregunta. Es indicativo de un funcionamiento inadecuado de la pregunta, pues existen otras alternativas "correctas" distintas de la prevista. Esto sucede, por lo general, cuando el enunciado no es el adecuado.

**Respuesta dicotómica:** Es el tipo de respuesta que presenta dos alternativas a una pregunta; por ejemplo, "correcto" o "incorrecto", "verdadero" o "falso". Es una medida de incidencia donde 1 significa un acierto y 0 no.

**Validez:** Propiedad de una prueba que permite establecer si ésta mide lo que está designada a medir. La validez puede determinarse de varias formas: analizando el contenido de la prueba (validez de contenido), relacionando los puntajes con los de otro criterio (validez concurrente), o por un análisis del grado en que la prueba mide cierto constructo psicológico (validez de constructo).

**Validez concurrente:** El grado en que el puntaje obtenido está relacionado simultáneamente con determinado puntaje de otra prueba (criterio) de las mismas características.

**Validez de contenido:** Es el grado de validez de una prueba de acuerdo con la opinión de un grupo de expertos relacionado con la materia de la prueba.

**Validez de constructo:** El grado en que los puntajes de una prueba que mide cierta característica está relacionado con las medidas de conductas en situaciones en que la característica, se supone, determina de manera importante la conducta.

# EVALUACIÓN PSICOMÉTRICA DE LAS PRUEBAS CRECER 1998

JORGE BAZÁN  
CESCAR MILLONES

Esta sección tiene por finalidad evaluar las pruebas CRECER 1998 desde el punto de vista psicométrico, es decir, determinar si éstas son óptimas para ser usadas en el análisis del rendimiento escolar desde el enfoque de normas que posteriormente se precisa. Este documento técnico es complementario de otro<sup>1</sup> en el que se detallan las características psicométricas de las preguntas de las pruebas.

## 1 EL MARCO PARA EL ANÁLISIS DE LAS PRUEBAS

Es importante aclarar los conceptos del marco de análisis para evaluar psicométricamente las pruebas, que incluye los alcances y usos que pudieran hacerse de los puntajes obtenidos en éstas.

### PROPÓSITO DE LAS PRUEBAS

Una de las primeras preguntas que debe resolverse en el análisis de las pruebas es qué se espera de ellas; es decir, qué se pretende hacer con los resultados y qué tipo de conclusiones podremos obtener. Este aspecto ha sido recientemente señalado como clave para determinar una de las propiedades que debe cumplir la prueba: su validez<sup>2</sup>.

Existen dos enfoques en la interpretación de los resultados: el de normas y el de criterios. En el caso de CRECER 1998, el propósito principal

de las pruebas es reportar niveles relativos entre grupos de la población. Así, en cuanto al análisis de los resultados se adoptó un criterio de normas más que uno de criterios.

La interpretación de normas de una prueba busca estimar las distancias relativas entre grupos de interés o subpoblaciones. Estas comparaciones pueden definirse de varias formas. De esta manera, para algunos usos es interesante referirse a una norma nacional o regional. En otros casos se puede hacer referencia a subdistribuciones de grupos más específicos como podrían ser, por ejemplo, los de la gestión pública o privada. Se dice que las interpretaciones basadas en tales comparaciones son referenciadas con base en normas. Estadísticas adicionales como percentiles o cuartiles de la distribución relevante son útiles para especificar mejor las comparaciones entre estos grupos de interés.

En la interpretación de criterios no se hace referencia directa al desempeño de otros exa-

- 1 Bazán, J. y O. Millones: Evaluación psicométrica de las preguntas de las pruebas CRECER 1998. Lima: Ministerio de Educación-Unidad de Medición de la Calidad de la Educación, 2000. Documento de trabajo, incluido en la sección anterior de este volumen.
- 2 AERA, APA, NCME: *Standards for Educational and Psychological Testing*, preparado por un comité conjunto de la American Educational Research Association, American Psychological Association y el National Council on Measurement in Education. Washington: AERA, 1999.

minados. El interés se centra en determinar la probabilidad de éxito (individual) respecto de algún dominio de preguntas. Este tipo de interpretación también toma una variedad de formas. Por ejemplo, es posible referirse a la probabilidad de éxito o de respuesta correcta para un subconjunto de preguntas de la prueba (un dominio de la prueba) o para un dominio más amplio de preguntas. Otras interpretaciones pueden generarse si se consideran otros tipos de pruebas construidas para otros ámbitos fuera del área del rendimiento educativo.

En general, las pruebas —sea para su interpretación en normas o en criterios— deben cumplir con un conjunto de requisitos como validez y confiabilidad. Adicionalmente, están sujetas a principios en el desarrollo de pruebas, en las escalas generadas desde los puntajes originales y en las condiciones de su administración.

Respecto de la construcción de pruebas, éste es un proceso que incluye el planeamiento de la prueba, la selección de áreas a incluirse en ella, la proposición de un conjunto de preguntas que cubren las áreas elegidas, la administración de una prueba piloto para el ensayo de las preguntas seleccionadas, el proceso de análisis de las preguntas (lo que lleva a la selección de las mejores) y una administración final a partir de una muestra que servirá para la última versión de la prueba. El lector que quiera tener una síntesis de estos pasos puede revisar el anexo 2 de esta sección, que resume este proceso de construcción de pruebas.

## CRITERIOS PARA LA EVALUACIÓN DE LAS PRUEBAS: VALIDEZ DE LAS PRUEBAS

El análisis de las pruebas CRECER 1998 se basa en criterios presentados en la literatura psico-

métrica tanto desde el enfoque clásico<sup>3</sup> como desde el moderno<sup>4</sup>. Un aspecto del nuevo enfoque es el concerniente al significado ampliado que tiene el concepto de validez y el término asociado de “constructo”.

### Validez de la prueba

Desde el enfoque clásico, la validez de una prueba es la medida en que ésta mide el constructo que pretende medir. El término “constructo” se refiere a las características que no pueden ser medidas directamente sino que se infieren a partir de un conjunto de observaciones. El enfoque moderno del concepto de validez es más amplio: validez es el grado en que la evidencia acumulada (teórica o empírica) soporta las interpretaciones derivadas de los puntajes obtenidos en las pruebas<sup>5</sup>. Estas interpretaciones se refieren a los constructos o los conceptos que las pruebas se proponen medir (por ejemplo, rendimiento en Matemática). En este sentido, ya no se habla de diferentes tipos de validez (por ejemplo validez de contenido, concurrente o de constructo), sino de diferentes líneas o formas de evidenciar validez.

Este documento presenta un conjunto de criterios que proveen información relevante para determinar la validez de las pruebas. Los criterios que se describen incluyen: (i) el juicio de expertos; (ii) el análisis de unidimensionalidad de las preguntas que componen las pruebas; (iii) la confiabilidad de las pruebas; (iv) otras características basadas en las propiedades psicométricas de las preguntas, como son el nivel de dificultad, el grado de discriminación y los índices de no respuesta; y, (v) las propiedades derivadas de la construcción de las escalas y las transformaciones hechas para los objetivos de las pruebas. También puede derivarse evidencia complementaria de validez del diseño muestral y de la administración de las pruebas.

#### (i) Juicio de expertos

La opinión de los expertos tiene como finalidad analizar la correspondencia entre el contenido de las pruebas y los constructos que éstas intentan medir. El juicio de expertos se basa

3 Lord y Novick: *Statistical Theories of Mental Test Scores*. New York: Adisson-Wesley, 1974.

4 Por ejemplo, Moss P.A.: *Concepciones cambiantes de validez en la medición educativa: Implicaciones para la medición del desempeño*. Traducido por Juan Esquivel Alfaro. Tomado de *Review of Educational Research*, otoño de 1992, (62), 3, 1992, pp. 229-258; AERA, APA, NCME: *Standards for Educational...*, *op. cit.*, 1999.

5 AERA, APA, NCME: *ibidem*.

en el análisis curricular y las tablas de especificaciones que generaron los especialistas responsables de las pruebas (véase ejemplos de especificaciones en el anexo 1). Estas especificaciones fueron sometidas a juicio de expertos, y participaron en su elaboración los especialistas del MED y diversos consultores nacionales e internacionales.

### *(ii) Análisis de unidimensionalidad*

En el esquema moderno del concepto de validez se incluye la evidencia de unicidad, es decir, la propiedad de una prueba de medir únicamente un constructo (unicidad de la prueba medible).

Para establecer si el conjunto de preguntas de una prueba mide una sola cosa —es decir, para evaluar la unidimensionalidad—, se usó el Modelo de Análisis de Correspondencias<sup>6</sup>. Este análisis indica el grado de homogeneidad de los conceptos medidos por el conjunto de preguntas que componen la prueba. El criterio para determinar la unidimensionalidad es el porcentaje de varianza explicada por el conjunto de preguntas de la prueba. Si en la primera solución (para la primera dimensión) esta varianza explicada es de 70% o más, se concluye que esta dimensión es suficiente para explicar la varianza total; es decir, no es necesario considerar más dimensiones para explicar la varianza de la prueba.

### *(iii) Confiabilidad de la prueba*

La confiabilidad de una prueba mide el grado en que es consistente con los puntajes que de ella se obtienen. Idealmente se determina tomando dos o más veces la misma prueba a un examinado y revisando si los puntajes obtenidos son consistentes (idénticos o similares). En la práctica, la consistencia se determina de diversas formas alternativas, una de las cuales se basa en la consistencia interna de la prueba; es decir, por ejemplo, cuán consistente es la mitad de una prueba respecto de su otra mitad. Este criterio de consistencia interna de la prueba puede ser calculado por el coeficiente “alfa” de Cronbach<sup>7</sup>. Otra alter-

nativa para un indicador de confiabilidad son los coeficientes derivados de formas paralelas de la prueba<sup>8</sup>.

El criterio usado en CRECER 1998 es el de consistencia interna, es decir, coeficientes basados en la relación entre las preguntas de la prueba o entre un grupo de éstas. El supuesto del análisis de confiabilidad de las pruebas es que otros factores que afectan la distribución de puntajes son constantes o estables. Entre estos otros factores se encuentran las condiciones de administración de la prueba, la influencia del examinador y la inestabilidad de los estados del examinado, que son ajenas a los objetivos de la medición<sup>9</sup>.

### *(iv) Criterios basados en índices psicométricos de las preguntas*

Algunos criterios usados para las pruebas se sustentaron en los promedios de los índices psicométricos de sus preguntas. Los índices psicométricos de las preguntas incluyen la correlación pregunta-prueba, la discriminación, el nivel de dificultad y los índices de no respuesta. Así, para un índice agregado de correlación pregunta-prueba se ha tomado el promedio de las correlaciones pregunta-prueba de las preguntas que componen la prueba. Para un índice agregado de discriminación de las preguntas de las pruebas se ha tomado el promedio de los coeficientes de discriminación de las preguntas que componen la prueba. El nivel de dificultad de las pruebas se ha estimado con el promedio de los niveles de dificultad de cada pregunta. El nivel de no respuesta es el promedio de no

6 Véase Nishisato, S.: *Dual Scaling*. Toronto: University of Toronto Press, 1994. Para el análisis de homogeneidad, Visauta, V.B.: *Análisis estadístico con SPSS para Windows*, vol. II: *Estadística multivariante*. Madrid: McGraw-Hill, 1998.

7 Véase el anexo 2 (Glosario de términos psicométricos) de la sección anterior, p. 164.

8 Véase, por ejemplo, Nunnally, J. e I. Berstein: *Teoría psicométrica*. México: McGraw-Hill, 1995.

9 Véase Bazán y Millones: *Evaluación psicométrica...*, *op. cit.*

respuesta de las preguntas que componen la prueba<sup>10</sup>.

(v) *Construcción de escalas, transformaciones y normalidad. Comparabilidad de puntajes*

Típicamente, los puntajes son las sumas de las respuestas correctas de la prueba. Así, los puntajes altos denotan mayor rendimiento en la prueba. Sin embargo, es necesario aclarar que los puntajes están determinados en parte por el número de preguntas, el tiempo que dura la prueba y las dificultades que presentan las preguntas. Estas características hacen que diferentes puntajes sean a veces difíciles de interpretar en ausencia de mayor información. En el caso de las pruebas CRECER, el puntaje está definido por el número de aciertos.

*Construcción de escalas y transformaciones.* La interpretación de los puntajes y su análisis estadístico pueden facilitarse convirtiendo los puntajes en un conjunto diferente de valores llamados puntajes derivados o puntajes de escala. La literatura presenta diversas escalas; las más populares son la escala de Rasch y la escala porcentual. La UMC emplea ambas.

La *escala de Rasch*<sup>11</sup> requiere el cumplimiento de ciertos principios en aspectos como: a) el comportamiento de los alumnos durante las pruebas; b) las características de las pruebas; y, c) la aplicación misma. Se puede comentar estos aspectos en las pruebas CRECER 1998.

Respecto de lo primero, se ha estimado que los alumnos que sabían las respuestas tuvieron más oportunidad de responder correctamente que los que no las sabían. Igualmente, no hubo evidencia de que los alumnos no resolvieran la prueba de manera independiente.

Con respecto a las características de las pruebas, se verificó que éstas evaluaron el rendimiento del alumno de manera unidimensional. Esto sugiere que una sola habilidad sería suficiente para explicar la ejecución de la prueba. Adicionalmente, por construcción, las preguntas miden sólo una variable (evidencia recogida de los índices de correlación pregunta-prueba). También se sostiene que la respuesta a una pregunta no es afectada por las respuestas a otras, resultado que incluye el caso de las pruebas de Lenguaje que corresponden a un mismo estímulo, sea éste un texto o una imagen.

En lo que concierne a la aplicación de las pruebas, los tiempos asignados para su resolución fueron suficientes. Una evidencia de esto son las bajas tasas de no respuesta encontradas. Sólo en los casos de las pruebas de Matemática para cuarto y quinto de secundaria la no respuesta es más alta, aun cuando no significativamente (véase los indicadores del acápite 3: Resultados de los indicadores en la evaluación de las pruebas).

El modelo de Rasch postula que la relación entre el rendimiento y la dificultad de una pregunta sigue una función determinada que permite obtener la probabilidad de acertar una pregunta determinada para un rendimiento específico. La escala de Rasch es una estimación de las habilidades de los alumnos a partir del modelo de Rasch. Sin embargo, es importante anotar que en las pruebas CRECER 1998 sólo se ha empleado la escala como una transformación no lineal de la escala porcentual. No se han usado las otras características e información generada por este modelo. La transformación realizada toma valores de 50 a 550, y corresponde a una transformación lineal estandarizada de la escala *logit* del modelo de acuerdo con ciertas ponderaciones. Tiene media 300 y varianza 50, y la correlación entre esta escala transformada y la escala porcentual está por encima de 0,98 en todas las pruebas.

La *escala porcentual* corresponde al porcentaje de acierto de la prueba. Esta escala va de 0 a 100. Con ésta se consigue uniformar la presentación de los resultados, independientemente del número de preguntas de las pruebas.

Se han señalado limitaciones en el uso de escalas porcentuales para la presentación

10 Véase, por ejemplo, M. Martin y D. Kelly, editores: *Technical Report*, vol. III: *Implementation and Analysis*. Final Test of Secondary School (Population 3). Third International Mathematics and Science Study, 1998.

11 Véase Muñoz, J.: *Teoría de respuesta a los ítemes: Un nuevo enfoque en la evolución psicológica y educativa*. Madrid: Ediciones Pirámide S.A., 1990.

de los resultados<sup>12</sup>; entre las principales están: 1) no existe, *a priori*, ningún valor que pueda considerarse como rendimiento insatisfactorio; 2) no indican qué es lo que saben o lo que ignoran los alumnos; 3) no tienen en cuenta la dificultad de las preguntas; 4) no pueden referirse de ninguna manera a los contenidos; 5) no indican la importancia de las preguntas no contestadas correctamente, ni cuántos son los sujetos que no las contestaron; 6) no permiten hacer comparaciones entre pruebas distintas: por ejemplo, sería erróneo interpretar que el resultado de Matemática en tercero de secundaria (51% medio de aciertos) es inferior a los resultados del quinto grado de primaria (53% medio de aciertos).

Aunque las escalas no porcentuales como la de Rasch pueden superar estas limitaciones, la ventaja principal de tener una escala no porcentual es que los usuarios no intentarán determinar cuántos alumnos han sido aprobados (que no es el sentido de la prueba), sino que buscarán saber qué grupos de alumnos han salido mejor que otros.

## Normalidad

Es deseable que las escalas sigan una distribución normal para el uso de la inferencia paramétrica y para la eventualidad de formar grupos de rendimiento utilizando las medias y desviaciones estándar.

Esta propiedad ha sido difícil de obtener con las distribuciones de las pruebas CRECER 1998.

Sin embargo, en términos estadísticos esta exigencia no es necesaria. Las características de los puntajes de las pruebas (el número de preguntas de la mayoría es 30) determinan que cualquiera de las escalas presente sólo 30 valores diferentes, pues éstas son transformaciones biyectivas de aquéllos. Además, el tamaño efectivo de las muestras (aproximadamente 17 000) hace que las escalas presenten un número grande de “empates” (valores repetidos). De esta manera, la escala porcentual no necesariamente sigue una distribución normal. Sin embargo, las escalas de Rasch, por construcción, siguen una distribución normal.

## Uso de las escalas para la comparabilidad

Se ha mencionado que uno de los objetivos de la evaluación basada en normas es sostener la comparabilidad entre grupos. Tanto la escala porcentual como la de Rasch garantizan las comparaciones de una misma prueba entre los estratos de la muestra. Adicionalmente, la escala de Rasch refuerza este objetivo cuando consigue que todas las pruebas presenten la misma media o valor central para la muestra nacional.

Los resultados de ambas escalas servirán para presentar los reportes globales (o nacional) y por los estratos de interés (gestión, región, departamento, etcétera). La presentación de los resultados por estratos puede incluir los reportes de los promedios, errores estándares, cuartiles de distribución y porcentaje de alumnos dentro del estrato.

Los resultados deben ser presentados para cada curso y para cada grado, para evitar la crítica 6) ya expuesta. Se deben considerar además los factores de ponderación, que toman en cuenta qué porcentaje tiene el estrato elegido en la población y en la muestra.

## Homogeneización de pruebas

La homogeneización de las pruebas es un conjunto de procedimientos que hacen posible convertir el sistema de puntajes de una prueba al sistema de puntajes de otra prueba para que la interpretación sea equivalente. Para que esto suceda, la prueba debe mapear y mostrar que ambas miden la misma variable.

Dos pruebas son equivalentes si el desempeño o los puntajes pueden ser directamente trasladados de una a otra para que la elección de la prueba sea independiente del desempeño. Si la misma variable subyace en ambas pruebas, es indiferente cuál se use para

12 De la Orden Hoz, A.; R. Bisquerra; J. Gaviria; G. Gil; J. Jornet; F. López; J. Sánchez; M. Sánchez; J. Sierra y F. Tourón: *Los resultados escolares. Diagnóstico del sistema educativo 1997*. Madrid: Ministerio de Educación y Cultura, Instituto Nacional de Calidad y Evaluación, 1998.

obtener la medida de la posición de una persona en la variable. Conjuntos diferentes de ítemes pueden ser usados para evaluar grupos de personas diferentes.

En CRECER 1998 se utilizaron dos formatos de prueba de Matemática para el cuarto y quinto de secundaria. En cuarto, ambos formatos presentan tres preguntas iguales y dos similares (varían en el orden de las alternativas y, por ende, en la clave de respuesta). En quinto, ambos formatos presentan tres preguntas iguales y tres similares (varían en el orden de las alternativas y, por tanto, en la clave de respuesta). En ninguno de los dos grados coincide el lugar que ocupan los ítemes iguales o similares.

La utilización de dos formatos fue sugerida a raíz de las altas tasas de no respuesta detectadas en el piloto (prueba única de 40 preguntas y 90 minutos de duración). De cada objetivo o área de la prueba se tomó la mitad de los ítemes para determinar las formas A y B. Adicionalmente, se incorporaron cinco ítemes nuevos a cada prueba. Las pruebas fueron administradas a muestras equivalentes: la mitad del aula por cada escuela seleccionada del país. Cada muestra equivalente es para aproximadamente 8000 sujetos.

En principio, se trata de un diseño de anclaje con muestras equivalentes, algo que no debía ser difícil de equiparar. Un proceso de equiparación busca homogeneizar los puntajes (escalas) de las formas para elaborar un solo reporte de ambas.

Para realizar este procedimiento se tiene que partir del supuesto de que las poblaciones o que las pruebas son equivalentes. Es importante verificar ambos y decidir cuál de ellos será la base para la construcción de la escala (a partir de las capacidades de los alumnos o la dificultad de los ítemes).

Lo que se hizo fue un análisis de Rasch de los datos para comparar los indicadores de los ítemes en común. En principio, los indicadores deberían arrojar resultados muy similares, lo que sugeriría que, como se planificó al construir las pruebas, la escala de habilidades subyacentes es común. Se encontró que las dificultades de los ítemes de anclaje no son las mismas en las dos pruebas. Los ítemes que aparecieron primero en cualquiera de las dos formas fueron menos difíciles. El efecto de orden parece importante.

Tomando esto en cuenta, hemos asumido que las dos formas de pruebas tienen ítemes diferentes, pues las dificultades de los ítemes de anclaje así lo indican. Para certificar que es posible sobreponer las estimaciones RASCAL (desde el enfoque de Rasch) basadas en las muestras equivalentes, nos hemos preguntado si las pruebas son realmente diferentes en la medición del constructo. Hemos encontrado que no: las pruebas tienen idéntica distribución. Esto nos permite suponer que las formas midieron lo mismo y son "idénticas" en distribución de habilidades: 2 *logits* de habilidad en una "significa" 2 *logits* en la otra; es decir, las pruebas son equiparables.

## DISEÑO MUESTRAL

Un requerimiento importante para el efectivo uso de las pruebas es obtener muestras y tamaños de muestras representativos y apropiados. Las pruebas de 1998 se aplicaron a una muestra representativa de centros educativos (CE) polidocentes urbanos en el ámbito nacional. En el nivel de educación primaria incluye a 17 370 estudiantes del cuarto grado y a igual número de alumnos de sexto, y en el nivel de educación secundaria a 17 400 estudiantes del cuarto grado y a igual cantidad de alumnos de quinto.

## Estratos considerados

El enfoque de normas seguido sugirió la selección de un conjunto de criterios de utilidad para las comparaciones y para los futuros usos de los resultados de las pruebas. Estos criterios sirvieron de base para la estratificación de la población y de la muestra. Los estratos de la muestra sólo se refieren a la zona urbana, y en este ámbito los criterios de clasificación fueron: gestión del CE, región y departamento a los que pertenecen los examinados.

De acuerdo con el criterio de regiones se definieron tres regiones longitudinales (costa, sierra y selva) y tres transversales (norte, centro y sur), que determinaron las nueve regiones de comparación: costa norte, centro y sur; sierra norte, centro y sur; y selva norte, centro

y sur. Los departamentos fueron veinticinco en total (incluyendo Callao), y la gestión se refirió a escuela pública (estatal) y escuela privada (particular).

## Selección de la muestra

La muestra se seleccionó con base en un sistema de muestreo bietápico, estratificado y por conglomerados, con las escuelas como unidades de primera etapa y los estudiantes como unidades de segunda etapa. Es decir, en cada estrato se seleccionó una muestra probabilística de CE, y, luego, en cada CE se escogió, mediante sorteo aleatorio simple, a los estudiantes que integrarían la muestra.

El procedimiento fue el siguiente: dados los estratos considerados, en cada uno de ellos se determinó la cuota de colegios por elegir (dividiendo entre 30 la cantidad de estudiantes por seleccionar en el estrato y redondeando). Esta operación da más probabilidad de que una escuela sea seleccionada en función del número de alumnos que estudian en ella. Además, se fijó un mínimo de dos colegios por estrato de sorteo, para poder estimar el error de muestreo en todos los casos.

La segunda etapa consistió en la selección de estudiantes. Los 30 estudiantes de cada sección fueron elegidos por sorteo aleatorio simple. En el caso de CE con más de una sección se realizó un sorteo simple de secciones, para evitar complicaciones excesivas de logística. Esta etapa adicional del sorteo no sesga la muestra (sigue siendo representativa del colegio), pero tampoco permite retirar del error de estimación el posible efecto de diferencias entre secciones del colegio, si las hubiera. Pero es un riesgo calculado por el que se optó para controlar los errores "no de muestreo".

Aunque para cada nivel primario y secundario son dos las poblaciones objetivo (cuarto y sexto grado en primaria, por ejemplo), por tratarse de CE polidocentes urbanos (que tienen tanto cuarto como quinto grado) se optó por un diseño que contemplaba el muestreo simultáneo en ambas poblaciones; esto es, las muestras de cuarto y sexto se tomaron en los mismos CE. Las con-

sideraciones anteriores, aplicadas a los estratos, dieron una lista de 579 CE.

Se trata, pues, de una muestra con probabilidad de selección proporcional al tamaño en la primera etapa, pero que asegura una igual oportunidad de formar parte de la muestra a todos los estudiantes en la segunda etapa. Por esto, las estimaciones que se hagan requieren el uso de ponderaciones. Esto quiere decir que en una muestra planificada por estratos de muestreo (por departamentos, regiones y gestión) es necesario calcular los pesos o ponderaciones de los estudiantes para garantizar la representatividad total de la muestra y restituir la proporcionalidad<sup>13</sup>.

## Factores de ponderación

Para los efectos de los cálculos agregados y cálculos de los promedios de los puntajes, se usaron las ponderaciones correspondientes para corregir la no proporcionalidad respecto de los tamaños de los estratos del universo.

## Errores de estimación

La estimación de las varianzas y el error estándar para el cálculo de los promedios y otros estadísticos debe ser tomada en consideración para cualquier inferencia paramétrica. Debido al diseño de muestreo utilizado, las estimaciones del error estándar pueden ser empleadas usando las fórmulas correspondientes del muestreo estratificado o, alternativamente, pueden ser estimadas por otros métodos como Jackknife o Bootstrap. Es decir, no es recomendable el uso de las fórmulas de las varianzas del muestreo simple aleatorio<sup>14</sup>.

13 Para mayores detalles, véase Calderón, Arturo; Cholly Farro y Jorge Bazán: "Diseño muestral en la aplicación CRECER 98". Lima: Ministerio de Educación-Unidad de Medición de la Calidad de la Educación, 2000. Documento de trabajo.

14 Para mayor información sobre la estimación de dichos errores véase la sección Estimación del error estándar en las pruebas CRECER 1998 en este mismo documento.

## APLICACIÓN DE LAS PRUEBAS

La versión final de las pruebas se aplicó en diciembre de 1998. La aplicación fue supervisada por especialistas y coordinadores en cada uno de los CE seleccionados. Ningún profesor del CE donde se realizaba la evaluación participó en la aplicación de las pruebas: la tarea recayó sobre profesores de otros ámbitos especialmente entrenados para esta aplicación.

Las pruebas fueron elaboradas en los meses previos y tomadas al final de un proceso metodológico que se inició en 1997 a través de la aplicación piloto de cinco a diez formas de prueba para cada una de las doce pruebas nacionales. El anexo 2 resume el proceso de construcción de las pruebas.

## OTROS ENFOQUES ALTERNATIVOS

Las propiedades psicométricas mencionadas están relacionadas con el marco referencial teórico (modelo) que se adopta en el análisis de las pruebas. Como se sabe, los modelos más usados son: (i) los de la Teoría Clásica de los Tests; y, (ii) los modelos de la familia de la Teoría de Respuestas a Ítemes o TRI. El modelo de Rasch ya mencionado es un caso especial de la familia de los TRI.

### La Teoría Clásica de los Tests

La Teoría Clásica de los Tests es un enfoque según el cual el resultado de la medición de una variable depende de la prueba utilizada y de los sujetos evaluados. El énfasis que pone esta teoría en las pruebas usadas ha motivado críticas, pues en esta estrategia una variable es inseparable del instrumento utilizado para medirla. Esto constituye una seria limitación, pues inevitablemente se acabaría definiendo de manera operativa la variable por el instrumento con que se mide.

La Teoría Clásica de los Tests, denominada también Teoría del Puntaje Verdadero, se

apoya en un modelo lineal con error de medición formulado por Spearman en 1904. El puntaje obtenido en la prueba tiene dos componentes: su verdadero valor y un error de medición. A partir de una axiomática simple y con base en la noción de pruebas paralelas se definen las propiedades mencionadas: confiabilidad, validez y discriminación.

Las propiedades de las preguntas que se incluyen en las pruebas son las definidas por el modelo clásico: validez, dificultad, índice de discriminación e índice de no respuesta<sup>15</sup>.

### TRI versus teoría clásica

Otras opciones metodológicas respecto del análisis de las pruebas podrían ser seguidas con el uso de la TRI. La ventaja de tomar en cuenta otros enfoques es la oportunidad de estimar mediciones psicológicas adicionales que no pueden ser proporcionadas por la teoría clásica. Es importante anotar, sin embargo, que el enfoque TRI no contradice ni los supuestos ni las conclusiones fundamentales de la teoría clásica. Son sólo enfoques que nos dan información adicional sobre si la metodología empleada y los requisitos adicionales se cumplen. Por ello, el carácter de estos modelos TRI es complementario al de la teoría clásica.

Siguiendo una tendencia reciente en reportes especializados sobre evaluaciones de sistemas educativos nacionales e internacionales (NAEP-National Assessment Educational Progress, IAEP-International Assessment Educational Progress, TIMMS-Third International Mathematics and Science Study y LLECE-Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación), las pruebas nacionales CRECER 1998 fueron sometidas a un análisis con el modelo de medición de Rasch, uno de los modelos de Teoría de Respuesta al Ítem. Los modelos de Rasch son utilizados en sistemas de evaluación educativa de países como Australia, Inglaterra, Alemania, Estados Unidos, Colombia, Holanda y Dinamarca.

### Programas computacionales

Respecto del *software* computacional, en el análisis de las preguntas de las pruebas CRECER

15 Para mayores detalles, véase el anexo 2 de la sección anterior, p. 164.

1998 se usaron tanto el ITEMAN (modelo clásico) como el RASCAL con un parámetro (modelo de Rasch)<sup>16</sup>. Cada equipo de especialistas revisó las preguntas para verificar las propiedades generadas por las corridas.

Con el ITEMAN se obtuvo la información sobre nivel de discriminación, dificultad de ítems y aquella sobre distractores. Con el RASCAL se evaluó la posibilidad de extender el análisis para estimar las habilidades de los examinados. Este análisis está supeditado a la verificación del supuesto de unidimensionalidad que se desprende del análisis de las pruebas. Sin embargo, el RASCAL fue usado sólo para la parte de la transformación de las variables, que se ha explicado anteriormente.

## ALCANCES

El objetivo de CRECER 1998 de comparar estratos relevantes constituye un primer paso para la evaluación del rendimiento escolar en el sistema educativo peruano. Permite conocer la estructura del rendimiento educativo para los grados elegidos en 1998 e identificar sectores críticos relativos que necesiten mayor atención y prioridad en la implementación de políticas de mejoramiento de la educación. En este sentido, los resultados de las pruebas CRECER 1998 se colocan en una secuencia que provee de información para delinear los siguientes pasos en futuras aplicaciones.

Con las propiedades presentadas y las escalas elaboradas se puede obtener una idea de cómo (en términos relativos) han salido los estudiantes de diferentes grupos. El propósito de generar estos resultados es la comparación entre los grupos que corresponden a los criterios seleccionados y que son presentados en el tercer acápite de este documento.

Con respecto a las pruebas desarrolladas para la fase de estudio piloto y las versiones finales de las pruebas, es importante notar un aspecto cualitativo del muestreo. Los valores que se obtienen para las propiedades estadísticas analizadas pueden variar de muestra a muestra. A este tipo de variación hay que añadir la variación debida a la modificación (mejoramiento) de las preguntas. Por tanto, es necesario volver a revisar las propiedades psi-

cométricas por cambios de muestra y por modificaciones como resultado del desarrollo o construcción de *tests*. El poder inferencial de los resultados sobre la población muestreada está supeditado tanto a las posibles variaciones y modificaciones que ocurren como efecto del análisis de las pruebas en sí, cuanto a las muestras utilizadas entre la fase piloto y definitiva.

Este aspecto está vinculado a una de las críticas dirigidas a la Teoría Clásica de los Tests por los seguidores de la Teoría de Respuesta a Ítems (TRI), en el sentido de que los resultados de los *tests* no sólo dependen de los *tests* mismos sino también de los examinados. De ahí la necesidad, según el enfoque TRI, de obtener mediciones que no varíen en función del instrumento utilizado y la necesidad de disponer de instrumentos de medida cuyas propiedades no dependan de los examinados.

## Elección de la escala

Dado que las pruebas son usadas sobre todo para comparar los resultados entre grupos relevantes, cualquier transformación de los puntajes obtenidos a una nueva escala es apropiada para los fines de comparación.

En esta nueva escala lo que se busca es determinar qué grupo de alumnos ha salido mejor. Los resultados de la escala servirán para presentar los reportes globales (o nacionales), así como los resultados por los estratos de interés (gestión, región, departamento, etcétera). La presentación de los resultados por estratos puede incluir los reportes de los promedios (previa ponderación por no proporcionalidad en el tamaño de los estratos en la población), desviaciones estándar, cuartiles de distribución para cada materia estudiada y para cada grado por separado.

Para el análisis de las preguntas siguiendo el enfoque clásico, la revisión de las propiedades de las pruebas y de las preguntas de cada prueba se hará con la escala original de

16 Para mayor información, véase <http://www.assess.com/softmenu.html>

puntaje total. Para el análisis de los resultados, sin embargo, se empleará la escala de Rasch y la escala porcentual, dependiendo de los análisis por implementar.

## 2 CONSTRUCTOS Y CONTENIDOS DE LAS PRUEBAS

A fines de noviembre y los primeros días de diciembre de 1998 se realizó la Aplicación Nacional CRECER 1998 a los estudiantes de cuarto y sexto grado de primaria y de cuarto y quinto de secundaria. Esta aplicación de pruebas fue complementada con encuestas a padres de familia o tutores y directores de CE en los que estudiaban los evaluados.

La Aplicación Nacional incluye 18 pruebas (12 de selección múltiple y 6 de respuesta abierta o de desempeño), 2 a directores, 4 a padres o apoderados, 4 a estudiantes y 10 a profesores (que incluyen 4 denominadas de oportunidades para aprender).

### CONTENIDO DE LAS PRUEBAS

En 1998 los estudiantes del cuarto grado de primaria estaban desarrollando sus aprendizajes con la nueva estructura curricular organizada en áreas de desarrollo. Así, las pruebas fueron aplicadas en las áreas de Comunicación Integral, Lógico-Matemática, Perso-

nal Social y Ciencia y Ambiente. Los estudiantes del sexto grado de primaria, en cambio, desarrollaban la estructura curricular organizada en líneas de acción educativa, por lo que las pruebas fueron aplicadas en las asignaturas de Lenguaje, Matemática, Ciencias Histórico-Sociales y Ciencias Naturales. Los estudiantes de cuarto y quinto de secundaria desarrollaban también la estructura curricular organizada en asignaturas, por lo que las pruebas fueron aplicadas en Lenguaje y Matemática.

### LAS ÁREAS O ASPECTOS EVALUADOS

#### 4° grado de primaria

*Personal Social.* Incluye cuidado de la salud personal y colectiva, convivencia democrática, sentimiento de pertenencia y conocimiento de su medio sociohistórico y natural.

*Ciencia y Ambiente.* Cubre las áreas de conservación de la salud, conservación del medio ambiente y la intervención humana en el medio.

*Comunicación Integral.* Incluye las áreas de comunicación escrita-lectura, reflexión sobre la lengua y lectura de imágenes.

*Lógico-Matemática.* Revisa el conocimiento de los números y la numeración; la habilidad operativa y el cálculo; la medición y la organización del espacio.

Cuadro 1 Pruebas consideradas en CRECER 1998					
Primaria	N° de ítems	Tiempo de prueba	Secundaria	N° de ítems	Tiempo de prueba
<b>4° de primaria</b>			<b>4° de secundaria</b>		
Comunicación Integral	30	60	Lenguaje y Literatura	38	60
Lógico-Matemática	30	75	Matemática Forma 1	25	60
Ciencia y Ambiente	30	60	Matemática Forma 2	25	60
Personal Social	27	60			
<b>6° de primaria</b>			<b>5° de secundaria</b>		
Lenguaje	31	60	Lenguaje y Literatura	38	60
Matemática	30	75	Matemática Forma 1	25	60
Ciencias Naturales	30	60	Matemática Forma 2	25	60
Ciencias Histórico- Sociales					

## 6° grado de primaria

*Ciencias Histórico-Sociales.* Incluye historia y cultura peruana, forjadores de la peruanidad (personajes modelos de patriotismo), el Estado peruano, el reconocimiento del universo, el medio geográfico y la convivencia en sociedad.

*Ciencias Naturales.* Abarca las áreas de transformación de la materia, funciones del cuerpo humano, interacciones materia-energía y conservación del medio ambiente.

*Lenguaje.* Las áreas temáticas son comprensión de lectura, nociones gramaticales, vocabulario y análisis de imágenes (identificación de textos gráficos, sus usos, análisis e intencionalidad).

*Matemática.* Incluye números naturales, fracciones, números decimales, medición y geometría.

## 4° de secundaria

*Lenguaje y Literatura.* Presenta áreas de comprensión lectora, nociones y reglas gramaticales, análisis de imágenes y razonamiento verbal.

*Matemática.* Incluye Aritmética, Álgebra, Estadística y Geometría.

## 5° de secundaria

*Lenguaje y Literatura.* Abarca temas de comprensión lectora, nociones y reglas gramaticales, análisis de imágenes y razonamiento verbal.

*Matemática.* Comprende conjuntos, Aritmética, Álgebra, Estadística, Geometría y Trigonometría.

## REPRESENTATIVIDAD DE CONTENIDOS

Uno de los aspectos fundamentales de la validez de un test se refiere al grado y extensión con que los constructos son representados por los contenidos de la prueba. Al respecto, es importante determinar si el conjunto de contenidos desarrollados en una prueba representa o no los constructos teóricos por

medir. Puede manifestarse carencia de validez por una subrepresentación o representación irrelevante del constructo. En presencia de subrepresentación de constructos, los puntajes de las pruebas pierden la capacidad de representar el verdadero rendimiento del área evaluada.

Este riesgo de no representatividad se presenta en diversos grados de acuerdo con el objetivo del uso de la prueba, es decir, según el enfoque sea de normas o de criterios. Al construir una prueba con el enfoque de normas hay menos especificidad en los contenidos correspondientes al área a ser evaluada. No obstante, las áreas representadas en la prueba deben ser una muestra representativa del dominio del constructo.

La diferencia entre los análisis del enfoque de normas y de criterios es que el análisis de criterios requiere de una completa descripción de la(s) conducta(s) medida(s), mientras que en el análisis de normas hay más flexibilidad o generalidad para definir las conductas por medir. Esta diferencia puede ilustrarse con el siguiente ejemplo. Si consideramos el currículo escolar de cuarto grado de primaria en el área Lógico-Matemática, distinguimos dos niveles de análisis: (i) las competencias; y, (ii) las capacidades y actividades. En el primer nivel encontramos competencias como “conocimiento de los números y la numeración”, “conocimiento de las operaciones con números naturales”, etcétera. En el segundo nivel hallamos capacidades como “elaborar sucesiones numéricas crecientes y decrecientes”, o “establecer y graficar relaciones numéricas y no numéricas”.

El análisis de normas se dirige a preguntas representativas del primer nivel de las competencias (seleccionadas del total de competencias y de acuerdo con los criterios de la Teoría Clásica de los Tests). Aunque muchas de estas competencias corresponden a capacidades y actitudes del segundo nivel, lo que se desea es tener representatividad general del nivel de competencias.

En un análisis de criterios se seleccionan capacidades (del nivel ii) que son relevantes por sí mismas; por ejemplo, “clasificar números naturales de acuerdo con diversos criterios”. Sobre la base de una correcta especificación de esta capacidad (nivel ii) se analiza

un conjunto de preguntas y se incorpora un criterio de logro (por ejemplo, 75%).

Como se verá luego, el análisis de las pruebas CRECER 1998 sigue la orientación de normas y, por tanto, su uso se restringe a los objetivos específicos que persiguen este tipo de pruebas.

Este punto es importante cuando se considera otra diferencia entre los análisis de los enfoques de normas y criterios. En el enfoque de normas importan principalmente los contrastes relativos entre los examinados, mientras que en el enfoque de criterios se busca estimar específicamente lo que los examinados pueden o no hacer (los logros). Adicionalmente, el grado de rigor requerido en las pruebas basadas en normas para definir las conductas medidas puede ser más global o general, aunque representativo de los dominios del análisis.

La construcción de pruebas es un proceso que incluye el planeamiento de la prueba, la selección de áreas a incluir en la prueba, la proposición de un conjunto de preguntas que cubren las áreas elegidas, la administración de una prueba piloto para el ensayo de las preguntas seleccionadas, el proceso de análisis de las preguntas (lo que lleva a la selección de las mejores preguntas) y una administración final a partir de una muestra que servirá para la versión final de la prueba.

### 3 RESULTADOS DE LOS INDICADORES EN LA EVALUACIÓN DE LAS PRUEBAS

Con respecto a los indicadores considerados, los siguientes fueron los resultados cuantitativos más relevantes de las pruebas CRECER 1998 (véase los cuadros 2, 3 y 4).

#### CONFIABILIDAD (CONSISTENCIA INTERNA)

En lo que a las pruebas se refiere, los resultados se presentan a partir del indicador de consistencia interna (alfa de Cronbach).

Tanto las pruebas de primaria como las de secundaria arrojaron altos índices de consistencia interna. El rango se sitúa entre 0,73 en Matemática Forma 1 en cuarto de secundaria y 0,85 en Lenguaje de sexto grado de primaria. Los coeficientes de consistencia interna resultaron relativamente menores en el nivel secundario respecto del primario, aunque esta diferencia no es significativa.

Es de notar también que en el área de Matemática de secundaria se incluyen temas como Aritmética, Álgebra, Geometría/Trigonometría y Estadística, mientras que en primaria los temas de Matemática son más homogé-

**Cuadro 2**  
**Criterios de validez cuantitativa de las preguntas de las pruebas CRECER 1998:**  
**Educación primaria**

Estadísticas finales	4° grado				6° grado			
	Comuni- cación Integral	Lógico- Matemá- tica	Personal Social	Ciencia y Medio Ambiente	Lenguaje	Matemá- tica	Ciencias Histórico- Sociales	Ciencias Naturales
<b>Confiabilidad</b>								
Alfa de Cronbach	0,816	0,773	0,826	0,764	0,852	0,806	0,773	0,774
<b>Índices psicométricos</b>								
Índice de dificultad medio (%)	53,96	54,53	58,44	52,78	56,06	48,68	58,17	61,90
Índice de discriminación medio (%)	45,69	36,42	48,69	40,65	49,57	45,61	41,01	39,92
Índice de validez medio (%)	39,75	35,56	42,69	35,67	42,90	39,26	36,37	36,61
Índice de no respuesta medio (%)	2,75	3,58	1,14	1,17	3,69	4,93	1,15	0,78
Total de preguntas	30	30	27	30	31	30	30	30
Total de evaluados	16 997	16 827	16 819	16 744	16 833	16 716	16 759	16 650

<b>Cuadro 3</b> <b>Criterios de validez cuantitativa de las preguntas de las pruebas CRECER 1998:</b> <b>Educación secundaria</b>						
Estadísticas finales	4° grado			5° grado		
	Lenguaje y Literatura	Matemática Forma 1	Matemática Forma 2	Lenguaje y Literatura	Matemática Forma 1	Matemática Forma 2
<b>Confiabilidad</b>						
Alfa de Cronbach	0,774	0,733	0,755	0,79	0,772	0,755
<b>Indicadores psicométricos</b>						
Índice de dificultad medio (%)	44,12	46,44	45,65	54,94	47,18	46,27
Índice de discriminación medio (%)	37,26	39,44	43,20	39,78	40,87	44,09
Índice de validez medio (%)	32,32	36,82	38,21	33,97	38,90%	39,83
Índice de no respuesta medio (%)	6,58	8,73	7,65	5,22	8,80	10,82
Total de preguntas	38	25	25	38	25	25
Total de evaluados	16 939	8294	8274	16 710	8206	8034

<b>Cuadro 4</b> <b>Criterios de validez basados en la unidimensionalidad de las pruebas CRECER 1998:</b> <b>Educación primaria y secundaria</b>				
Pruebas	% varianza explicada (1ª dimensión)	Ratio de 1ª sobre 2ª dimensión	Ratio de 2ª sobre 3ª dimensión	Coefficiente de confiabilidad Theta
<b>4° de primaria</b>				
Lógico-Matemática	73,52	2,69	1,25	0,89
Comunicación Integral	81,85	3,80	1,11	0,87
Personal Social	84,71	3,89	1,28	0,84
Ciencia y Ambiente	77,33	3,23	1,17	0,89
<b>6° de primaria</b>				
Matemática	81,15	3,59	1,13	0,86
Lenguaje	86,45	4,42	1,08	0,84
Ciencias Histórico-Sociales	80,98	3,92	1,05	0,88
Ciencias Naturales	80,64	3,51	1,21	0,88
<b>4° de secundaria</b>				
Matemática Forma 1	70,23	2,54	1,23	0,89
Matemática Forma 2	74,77	2,89	1,25	0,88
Lenguaje y Literatura	76,73	3,26	1,10	0,91
<b>5° de secundaria</b>				
Matemática Forma 1	76,56	3,26	1,06	0,88
Matemática Forma 2	78,01	3,25	1,19	0,87
Lenguaje y Literatura	79,47	3,37	1,19	0,90

neos. Por otro lado, en Lenguaje de sexto de primaria las áreas temáticas son similares a las de secundaria.

## UNIDIMENSIONALIDAD

Todas las pruebas arrojaron un grado de unidimensionalidad (porcentaje de la varianza explicada de la primera dimensión) de 70% o más, lo que nos permite concluir que el conjunto de preguntas mide un solo aspecto. Las pruebas de Matemática arrojaron unidimensionalidad relativamente baja (entre 70 y 80%). Otras áreas como Lenguaje, Personal Social, Comunicación Integral, Ciencias Histórico-Sociales y Ciencias Naturales arrojaron una unidimensionalidad mayor de 80%.

Como en el índice anterior, es posible que las pruebas de Matemática midan áreas relativamente más heterogéneas que las otras pruebas. Sin embargo, como la consistencia interna de las pruebas es alta, se puede concluir que la principal fuente de variación no es el contenido de las pruebas sino la variación proveniente de los estudiantes. En el primer acápite se comentó, además, que las variaciones provenientes de otros factores situacionales del muestreo no fueron importantes. Un paso siguiente en el análisis de unidimensionalidad es aplicar el esquema de la Teoría de Respuesta al Ítem en el sentido de analizar la varianza proveniente de los sujetos (habilidades y actitudes).

## OTROS ANÁLISIS BASADOS EN LAS PREGUNTAS

Los siguientes datos en el nivel de prueba se basan en los resultados del análisis de las preguntas<sup>17</sup>; esto quiere decir que los resultados de dificultad media, discriminación media, correlación pregunta-prueba media e índice

de respuestas medio son los promedios de los indicadores respectivos de las preguntas que componen cada prueba.

## Correlación pregunta-prueba

Los coeficientes de validez se encontraron en el rango 32,32%-42,90%, y corresponden a las pruebas de Lenguaje y Literatura de cuarto de secundaria y Lenguaje de sexto de primaria respectivamente. El rango está sobre el nivel del 20%, que califica una buena validez.

## Dificultad<sup>18</sup>

El rango de dificultad medio varió desde 44,12% para Lenguaje y Literatura de cuarto grado de secundaria a 61,90% para Ciencias Naturales de sexto de primaria. Este rango no se aparta mucho del nivel de dificultad promedio recomendable de 50%. Merece notarse que las pruebas de primaria, salvo la de Matemática de sexto grado, estuvieron por encima de una dificultad del 50%. Por otro lado, y contrariamente, las pruebas de secundaria, excepto la de Lenguaje y Literatura de quinto de secundaria, estuvieron por debajo del valor medio de dificultad.

## Discriminación<sup>19</sup>

La prueba que más discriminó fue la de Lenguaje de sexto de primaria, con un coeficiente de discriminación medio de 49,57%, mientras que la que menos discriminó fue la prueba de Lógico-Matemática de cuarto de primaria. El promedio general de discriminación fue de 36,42%.

## No respuesta<sup>20</sup>

El rango de no respuesta se encuentra entre 0,78% para la prueba de Ciencias Naturales de sexto de primaria y 10,82% para la prueba de Matemática Forma 2 de quinto de secundaria. Sin embargo, la mediana para primaria es de sólo 1,96%, y para secundaria de 7,11%, lo que indica que la no respuesta fue un pro-

17 Véase la sección precedente, Evaluación psicométrica de las preguntas de las pruebas CRECER 1998.

18 Véase el anexo 2 (glosario) de la sección anterior, p. 164.

19 *Idem* nota 18.

20 *Idem* nota 18.

blema sólo en las pruebas de secundaria. Se observó que la Forma 2 de Matemática de quinto de secundaria, junto con las pruebas de Matemática de cuarto grado, arrojaron tasas de no respuesta relativamente altas comparadas con el resto de pruebas. Estos resultados nos llevan a concluir que son las pruebas de Matemática de secundaria las que arrojaron las más altas tasas de no respuesta.

## **4 CONCLUSIONES Y SUGERENCIAS**

### **CONCLUSIONES**

Respecto de la validez de las pruebas, y dado que el objetivo trazado es comparar grupos de interés, los resultados obtenidos del análisis de las preguntas a través de una variedad de tipos de evidencias (que van desde aspectos cualitativos como la opinión de expertos sobre los contenidos de las pruebas hasta indicadores más cuantitativos en el análisis de las preguntas) sugieren que las pruebas finales presentan propiedades psicométricas óptimas para su empleo en el análisis de resultados.

En cuanto a las escalas y el uso de los puntajes, los resultados de las pruebas pueden ser reportados con el uso de los porcentajes y la transformación a la escala de Rasch, teniendo en cuenta las ventajas y desventajas de cada cual ya expresadas en este documento. Cuando se quiera estimar totales u otros estadísticos agregados debe usarse el sistema de ponderaciones para recuperar la proporcionalidad del universo.

La estimación de las varianzas y el error estándar para el cálculo de los promedios y otros estadísticos debe tomar en consideración el diseño de muestreo usado. Pueden utilizarse las fórmulas correspondientes del muestreo estratificado o, alternativamente, pueden ser estimados por otros métodos como el de Jackknife<sup>21</sup>. Dicho de otro modo, no es recomendable el uso de las fórmulas de las varianzas del muestreo simple aleatorio.

Otro aspecto en la inferencia es la distribución subyacente en la muestra. La ausencia de normalidad en las distribuciones de los porcentajes de la muestra tiene algunas explicaciones. En primer lugar, el rango posible de va-

lores para los puntajes es truncado, es decir, se limita a uno que va de 0 a aproximadamente 30 (que es el máximo puntaje que se puede obtener). Considerando que en cada prueba hay 16 000 alumnos y muchos empates, la distribución de esta frecuencia es limitada al rango mencionado. Es importante anotar que, en la práctica, muchas de las comparaciones se hacen en un nivel más agregado, como por ejemplo la escuela que toma los promedios en las aulas. En este caso se presentan dos efectos favorables hacia la normalidad: una disminución de casos (existen aproximadamente 580 valores correspondientes a las escuelas) y una suavización de la distribución de casos (los promedios se distribuyen más continuamente), lo que mejora la densidad de frecuencia del intervalo de la escala.

La consecuencia de esto es que el uso de estadísticas paramétricas en las pruebas debe tomarse con cuidado. Es recomendable utilizar métodos complementarios sustentados en el análisis de datos categóricos o cualitativos en los que la asociación y correlación entre variables es estimada a partir de supuestos más flexibles. En el caso de los análisis paramétricos se sugiere el análisis parcial referido a grupos pequeños de categorías relevantes o modelos integrados que estiman efectos parciales en forma escalonada (modelos jerárquicos de análisis).

En síntesis, de los cuadros 2, 3 y 4 podemos concluir que:

- Las pruebas son unidimensionales y presentan alta confiabilidad. Los índices de validez basados en las preguntas son aceptables, lo que permite concluir que en el nivel global de prueba éstas poseen buenas características desde el punto de vista psicométrico.
- Las características adicionales presentadas reflejan un conjunto de pruebas con buena discriminación, centradas en una dificultad intermedia que garantiza las comparaciones por estratos de la muestra. Adicionalmente, las tasas de no respuesta, con excepción de las pruebas de Mate-

21 Véase Calderón, Farro y Bazán: "Diseño muestral...", *op. cit.*

mática de secundaria, han resultado poco significativas.

- El número de preguntas de las pruebas CRECER 1998 resultó apropiado, pues se minimizó la tasa de no respuesta de la etapa piloto, especialmente en el área de Matemática (en primaria el problema ha desaparecido y en secundaria se ha minimizado). La estrategia de utilizar dos formas de las pruebas permite obtener mayor número de preguntas.

## SUGERENCIAS

En el aspecto metodológico fue útil recoger una serie de recomendaciones que nos servirán para futuras aplicaciones. Las principales sugerencias son:

- Procurar una correspondencia entre el enfoque de evaluación (normas y criterios) y la metodología de construcción de pruebas (Teoría Clásica de los Tests, Teoría de Respuesta al Ítem), de manera que esta última proporcione las condiciones para el uso final de la información.
- Realizar una supervisión continua de los equipos responsables de las pruebas para garantizar el seguimiento similar (uniforme) y correcto de la metodología de cons-

trucción de pruebas y así eliminar diferencias en los procedimientos seguidos.

- Realizar análisis cuantitativos y cualitativos, de manera que las decisiones a tomar durante el seguimiento de la metodología de construcción de pruebas sean las más adecuadas.
- Implementar una etapa de verificación de cambios entre la etapa piloto y la aplicación definitiva, de manera que se anticipen el comportamiento final de las pruebas en términos de validez, confiabilidad e indicadores psicométricos agregados.
- Incorporar en la etapa piloto el análisis de sesgos (validez transcultural) y estudios complementarios de diferenciabilidad de las preguntas y prueba.

Desde una perspectiva más amplia, las bondades que nos ofrecen las estimaciones de los indicadores del modelo clásico pueden ser complementadas por otros modelos como el de Teoría de Respuesta a los Ítems. Para ello será necesario evaluar también las preguntas desde el punto de vista de si los supuestos implícitos de estos modelos se cumplen o no. Sin embargo, dado que se ha planteado el enfoque de normas, que conduce a una metodología útil para las comparaciones de grupos, el uso de las propiedades que se desprenden de la Teoría Clásica de los Tests es suficiente para la evaluación de estas pruebas.

# ANEXO 1

## TABLAS DE ESPECIFICACIONES DE LAS PRUEBAS CRECER 1998

Prueba de Lógico Matemática de 4° grado de educación primaria					
Competencias		Contenidos	Ítems	Total	%
1. Conocimiento de los números y la numeración	101	Relación de orden en los números naturales	2	5	16,67
	102	Relación de orden con los números fraccionarios	2		
	103	Relación de orden con los números decimales	1		
2. Habilidad operativa y cálculo	201	Resolución de ejercicios	2	12	40,00
	202	Resolución de problemas utilizando las cuatro operaciones fundamentales con los números naturales	4		
	203	Resolución de problemas utilizando las operaciones fundamentales con fracciones homogéneas	2		
	204	Conversión de fracciones a números decimales	2		
	205	Resolución de problemas utilizando las operaciones fundamentales con números decimales	2		
3. Medición	301	Estima longitud y masa en unidades convencionales y no convencionales	3	9	30,00
	302	Resuelve problemas de compra y venta relacionados con las unidades de masa	2		
	303	Resuelve problemas de compra y venta relacionados con las unidades de longitud	2		
	304	Resuelve problemas y ejercicios de compra y venta relacionados con las unidades de tiempo	2		
4. Geometría	401	Clasificación de ángulos y rectas paralelas y perpendiculares	2	4	13,33
	402	Identifica figuras geométricas en el plano	2		
			<b>Total</b>	<b>30</b>	<b>100,00</b>

Prueba de Comunicación Integral de 4° grado de educación primaria					
Competencias	Contenidos		Ítems	Total	%
1. Comunicación escrita: lectura	101	Identifica tipos de textos	2	13	43,30
	102	Reconoce información específica de un texto	5		
	103	Extrae la idea principal	2		
	104	Reconoce e infiere la causa y efecto de los sucesos	2		
	105	Encuentra el significado de palabras y frases por el contexto	2		
2. Reflexión sobre la lengua	201	Construye oraciones con sentido	2	13	43,30
	202	Identifica el sujeto y el predicado de la oración	2		
	203	Reconoce sustantivos y adjetivos	2		
	204	Identifica y usa los tiempos fundamentales de los verbos	2		
	205	Reconoce y utiliza los sinónimos y antónimos	2		
	206	Ordena alfabéticamente palabras en un contexto dado	3		
3. Lectura de imágenes	301	Identifica diversos textos gráficos	1	4	13,30
	302	Obtiene información a partir de ilustraciones	1		
	303	Analiza los elementos de las ilustraciones y gráficos	1		
	304	Descubre la intencionalidad del autor	1		
			<b>Total</b>	<b>30</b>	<b>100,00</b>

Prueba de Personal Social de 4° grado de educación primaria					
Competencias	Contenidos		Ítems	Total	%
1. Cuidado de su salud personal y colectiva	101	Riesgos para su salud física y comportamientos que previenen enfermedades	3	7	25,90
	102	Identificación y respeto de las normas de seguridad vial	2		
	103	Cumplimiento de las indicaciones de Defensa Civil	2		
2. Convivencia democrática	201	Comprensión de los deberes y derechos al interior de la escuela, en la familia y comunidad	4	6	22,20
	202	Respeto de las diferentes costumbres, opiniones y gustos	2		
3. Sentimiento de pertenencia	301	Identificación de los principales acontecimientos del proceso histórico peruano y ubicación cronológica	3	9	33,30
	302	Identificación de los símbolos patrios y sus características	2		
	303	Análisis de las funciones de las principales instituciones de la localidad	2		
	304	Reconocimiento de la importancia de la conservación del patrimonio cultural del Perú	2		
4. Conocimiento de su medio sociohistórico y natural	401	Ubicación de los países y departamentos limítrofes del Perú	2	5	18,50
	402	Reconocimiento de las actividades económicas de las diversas regiones del Perú	3		
			<b>Total</b>	<b>27</b>	<b>100,00</b>

Prueba de Ciencia y Ambiente de 4° grado de educación primaria					
Competencias	Contenidos		Ítems	Total	%
1. Conservación de su salud	101	Recuperación de energías mediante la alimentación	4	6	20,00
	102	Normas y códigos para conservar la salud	2		
2. Conservación del medio ambiente	201	Plantas y animales más representativos de diferentes medios	3	17	56,67
	202	Plantas y animales como recursos fundamentales para la supervivencia humana. Sus cuidados y uso racional	6		
	203	Contaminantes más frecuentes del aire, agua y suelo, y su influencia en los seres vivos	3		
	204	La importancia del agua para la vida y sus cambios de estado físico por acción del calor	5		
3. Intervención humana en el medio	301	Recursos naturales de la región, recursos renovables y no renovables	3	7	23,33
	302	Las transformaciones del paisaje y los recursos naturales producidas por los seres humanos para satisfacer sus necesidades	4		
			<b>Total</b>	<b>30</b>	<b>100,00</b>

Prueba de Matemática de 6° grado de educación primaria					
Objetivos	Contenidos		Ítems	Total	%
1. Números naturales	101	Relación de orden en los números naturales	2	8	26,67
	102	Ejercicios de multiplicación y división	2		
	103	Operaciones combinadas de adición, sustracción, multiplicación y división (hasta dos operaciones)	2		
	104	Resolución de problemas que requieren de dos operaciones fundamentales	1		
	105	Uso del algoritmo de mínimo común múltiplo	1		
2. Fracciones	201	Relación de orden y simplificación de fracciones	2	6	20,00
	202	Ejercicios de operaciones con fracciones	2		
	203	Resolución de problemas de la vida cotidiana utilizando operaciones con fracciones	2		
3. Números decimales	301	Relación de orden de decimales	1	4	13,33
	302	Operaciones combinadas de decimales	1		
	303	Resolución de problemas de la vida cotidiana utilizando operaciones con decimales	2		
4. Proporcionalidad	401	Resolución de problemas de la vida cotidiana utilizando regla de tres simple	1	2	6,67
	402	Resolución de problemas de la vida cotidiana utilizando porcentajes	1		
5. Medición	501	Resolución de problemas de la vida cotidiana utilizando unidades de medición	3	4	13,33
	502	Resolución de problemas de compra usando la conversión de unidades de medición	1		
6. Geometría	601	Clasificación de ángulos y rectas	2	6	20,00
	602	Cálculo y resolución de problemas con base en el perímetro de un cuadrilátero	2		
	603	Concepto de sólidos geométricos	1		
	604	Cálculo del volumen conjunto de cubos	1		
			<b>Total</b>	<b>30</b>	<b>100,00</b>

Prueba de Lenguaje de 6° grado de educación primaria					
Objetivos	Contenidos		Ítems	Total	%
1. Comprensión de lectura	101	Identificar los hechos o ideas principales de textos leídos	4	11	35,48
	102	Extraer información específica de un texto	4		
	103	Descubrir la intencionalidad del autor	2		
	104	Reconocer las marcas de cohesión textual	1		
2. Nociones gramaticales	201	Reconocer la oración gramatical como unidad de sentido	2	10	32,26
	202	Reconocer los núcleos del sujeto y del predicado	2		
	203	Reconocer las clases de palabras y sus accidentes	2		
	204	Conocer el empleo de los signos de puntuación	2		
	205	Usar correctamente las grafías: g-j, x, s-c-z, h, b-v, ll-y	2		
3. Vocabulario	301	Encontrar el significado de palabras por el contexto	2	6	19,35
	302	Identificar las familias de palabras	2		
	303	Reconocer las relaciones semánticas entre las palabras	2		
4. Análisis de imágenes	401	Identificar diversos textos gráficos	1	4	12,90
	402	Obtener información a partir de ilustraciones	1		
	403	Analizar los elementos de las ilustraciones	1		
	404	Descubrir la intencionalidad del autor	1		
			<b>Total</b>	<b>31</b>	<b>100,00</b>

Prueba de Ciencias Naturales de 6° grado de educación primaria					
Objetivos	Contenidos		Ítems	Total	%
1. Transformaciones de la materia; mezcla y combinación	101	Transformaciones de la materia. Cambios de estado	3	3	10
2. Funciones del cuerpo humano	201	Sistema digestivo: órganos, funciones, relaciones con otros sistemas	4	14	47
	202	Sistema respiratorio: órganos, relaciones: respiración, aire y fotosíntesis	2		
	203	Sistema circulatorio: órganos y funciones	2		
	204	Función excretora, sistema urinario	2		
	205	Función de relación: irritabilidad (sentidos)	2		
	206	Función de reproducción, órganos, paternidad responsable	2		
3. Interacciones entre materia y energía	301	Fotosíntesis, luz, propiedades	4	10	33
	302	El calor y climas. Buenos y malos conductores del calor	2		
	303	La energía eléctrica, fuentes y canales de transmisión	2		
	304	Imanes, propiedades. Magnetismo, sus aplicación en los aparatos inventados por el hombre	2		
4. Conservación del medio ambiente	401	Contaminación ambiental, problemas y alternativas de solución	3	3	10
			<b>Total</b>	<b>30</b>	<b>100</b>

Prueba de Ciencias Histórico-Sociales de 6° grado de educación primaria					
Objetivos		Contenidos	Ítems	Total	%
1. Historia y cultura peruana	101	Identificación del origen de la cultura peruana y el proceso histórico nacional	3	8	26,70
	102	Explicación de los factores ideológicos, políticos, económicos y sociales en la independencia	2		
	103	Análisis del proceso de formación de la conciencia nacional	2		
	104	Identificación de los símbolos patrios	1		
2. Forjadores	201	Comprensión de las acciones de los personajes modelos de patriotismo	3	3	10,00
3. Estado peruano	301	Análisis de los principales tratados limítrofes firmados por el Estado peruano	2	4	13,30
	302	Reconocimiento de la labor de los principales poderes del Estado	2		
4. Reconocimiento del universo	401	Reconocimiento de la estructura y características del sistema planetario solar	2	4	13,30
	402	Identificación de los elementos y movimiento de la tierra	1		
	403	Reconocimiento de las líneas imaginarias de la tierra y su utilización para la ubicación de puntos geográficos	1		
5. Medio geográfico	501	Ubicación de las regiones donde se encuentran los distintos recursos naturales y actividades económicas	4	6	20,00
	502	Análisis de las causas y efectos de la migración y el crecimiento urbano	2		
6. Convivencia en sociedad	601	Comprensión de los deberes y derechos de la persona al interior de la escuela, en la familia y la comunidad	3	5	16,70
	602	Reconocimiento de las normas de Defensa Civil	1		
	603	Comprensión de las principales normas de educación y seguridad vial	1		
			<b>Total</b>	<b>30</b>	<b>100,00</b>

Prueba de Lenguaje y Literatura de 4° grado de educación secundaria					
Objetivos	Contenidos		Ítemes	Total	%
1. Comprensión lectora	101	Extraer, deducir e interpretar información del texto	3	17	44,74
	102	Identificar y deducir características de los personajes	4		
	103	Identificar la intencionalidad del autor	2		
	104	Identificar la idea central y derivar conclusiones de un texto	1		
	105	Reconocer géneros, estructuras, técnicas y estilos literarios	7		
2. Nociones y reglas gramaticales	201	Usar los modos y tiempos verbales	1	7	18,42
	202	Establecer y aplicar reglas de concordancia	2		
	203	Identificar, clasificar y usar correctamente las categorías gramaticales	4		
3. Análisis de imágenes	301	Interpretar y analizar imágenes fijas, deduciendo su intencionalidad	4	4	10,53
4. Razonamiento verbal	401	Identificar el significado de palabras y expresiones según el contexto	2	10	26,32
	402	Relacionar palabras a través de analogías	2		
	403	Diferenciar la formación de palabras	2		
	404	Establecer relaciones semánticas entre palabras	2		
	405	Reconocer las familias de palabras	2		
			<b>Total</b>	<b>38</b>	<b>100,00</b>

Prueba de Lenguaje y Literatura de 5° grado de educación secundaria					
Objetivos	Contenidos		Ítemes	Total	%
1. Comprensión lectora	101	Extraer, deducir e interpretar información del texto	5	18	47,37
	102	Identificar y deducir características de los personajes	3		
	103	Identificar la intencionalidad del autor	2		
	104	Identificar la idea central y derivar conclusiones de un texto	5		
	105	Reconocer géneros, estructuras, técnicas y estilos literarios	3		
2. Nociones y reglas gramaticales	201	Usar los modos y tiempos verbales correctamente	2	8	21,05
	203	Reconocer las oraciones compuestas	2		
	203	Utilizar las preposiciones y adverbios correctamente	2		
	204	Reconocer la función sintáctica de la oración subordinada o proposición	2		
3. Análisis de imágenes	301	Interpretar imágenes fijas y reconocer el contexto en que se realizan	1	2	5,26
	302	Identificar la intencionalidad del autor o emisor	1		
4. Razonamiento verbal	401	Deducir el significado de palabras y expresiones según el contexto	2	10	26,32
	402	Relacionar palabras a través de analogías	2		
	403	Reconocer elementos comunes en la formación de palabras	2		
	404	Establecer relaciones semánticas entre palabras	2		
	405	Reconocer las familias de palabras	2		
			<b>Total</b>	<b>38</b>	<b>100,00</b>

Prueba de Matemática de 4° grado de educación secundaria F1 y F2					
Objetivos	Contenidos		Ítems	Total	%
1. Arimética	101	Resolver problemas de la vida real usando las operaciones fundamentales con números racionales (Q)	1		
	102	Resolver problemas de la vida real usando las operaciones fundamentales con números decimales	1		
	103	Resolver situaciones problemáticas usando las operaciones fundamentales con números reales (R)	1		
	104	Resolver problemas de la vida real que impliquen el uso de los conceptos de proporcionalidad directa e inversa	1		
	105	Resolver problemas de la vida real usando el sistema internacional de unidades y sistema monetario peruano	2		
	106	Resolver ejercicios de racionalización	1	7	28,00
	2. Álgebra	201	Factorizar polinomios	1	
202		Operar con fracciones algebraicas	1		
203		Resolver ecuaciones de segundo grado	1		
204		Resolver problemas de la vida cotidiana utilizando ecuaciones de primer grado	1		
205		Resolver situaciones problemáticas de la vida real utilizando sistemas de ecuaciones lineales con dos incógnitas	1		
206		Resolver ecuaciones lineales e inecuaciones de primer grado con o sin valor absoluto	2	7	28,00
3. Estadística		301	Interpretar gráficas y cuadros estadísticos	2	
	302	Resolver problemas de la vida real con medidas de tendencia central	1	3	12,00
4. Geometría	401	Resolver problemas utilizando conceptos geométricos y propiedades de ángulos y rectas	3		
	402	Resolver situaciones problemáticas aplicando propiedades de triángulos	1		
	403	Resolver problemas aplicando propiedades de la circunferencia	1		
	404	Resolver situaciones problemáticas que impliquen el uso de propiedades de proporcionalidad entre figuras geométricas	1		
	405	Resolver situaciones problemáticas que impliquen el cálculo de perímetros y áreas de cuadriláteros y hexágonos	1		
	406	Resolver situaciones problemáticas que impliquen el cálculo de áreas y volúmenes de prismas y cilindros	1	8	32,00
				<b>Total</b>	25

Prueba de Matemática de 5° grado de educación secundaria F1 y F2					
Objetivos		Contenidos	Ítems	Total	%
1. Conjuntos	101	Resolver situaciones problemáticas utilizando las operaciones fundamentales de conjuntos	1	1	4,00
2. Arimética	201	Desarrollar ejercicios sobre progresión aritmética y geométrica	1	6	24,00
	202	Resolver problemas de la vida cotidiana usando las operaciones fundamentales con números reales	1		
	203	Resolver situaciones problemáticas de la vida real utilizando las unidades del sistema internacional	2		
	204	Resolver situaciones problemáticas que impliquen la aplicación de conceptos y propiedades de proporcionalidad	1		
	205	Resolver problemas de la vida real usando regla de interés	1		
3. Álgebra	301	Realizar operaciones combinadas de adición, sustracción, multiplicación y división de polinomios	1	5	20,00
	302	Resolver problemas de la vida cotidiana con expresiones algebraicas	2		
	303	Resolver ecuaciones de segundo grado	1		
	304	Resolver problemas de la vida real utilizando ecuaciones y sistema de ecuaciones lineales con dos incógnitas	1		
4. Estadística	401	Resolver problemas de la vida real con medidas de tendencia central e interpretación de cuadros y gráficos	3	4	16,00
	402	Resolver problemas de la vida real utilizando propiedades fundamentales del análisis combinatorio	1		
5. Geometría	501	Resolver problemas sobre triángulos y cuadriláteros aplicando ángulos formados por dos paralelas y una secante	2	6	24,00
	502	Resolver problemas utilizando la relación de proporcionalidad entre figuras geométricas	2		
	503	Resolver situaciones problemáticas que impliquen calcular el área de un polígono cualquiera	1		
	504	Resolver problemas sobre áreas de polígono regular inscrito y circunscrito en una circunferencia	1		
6. Trigonometría	601	Resolver problemas sobre longitud de arco y sector circular usando el concepto y las propiedades de ángulo trigonométrico	1	3	12,00
	602	Resolver problemas de la vida real sobre ángulos de elevación y depresión usando propiedades de funciones trigonométricas	1		
	603	Resolver problemas que impliquen la reducción de ángulos al primer cuadrante y el uso de ángulos notables	1		
			<b>Total</b>	<b>25</b>	<b>100,00</b>

## ANEXO 2

### DESARROLLO Y CONSTRUCCIÓN DE LAS PRUEBAS

En la elaboración de las pruebas participaron prestigiados expertos nacionales e internacionales en aspectos de evaluación, sistemas nacionales de evaluación, política educativa y análisis de datos, quienes trabajaron en coordinación con especialistas de diferentes dependencias del MED. También se contó con la participación de profesores en diversos talleres de sensibilización y capacitación en los que se generó parte de las preguntas de las pruebas.

#### 1. SELECCIÓN DE LAS «REAS ELEGIDAS

##### ANÁLISIS CURRICULAR

La primera etapa consistió en revisar la estructura curricular (los programas curriculares) correspondiente a las áreas de desarrollo evaluadas. Este proceso derivó en la selección de aquellos aspectos que fueron contemplados en las pruebas. La tarea estuvo a cargo de especialistas de la UMC, en coordinación con los expertos en currículo de la Dirección de Educación Primaria y con profesores especialistas de cada una de las áreas de desarrollo evaluadas.

La selección curricular se expresa en los cuadros de especificaciones elaborados a manera de matrices para guiar la generación o redacción de las preguntas. De manera específica, en el segundo acápite y en el anexo 1 podemos ver los temas seleccionados para las pruebas en cada área escogida.

##### REDACCIÓN DE LAS PREGUNTAS DE LAS PRUEBAS

Una vez definidos los aspectos por evaluar, sintetizados en los cuadros de especificaciones, se prepararon pautas para la elaboración de las preguntas. En esta parte del proceso, profesores especialistas de diferentes lugares del país participaron en talleres de sensibilización. Se trata de una selección de maestros de aula de las 16 USE de Lima y de las ciudades de Huaraz, Tumbes, Iquitos, Huánuco, Pucallpa, Tacna, Puno, Ayacucho, Junín, Cajamarca, Madre de Dios y Piura.

Asistieron 432 profesores que fueron capacitados para la tarea encomendada y elaboraron preguntas que, después de ser revisadas, pasaron a formar parte del banco de preguntas para cada área o asignatura. Por su parte, además de revisar las preguntas generadas en los talleres, los diferentes equipos de la UMC elaboraron sus propias preguntas para someterlas a prueba durante la aplicación piloto.

En las 12 pruebas aplicadas las preguntas generadas son del tipo de selección de respuesta (o de opción múltiple). Este tipo de pregunta consta de un enunciado que interroga o plantea una situación problemática y de cuatro alternativas de respuesta, de las cuales sólo una es correcta. La calificación de las preguntas asigna 1 punto a la respuesta correcta y 0 puntos a la respuesta incorrecta o a la pregunta sin respuesta.

En cuanto a la redacción o formulación de este tipo de preguntas, se tomaron en consideración diversas pautas para formular el enunciado y para estructurar las alternativas. Asimismo, y gracias a la tabla de especificaciones, se pudieron redactar preguntas que exploraran distintos niveles del dominio cognoscitivo.

## **2. LAS PRUEBAS PILOTO**

Una vez construidas y revisadas las preguntas se elaboraron 10 formatos de prueba para cada área de desarrollo evaluada, con un promedio de 35 preguntas cada una. Estas pruebas preliminares se aplicaron en el mes de noviembre de 1997 a una muestra nacional de 3240 estudiantes (324 por formato) distribuidos en 156 escuelas del país.

## **3. EL ANÁLISIS DE LAS PREGUNTAS**

Las respuestas dadas por los sujetos a las pruebas piloto fueron sometidas a análisis estadísticos (análisis psicométricos) con el fin de conocer el nivel de dificultad del ítem, su poder discriminatorio (que posibilite la comparación entre grupos de estudiantes), el porcentaje de no respuesta, entre otros. Esto permitió eliminar parte importante de las preguntas de la prueba y mantener, por consiguiente, las más pertinentes. Los resultados de esta etapa se presentan en la sección Evaluación psicométrica de las preguntas de las pruebas CRECER 1998.

Las preguntas también fueron evaluadas por especialistas que vieron el grado de correspondencia de cada una con la tabla de especificaciones. Este proceso permitió seleccionar las preguntas más apropiadas y satisfacer los requerimientos de validez interna y de contenido de aquellas que pasaron a formar parte de la prueba final.

## **4. SELECCIÓN DE TEMAS DEFINITIVOS**

Las mejores preguntas (a partir de criterios presentados en el acápite Diseño muestral) sirvieron de base para la confección de las pruebas finales. Esta última versión cuenta con criterios óptimos de validez, pertinencia pedagógica y la confiabilidad del caso.

La versión final de las pruebas se aplicó en diciembre de 1998. La aplicación fue supervisada por especialistas y coordinadores en cada uno de los CE seleccionados. Ningún profesor del CE donde se realizaba la evaluación participó en la aplicación de las pruebas; la tarea recayó en profesores de otros ámbitos especialmente entrenados para esta aplicación. GRADE participó y apoyó logísticamente este proceso.

## **5. PROCESAMIENTO Y ANÁLISIS**

Una vez aplicadas todas las pruebas de rendimiento, los cuadernillos se remitieron a la UMC para su procesamiento computarizado. En vista de que los estudiantes marcaron sus respuestas en formatos de lectora óptica, éstos fueron leídos y luego ingresados automáticamente en la base de datos.

La información incluida se encuentra codificada de tal forma que posibilita el desarrollo de los análisis planificados. Las bases de datos han sido depuradas con el fin de identificar adecuadamente las respuestas a las pruebas.

La información ingresada a la base de datos permitirá generar múltiples reportes y comunicaciones oficiales acerca de los resultados de las pruebas CRECER 1998. Se trata de reportes diversos que pretenden ofrecer insumos relevantes para la toma de decisiones a los diferentes actores del entorno educacional y, asimismo, motivar la reflexión abierta sobre nuestras principales debilidades y fortalezas.