



Consejo Directivo Central
Administración Nacional de Educación Pública
Dirección Sectorial de Planificación Educativa
Dirección de Investigación, Evaluación y Estadística

EVALUACIÓN NACIONAL DE APRENDIZAJES
EN LENGUAJE Y MATEMÁTICA
6º AÑO ENSEÑANZA PRIMARIA - 2005

ANEP- PROYECTO MECAEP

Área de Evaluación de los Aprendizajes
Río Negro 1308
Tels.: 903 04 46 – 903 04 47
Email: invyeval@adinet.com.uy
Montevideo. – República Oriental del Uruguay
Marzo, 2007

Director Nacional de Educación Pública
DR. LUIS YARZÁBAL

Sub. Directora Nacional de Educación Pública
Q.F. MARISA GARCÍA

Consejeros del Consejo Directivo Central
PROF. LILIÁN D'ELÍA
MTRO. HÉCTOR FLORIT

Directora General de Educación Primaria
MAG. EDITH MORAES

Consejeras del Consejo de Educación Primaria
MTRA. MARÍA INÉS GIL
MTRO. OSCAR GÓMEZ

Dirección Sectorial de Planificación Educativa
MAG. PABLO MARTINIS

Dirección de Investigación, Evaluación y Estadística
DR. ANDRÉS PERI

ÁREA DE EVALUACIÓN DE LOS APRENDIZAJES

Redactores Responsables

Mtra. Graciela Loureiro
Prof. Ivanna Centanino
Mtro. Óscar Luaces
Mag. Carmen Haretche
Soc. Federico Rodríguez

Índice

Presentación		3
Capítulo 1	Un nuevo enfoque para la evaluación de los aprendizajes	5
	1.1 Las pruebas aplicadas en Uruguay entre 1996 y 2002	5
	1.2 La Teoría de Respuesta al Ítem	6
Capítulo 2	Las pruebas de evaluación de aprendizaje 2005	9
	2.1 El diseño de las pruebas de Lenguaje y de Matemática para la evaluación 2005	9
	2.2 Lo que no varía en la Evaluación de 2005	10
	2.3 Lo nuevo de la Evaluación de 2005	10
	2.4 La prueba de Lenguaje	12
	2.5 La prueba de Matemática	12
	2.6 Formato de presentación de resultados	12
	2.7 Descripción de la Prueba de Matemática	14
	2.8 Los niveles de desempeño en la prueba de Lenguaje y de Matemática	26
Capítulo 3	Resultados de los aprendizajes en Matemática y Lenguaje por Contexto sociocultural de las escuelas	35
	3.1 Principales características de la muestra	35
	3.2 Características sociales de la población evaluada	37
	3.3 Resultados académicos por contexto sociocultural	39
Capítulo 4	Comparación entre las muestras nacionales de 2005 y 2005	45
	4.1 Diseño muestral y de pruebas	45
	4.2 Cómo interpretar los resultados	46
	4.3 Comparación entre 2002 y 2005 del porcentaje de respuestas correctas a cada ítem del texto explicativo de la prueba de 2002	47
	4.4 Comparación entre 2002 y 2005 del porcentaje de respuestas correctas a cada ítem del texto narrativo de la prueba de 2002	50
	4.5 Comparación entre 2002 y 2005 del porcentaje de respuestas correctas a cada ítem de Matemática de la prueba de 2002	52
Anexo 1		57
Anexo 2		61

Presentación

Este informe representa la continuidad de una línea de evaluación del sistema educativo en la que el Uruguay ha incursionado desde hace ya una década. Cada tres años desde 1996 se ha realizado una evaluación de los aprendizajes de los alumnos de sexto año de Primaria para monitorear periódicamente la evolución del sistema educativo en su conjunto. En la evaluación de 2005, al igual que en las anteriores, el énfasis no ha sido puesto en su carácter de externo con respecto a las escuelas, sino en su carácter de evaluación del sistema para realimentar con información relevante al cuerpo docente y a las autoridades.

Esta última evaluación, realizada en octubre del 2005, introdujo un cambio importante en la metodología que, si bien trajo aparejada una demora considerable en la devolución de los resultados, va a permitir un mejor aprovechamiento pedagógico de las evaluaciones de aprendizaje. Por primera vez en Primaria se utilizó como metodología para el diseño de las pruebas la Teoría de Respuesta al Ítem (TRI) en lugar de la Teoría Clásica de los Test (TCT).

Este cambio aporta una nueva mirada a un debate impostergable: ¿cuál es el perfil de egreso que debe tener un estudiante al finalizar el ciclo primario?; ¿qué competencias y habilidades debe tener un estudiante para poder enfrentar la educación media con buenos fundamentos? Las tablas de desempeño construidas a través de la TRI describen lo que los estudiantes que cursan el último grado de la enseñanza primaria son capaces de realizar. Sería interesante discutir con los equipos docentes, utilizando estas tablas de desempeño, cuál es el nivel de egreso esperado.

La comparación de los niveles de desempeño por contexto sociocultural sigue mostrando marcadas diferencias entre los resultados de los alumnos de los contextos socioculturales muy favorables y los muy desfavorables. A pesar de los esfuerzos por garantizar equidad de oportunidades, las brechas de aprendizajes persisten. En el siguiente informe se analizará los resultados por distintas categorías de escuelas.

Esta comunicación de resultados muestra, por primera vez, un estancamiento relativo de los aprendizajes de los alumnos de sexto año en relación con resultados anteriores. Mientras que en todos los relevamientos previos el porcentaje de estudiantes en el nivel de suficiencia, tanto en Matemática como en Lenguaje, había crecido —sobre todo en los sectores sociales más vulnerables—, en el realizado en 2005 hay signos de que en ese año los desempeños se han estancado o han bajado levemente —aunque hay que tener en cuenta las diferencias en la longitud de las pruebas para que las comparaciones sean válidas. Si bien aún se está trabajando en la comparación de las pruebas, los primeros indicios muestran que no habría diferencias significativas en los aprendizajes entre 2002 y 2005.

Toda evaluación se justifica si contribuye al aprendizaje institucional. Los sistemas complejos necesitan de sistemas de monitoreo permanente para evaluar su marcha y corregir su rumbo. No alcanza con meras impresiones, sino que se necesita información técnica para tomar decisiones. Pero las evaluaciones serían fútiles si se quedaran en los archivos de los organismos especializados y no retroalimentaran a los distintos actores del sistema. El compromiso de la Dirección de Investigación, Evaluación y Estadística es apoyar a los distintos subsistemas de la ANEP para diseñar entre todos un Sistema Nacional de Evaluación que permita incorporar el aprendizaje institucional como forma de llegar al mejoramiento continuo de las instituciones educativas.

Montevideo, 12 de junio de 2006

Capítulo 1

Un nuevo enfoque para la evaluación de los aprendizajes

1.1 Las pruebas aplicadas en Uruguay entre 1996 y 2002

El Programa de Evaluación de Aprendizajes ha desarrollado desde 1996 distintos tipos de pruebas de aprendizaje. Todas ellas han tenido como finalidad principal el diagnóstico general de los aprendizajes de los alumnos, para apoyar el mejoramiento de la enseñanza, y la evaluación en las escuelas por parte de los maestros.

Algunas de estas pruebas fueron principalmente ejercicios de evaluación sin pretensiones de establecer comparaciones en el tiempo: la evaluación en 3º de 1998, en 4º en 2001 y en el Primer Nivel de la Escolaridad, también en 2001. En todos estos casos se trabajó con pruebas de respuesta abierta.

Paralelamente, las pruebas aplicadas en 6º grado en 1996, 1999 y 2002 tuvieron una finalidad adicional: la del seguimiento de los aprendizajes de los alumnos en el momento en que se encuentran próximos a egresar de la escuela primaria.

Estas pruebas fueron de elección múltiple y construidas en el marco de la denominada Teoría Clásica de los Tests (TCT), el enfoque disponible en nuestro país y predominante en la región en ese momento.

Para poder establecer comparaciones en el tiempo en el marco de la TCT es necesario trabajar con formas de prueba denominadas "paralelas" o "equivalentes". Esto significa que un conjunto de elementos debe permanecer igual para que eventuales diferencias en los resultados no sean atribuibles a cambios en el instrumento. Estos elementos son:

- los conocimientos y competencias evaluados;
- la cantidad total de ítems o preguntas de la prueba;
- la cantidad, extensión y tipo de textos, en el caso de las pruebas de Lenguaje;
- la distribución de los ítems entre los diferentes contenidos y competencias;
- la dificultad media del conjunto de los ítems que conforman la prueba como conjunto.

Cualquier cambio en alguno de los elementos anteriores puede producir variaciones en los resultados que no obedecerían a cambios en los conocimientos y competencias de los estudiantes.

Además, cualquiera sea el enfoque para la construcción de las pruebas, los procedimientos de aplicación y el momento del año escolar en que se aplican también deben permanecer constantes.

La TCT se caracteriza porque cada respuesta correcta vale un punto, independientemente de que la pregunta sea fácil o difícil. El puntaje de un alumno se calcula sumando el total de respuestas correctas. En las evaluaciones realizadas en 6º entre 1996 y 2002 se trabajó con pruebas equivalentes de 24 preguntas tanto para Lenguaje como para Matemática.

Las pruebas equivalentes fueron construidas a partir de un operativo piloto realizado en el año 1998. En este operativo fue posible construir dos pruebas equivalentes a la de 1996, las que fueron aplicadas en 1999 y 2002. La metodología seguida está explicada en los Informes de Resultados correspondientes a dichos años.

A diferencia de otros países de la región, que aplicaban pruebas de tipo "normativo" (que son pruebas que sirven principalmente para establecer comparaciones y ordenamientos de alumnos), en Uruguay se definió desde el comienzo un criterio de "suficiencia" en cada prueba, es decir, qué cantidad de preguntas se consideraba necesario que un alumno fuera capaz de responder, al finalizar el ciclo primario, como indicador de que logra un nivel de aprendizaje suficiente.

Este criterio fue fijado, a partir de la opinión de un grupo de especialistas y docentes que analizaron las pruebas, en 14 respuestas correctas sobre un total de 24. En todos los informes de resultados se empleó como dato principal el porcentaje de alumnos que alcanzaba este nivel. Entre 1996 y 2002 fue posible observar incrementos modestos pero significativos en dichos porcentajes, tanto en Lenguaje como en Matemática, con incrementos mayores en las escuelas de contextos sociales desfavorables.

1.2 La Teoría de Respuesta al Ítem

Desde hace mucho tiempo en el mundo desarrollado, y recientemente en la región, existe una metodología más compleja y avanzada para el desarrollo de pruebas estandarizadas, denominada Teoría de Respuesta al Ítem (TRI). Se usa en Brasil desde hace mucho tiempo y han comenzado a emplearla recientemente México, Chile, Argentina y Perú. Es la que se emplea en evaluaciones internacionales como PISA.

La TRI es un modelo matemático complejo para estimar la capacidad de los estudiantes en un determinado ámbito de conocimientos y competencias, que supera las limitaciones de la Teoría Clásica en los siguientes aspectos:

- a. En primer lugar, la puntuación de cada alumno tiene en cuenta la complejidad de las preguntas que respondió y no solamente la cantidad de respuestas que dio.
- b. Lo anterior significa que, a diferencia de la TCT, en la que cada respuesta vale un punto, en la TRI el "valor" de cada respuesta es distinto según la complejidad y dificultad de la pregunta.
- c. Como se trabaja con un modelo probabilístico, con la TRI no es necesario que todos los alumnos respondan exactamente las mismas preguntas. Basta con un subconjunto de éstas para estimar las puntuaciones en una escala común para todos los alumnos.
- d. Lo anterior hace posible, en primer lugar, trabajar con una cantidad mucho mayor de preguntas, porque los alumnos responden a distintos subconjuntos. Por ejemplo, en PISA se trabaja con alrededor de 200 preguntas, pero cada alumno responde cerca de 50.
- e. Esto hace posible describir los desempeños de los alumnos con mucho mayor detalle que lo que permiten las 24 preguntas empleadas hasta ahora en las pruebas de 6º en Uruguay, a través de la construcción de los denominados "niveles de desempeño", que constituyen una descripción progresiva de las capacidades de los alumnos en el área evaluada, desde los niveles más simples hasta los más complejos.
- f. De esta manera, la información que proveen las pruebas resulta mucho más rica desde el punto de vista pedagógico. Una de las debilidades de las pruebas empleadas en 6º hasta el momento es que no es posible describir con detalle el

significado de la "suficiencia", más allá del hecho de que el alumno respondió correctamente 14 preguntas.

- g. Una segunda consecuencia relevante de lo indicado en el numeral "c" es que con la TRI las comparaciones en el tiempo son más precisas, a condición de mantener un subconjunto de preguntas como confidenciales. Estas preguntas, denominadas "ítems de anclaje", se aplican en sucesivas evaluaciones y permiten "equiparar" las puntuaciones, es decir, colocarlas en la misma escala que en las evaluaciones anteriores a los efectos comparativos.
- h. Otra ventaja de la TRI es que permite incorporar nuevas preguntas en cada ciclo de evaluación y enriquecer en forma permanente el banco de ítems. Este aspecto es una limitación importante de la TCT, porque no permite ir aumentando el banco de ítems a la vez que se realizan las distintas evaluaciones. Como se señaló anteriormente, las evaluaciones de 1996, 1999 y 2002 se realizaron con formas paralelas diseñadas a partir de una evaluación (1998) cuyo objetivo no fue devolver resultados, sino crear las pruebas equivalentes.
- i. Finalmente, la TRI permite trabajar con preguntas de respuesta construida por el alumno (no es necesario limitarse a las preguntas de elección múltiple) y, además, permite considerar respuestas parcialmente correctas.

En el marco de la TRI las pruebas se diseñan a partir de distintas combinaciones de pequeños grupos de ítems, denominados bloques.

Un bloque es un conjunto de ítems que representa la matriz de contenidos que se quiere evaluar. Para seleccionar los ítems que componen cada bloque, además de su contenido se considera su dificultad, de manera tal que los bloques sean lo más homogéneos entre sí en cuanto a su dificultad promedio.

De las distintas combinaciones de estos bloques surgen los diferentes cuadernillos de prueba, de los cuales cada alumno contesta solamente uno. Todos los cuadernillos tienen la misma cantidad de bloques. Por último, como cada bloque aparece en más de un cuadernillo, es posible calibrar las respuestas de los alumnos que contestaron distintos cuadernillos en una misma escala.

Tabla 1.1
Ejemplo de un diseño de prueba con 5 bloques
distribuidos en 5 cuadernillos

	Bloque	Bloque	Bloque
Cuadernillo 1	A	B	D
Cuadernillo 2	C	E	A
Cuadernillo 3	B	D	C
Cuadernillo 4	E	A	B
Cuadernillo 5	D	C	E

En el momento de la evaluación los cuadernillos se van rotando entre los alumnos, lo cual garantiza la obtención de grupos equivalentes para los diferentes cuadernillos.

Para el ejemplo recién presentado, suponiendo que cada bloque se componga por 10 ítems, podríamos decir que cada alumno contesta 30 y que la prueba se compone por un total de 50.

Una vez realizada la evaluación, se selecciona uno o más bloques para volverlos a aplicar en una próxima evaluación, a los cuales se los denomina “de anclaje”. Es por lo tanto a través de los bloques de anclaje que es posible equiparar los resultados de dos evaluaciones diferentes en la misma escala.

Capítulo 2

Las pruebas de evaluación de aprendizaje 2005

2.1 El diseño de las pruebas de Lenguaje y de Matemática para la evaluación 2005

Para la evaluación de 2005 el Área de Evaluación de Aprendizajes enfrentó la siguiente disyuntiva:

- a) volver a construir pruebas equivalentes a las aplicadas en evaluaciones nacionales anteriores, en el marco de la Teoría Clásica, con las limitaciones antes anotadas, pero haciendo posible mantener la comparación de resultados con la tendencia 1996-2002;
- b) construir una nueva generación de pruebas, con el enfoque de la TRI, que permitiera:
 - aportar descripciones más detalladas sobre los desempeños de los alumnos,
 - trabajar con mayor cantidad de preguntas e incluir preguntas de respuesta construida,
 - proveer información que aporte a una discusión sobre el significado de la "suficiencia", a través de niveles de desempeño,
 - garantizar comparaciones más precisas hacia el futuro.

Se optó por esta segunda opción, al costo de perder la comparabilidad con las evaluaciones realizadas entre 1996 y 2002.

Sin embargo, para no perder completamente la comparabilidad con las evaluaciones anteriores, se agregó a las pruebas 2005 un conjunto de cuadernillos que incluyeron parte de las preguntas de la prueba de 2002. Los datos de los alumnos que respondieron a estos cuadernillos no forman parte del resultado de la "muestra nacional" de 2005, sino que se utilizan como una submuestra solamente a los efectos de comparar los porcentajes de respuestas correctas a cada uno de los ítems comunes entre las evaluaciones de 2002 y 2005. Estos resultados permiten una aproximación, al menos indicativa, a los cambios producidos en el último trienio.

Por lo tanto, en la prueba de 2005 hay dos tipos de situaciones. Por un lado, aquellos alumnos que se encuentran en la muestra nacional, que se considerará la línea de base para las futuras evaluaciones. Estos alumnos no respondieron preguntas de las evaluaciones anteriores. Por otro lado, aquellos que contestaron cuadernillos con ítems de la evaluación de 2002. Estos últimos permitirán comparar el porcentaje de respuestas correctas en 2002 y en 2005.

A los efectos del desarrollo de los nuevos instrumentos de evaluación, se conformó un equipo de docentes para redactar las nuevas actividades de Lenguaje y Matemática.

A través de la prensa se realizó un llamado a maestros aspirantes a ser redactores de actividades. Para participar se debió cumplir con los siguientes requisitos:

- ser maestro,
- tener cinco años de antigüedad docente como mínimo,
- experiencia en 5º y/o 6º en los últimos cinco años (maestros) / 3 años en la dirección de la escuela.

Se valoró además:

- otros títulos relacionados con las disciplinas a evaluar,
- calificación de la Inspección,
- haber participado anteriormente en la elaboración de actividades de pruebas de evaluación.

Los docentes preseleccionados debieron participar de un seminario de capacitación de tres días con jornadas de aproximadamente 10 horas de trabajo. Se trabajó en las disciplinas a evaluar (Matemática y Lenguaje) y en el diseño de instrumentos de evaluación. Hubo al final una instancia de evaluación del seminario, luego de la cual fueron seleccionados 15 docentes para Matemática y 15 para Lenguaje a los que se les encomendó el diseño de actividades de evaluación. Trabajaron con supervisión continua del Programa de Evaluación de Aprendizajes.

Finalmente, a partir de las actividades elaboradas por los maestros se diseñaron las pruebas atendiendo las Tablas de Especificaciones (grilla de conocimientos y competencias) de las evaluaciones anteriores. Igualmente se introdujeron algunas modificaciones tales como el trabajo con textos explicativos y discontinuos, además del narrativo en el caso de Lenguaje. En Matemática se incluyeron preguntas de respuesta abierta.

2.2 Lo que no varía en la Evaluación de 2005

En la construcción de este instrumento se han respetado los siguientes criterios:

- a) El instrumento fue construido manteniendo el mismo referente: tabla de especificaciones que fue acordada por la totalidad del cuerpo docente del país a partir de sucesivas consultas iniciadas en 1996.
- b) Este referente fue acordado en base a los siguientes principios generales:
 - se privilegia la evaluación de competencias sobre la de contenidos;
 - la selección de contenidos que permiten evaluar las competencias se efectuó a partir de su relevancia y con criterio acumulativo (contenidos curriculares seleccionados de los programas de 4º a 6º año, inclusión de contenidos emergentes);
 - las competencias seleccionadas para evaluar el área de Matemática son tres: Resolución de Problemas, Comprensión de Conceptos y Aplicación de Algoritmos. En Lenguaje se evalúa la Comprensión de Textos.
- c) Se mantiene el mismo criterio de jerarquización de las competencias en la prueba a través del porcentaje de ítems que evalúan cada una de éstas.

2.3 Lo nuevo de la Evaluación de 2005

- a) El modelo de análisis TRI, como se expresó anteriormente, es probabilístico y asigna un puntaje al ítem y al desempeño del alumno en la prueba. Este valor numérico es indicador de la dificultad del ítem y de la habilidad del sujeto en el área evaluada. Ambas categorías: dificultad de la actividad y habilidad del sujeto están en una misma métrica. Por ejemplo: si se tiene que el puntaje de un ítem es 500 puntos y el indicador de competencia matemática que obtuvo un alumno en la prueba es también de 500 puntos, este alumno tiene una probabilidad del 50% de resolver el ítem en forma correcta. Asimismo, este alumno va a tener mayor probabilidad de responder ítems con puntajes inferiores a 500 y probabilidad menor que 50% de responder correctamente ítems con puntajes mayores que 500 puntos.

- b) Las actividades seleccionadas cubren la totalidad de la matriz de especificaciones. Esto es posible debido a que los alumnos no responden a todas las preguntas que componen la prueba. Como se dijo, responden a un subgrupo de actividades presentadas en el cuadernillo de prueba. Estos subconjuntos de ítems se organizan en bloques. Un cuadernillo está compuesto por varios bloques. Cada uno de estos bloques se presenta en varios cuadernillos y en posiciones diferentes (al final, al medio y al comienzo de la prueba). Esta organización permite: i.- calibrar los ítems en una misma métrica, ii.- cubrir la totalidad de contenidos de la tabla de especificaciones sin necesidad de que un alumno responda todos y iii.- compensar el factor cansancio y reducir así el peso de la omisión de respuesta por ítem.
- c) Los bloques se construyen manteniendo el criterio de equivalencia según su nivel de dificultad. Todos poseen la misma dificultad promedio.
- d) Cada bloque reproduce los criterios de distribución de ítems por competencias. Por ejemplo, en Matemática cada bloque está integrado por 10 ítems de los cuales cinco corresponden a la competencia Resolución de Problemas, tres a Comprensión de Conceptos y dos a Aplicación de Algoritmos.
- e) Para comparar con los resultados de la evaluación en 6° año del 2002, en Matemática se seleccionó un bloque de 12 de los 24 ítems de aquella prueba. En Lengua se incluyeron todos los ítems de 2002 para comparar con los resultados de 2005. Estos bloques de ítems 2002 se combinaron con los nuevos bloques de ítems 2005.
- f) Algunos de los bloques nuevos de la prueba de 2005 se reservan confidenciales para equiparar con resultados en futuras evaluaciones. A estos bloques se los denomina “de anclaje”.

En el correr del año próximo se hará pública una prueba de Lenguaje y una de Matemática, con un nivel de dificultad equivalente al de los cuadernillos utilizados en la evaluación.

Estas pruebas podrán ser utilizadas en actividades de evaluación autónoma y en procesos de retroalimentación técnica.

Las escuelas y los docentes de sexto año que deseen aplicar, en forma autónoma, esta evaluación a sus alumnos recibirán: los formularios de pruebas para cada uno de los alumnos de su grupo y el manual de aplicación, corrección e interpretación de respuestas a la prueba. Estos manuales contarán con información relativa al porcentaje de respuestas correctas que obtuvieron los alumnos que participaron en la muestra nacional, diferenciados por contexto sociocultural. Esto último permitirá a los docentes analizar los resultados de su grupo comparándolo con los resultados de alumnos pertenecientes a un contexto sociocultural similar al de su escuela.

2.4 La prueba de Lenguaje

La prueba de Lenguaje aplicada en octubre de 2005 estuvo compuesta por seis cuadernillos con dos bloques de doce preguntas cada uno. Por lo tanto, al igual que en las evaluaciones anteriores cada alumno respondió 24 preguntas. Pero en total la prueba consta de 72 preguntas.

En el diseño de los cuadernillos se tuvo en cuenta el tipo de texto propuesto en cada bloque. En la tabla que sigue se presenta el diseño de la prueba de Lenguaje aplicada en octubre de 2005.

Tabla 2.1
Estructura de la prueba de Lenguaje - 6to 2005

Cuadernillo	1er bloque			2do bloque		
	Bloque	Tipo de texto*	Cantidad de textos	Bloque	Tipo de texto*	Cantidad de textos
I	A	E/N	2	F	E/D/D	3
II	B	E/N	2	E	N/D/D	3
III	C	N/D/D	3	D	E/D/D	3
IV	D	E/D/D	3	A	E/N	2
V	E	N/D/D	3	B	E/N	2
VI	F	E/D/D	3	C	N/D/D	3

* Tipos de textos: "E": explicativo, "N": narrativo y "D": discontinuo. Los textos discontinuos, a diferencia de los otros dos, son textos cortos y no se incluyeron en las evaluaciones anteriores.

2.5 La prueba de Matemática

Para la prueba de Matemática se conformaron ocho bloques de diez preguntas cada uno (una de ellas de respuesta abierta). Los bloques fueron combinados en ocho cuadernillos de tres bloques cada uno.

El cuadernillo de prueba de Matemática estuvo compuesto por 30 preguntas. La prueba incluyó un total de 80 preguntas.

Tabla 2.2
Estructura de la prueba de Matemática - 6to 2005

Cuadernillo	1er bloque	2do bloque	3er bloque
I	A	C	F
II	B	D	G
III	C	E	H
IV	D	F	A
V	E	G	B
VI	F	H	C
VII	G	A	D
VIII	H	B	E

2.6 Formato de presentación de resultados

En las evaluaciones de 1996, 1999 y 2002 los resultados se presentaron como "porcentaje de alumnos que alcanzan la suficiencia". Según se indicó anteriormente, para que el desempeño de un estudiante fuese considerado "suficiente" se estableció, a partir del juicio

experto de docentes y especialistas, el criterio de responder correctamente al 60% de las preguntas.

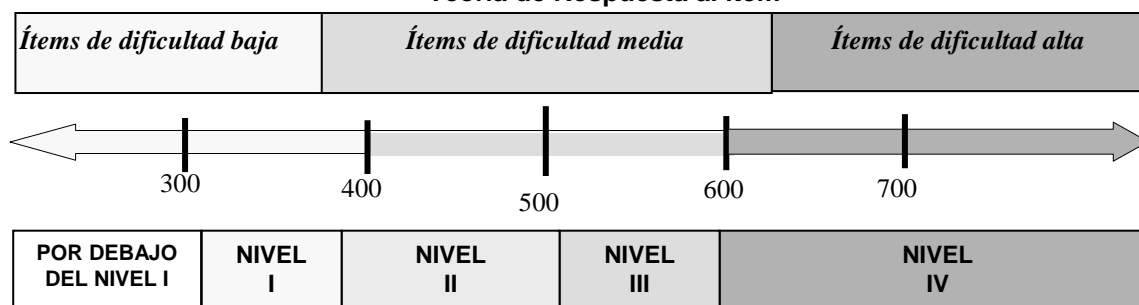
El cambio de metodología en la construcción de las pruebas implica modificaciones en la forma en que se presentan los resultados.

Mientras en el marco de la Teoría Clásica las puntuaciones de los alumnos están constituidas por la suma del total de respuestas correctas, en el marco de la TRI las puntuaciones se estiman a partir de un modelo matemático y se expresan en una escala que no tiene ni un cero absoluto ni un máximo. El punto de referencia de la escala no es la cantidad de preguntas, sino el desempeño promedio de la población evaluada.

Normalmente este promedio se fija arbitrariamente en 500 puntos, y se establece que una desviación estándar equivale a 100 puntos. Alrededor de dos tercios de los estudiantes tendrán puntajes ubicados entre 400 y 600 puntos. Un alumno cuyo puntaje es superior a 600 tuvo un muy buen nivel de desempeño en el área evaluada, un alumno con un puntaje de entre 400 y 600 puntos tuvo un desempeño de nivel medio, en tanto que un alumno con un resultado por debajo de los 400 puntos tuvo un desempeño pobre (véase la Figura 2.1).

Del mismo modo, la escala de puntuaciones se aplica a las preguntas de la prueba. El modelo TRI asigna a cada pregunta un puntaje en la misma escala que los desempeños de los estudiantes. Una pregunta con una puntuación de alrededor de 500 puntos es un ítem de dificultad media, lo cual significa que puede ser respondido por la mayoría de los alumnos de desempeño medio o superior. Una pregunta con una puntuación de más de 600 puntos es una pregunta difícil, que solo podrá ser respondida por alumnos de desempeño superior. Una pregunta con una puntuación de menos de 400 puntos es una pregunta fácil, que podrá ser respondida por la mayoría de los alumnos, incluso por aquellos de desempeño más pobre (véase la Figura 2.1).

Figura 2.1
La escala de puntuaciones en la
Teoría de Respuesta al Ítem



Una vez establecida la escala se ordenan todas las preguntas de la prueba de las más fáciles a las más difíciles. Un grupo de expertos analiza las preguntas y establece dos, tres o cuatro puntos de corte en la escala, que distinguen conjuntos de preguntas que reflejan diversos “niveles de desempeño” de los alumnos. Paralelamente se calculan los porcentajes de respuesta correcta de cada ítem y otras medidas estadísticas que permiten detectar los ítems que pertenecen y describen a cada nivel de desempeño. El paso siguiente consiste en describir, a partir de esos subconjuntos de preguntas, qué son capaces de hacer los alumnos que, de acuerdo a la puntuación que obtuvieron, se ubican en cada uno de los niveles de desempeño.

En este proceso cada ítem se aborda desde un punto de vista estadístico y pedagógico. El resultado de este trabajo es un conjunto de niveles de desempeño, similar al realizado en la evaluación del Primer Nivel de la Escolaridad en 2001. Permite dar sentido pedagógico a las

puntuaciones de TRI y aporta mejor información que la Teoría Clásica, en la que se podían definir tramos de puntaje, pero no realizar una descripción detallada de los desempeños asociados.

En este Informe los resultados en Lenguaje y en Matemática serán presentados a través del porcentaje de alumnos en cada nivel de desempeño, previa descripción de qué implica cada uno de los niveles.

Finalmente, es muy importante señalar que los niveles de desempeño tienen carácter exclusivamente descriptivo, es decir, no existe a priori un nivel que pueda ser considerado como de "suficiencia". Esto significa que, a partir del trabajo que ha sido realizado, el Área de Evaluación de Aprendizaje espera generar una discusión entre los actores del sistema que permita establecer cuál de los niveles de desempeño es el que deberían alcanzar todos los niños uruguayos al finalizar la escuela primaria, como condición necesaria para que la escuela cumpla con su papel democratizador.

Esta definición deberá surgir del análisis y la discusión a partir de los resultados. Una de las ventajas de la TRI consiste justamente en que permite aportar descripciones detalladas de los niveles de desempeño, que facilitan la discusión acerca de cuáles tendrían ser los conocimientos y capacidades que deberían lograr todos los niños escolarizados, independientemente de su origen social.

2.7 Descripción de la Prueba de Matemática

El avance continuo y sostenido de las Ciencias y de las aplicaciones de la Matemática a partir del siglo XX hace imposible que una sola persona domine la totalidad de contenidos que éstas abarcan. La especificidad del saber matemático ha adquirido niveles que hacen imprescindible la especialización en diferentes tópicos.

Paralelamente con la creciente complejidad del corpus disciplinario, se hace cada vez más compleja la problemática que conlleva su enseñanza y su aprendizaje. La enseñanza de la Matemática constituye hoy en día uno de los mayores retos a los que se ven enfrentados los docentes en todos los niveles del sistema educativo.

Es común que los distintos sistemas educativos destaquen la necesidad de que los alumnos desarrollen tempranamente competencias en el área que les permitan actuar en el mundo de hoy utilizando de manera inteligente el saber matemático en variadas circunstancias y contextos.

El término "**alfabetización matemática**", utilizado como postulado educativo, propuesto por las asociaciones internacionales de docentes de Matemática, impone la necesidad de tomar conciencia de que su conocimiento es tan importante en el mundo de hoy como el de la lengua oral y escrita.

Entre las dificultades comunes a todos los sistemas educativos que conlleva la enseñanza y el aprendizaje de Matemática se destacan las siguientes:

- 1.- La necesidad de cambios en los niveles obligatorios de educación (Primaria y Secundaria), referidos fundamentalmente al tiempo dedicado al aprendizaje en Matemática y a considerar la diversidad de intereses de los alumnos.
- 2.- La necesidad de reforzar instancias de capacitación y actualización de los docentes del nivel primario y secundario en lo que respecta al dominio disciplinar y al conocimiento de los

avances que se han dado en su didáctica, a fin de que el sistema educativo pueda hacer frente a los cambios.

3.- La necesidad de acordar un programa de estudios actualizado. Esta discusión debe convocar, necesariamente, a los diferentes agentes sociales (educadores, científicos, etc.) a fin de llegar a acuerdos explícitos sobre las competencias básicas necesarias para ejercer la ciudadanía y seguir aprendiendo, así como sobre los modos de hacer posible que esas competencias sean desarrolladas.

4.- Estimular la investigación en educación matemática. Atendiendo a estas necesidades de cambios, desde el año 1996, nuestro país encara la realización de evaluaciones nacionales periódicas en Matemática¹ al finalizar el ciclo primario. Estas evaluaciones cumplen con la función de generar insumos a partir de los cuales los diferentes agentes del sistema educativo puedan realimentar procesos de discusión y transformación en pos de una educación matemática actualizada.

Las pruebas en Matemática son diseñadas atendiendo a evaluar en nuestros alumnos el grado de desarrollo de competencias que en el nivel primario han sido definidas como fundamentales por parte del cuerpo docente nacional, y el grado de manejo de una selección de contenidos relevantes que hacen posible la evaluación de las competencias seleccionadas.

La prueba aplicada en el año 2005 fue elaborada ya teniendo en cuenta que la TRI sería el modelo estadístico que se usaría para analizar los datos. No obstante, se tuvo en cuenta, como antes se ha explicado, el mismo marco teórico que sirvió de fundamento para la construcción de los instrumentos utilizados en los años 1996, 1999 y 2002.

A continuación se presenta una tabla con la descripción de los niveles de desempeño establecidos y, para cada uno de ellos, las habilidades y los procesos que los alumnos deben haber puesto en marcha para resolver las actividades correspondientes, así como los contenidos que abordaron al resolverlas.

Puede observarse que en el Nivel I básicamente se encuentran los logros de aquellos alumnos que solamente fueron capaces de resolver algoritmos básicos en el campo de los números naturales y decimales. Estos logros se van potenciando hasta un máximo que se describe a través de lo que fueron capaces de realizar los alumnos que se ubican en el Nivel IV.

Es de destacar que, si se considera que evaluar es *la adjudicación de un juicio de valor de experto a un hecho u observación*, la tabla de niveles que a continuación se presenta no puede ser considerada como un producto de evaluación, en tanto que no establece cuál es el nivel que debe alcanzar un alumno que finaliza el ciclo primario. Sin embargo, si consideramos a la evaluación *como un proceso de producción de información para la toma de decisiones*, entonces esta tabla de niveles es el producto de una actividad de evaluación. El proceso de evaluación se cerrará cuando los actores que tienen a su cargo las actividades de enseñanza y de aprendizaje de la disciplina discutan y determinen el criterio a partir del cual se pueda establecer el nivel mínimo que debe alcanzar un alumno de sexto año de primaria para ser considerado competente en Matemática y lograr así nivel de suficiencia en el área.

¹ Las actividades de evaluación en Matemática a nivel de sistema realizadas hasta el momento en la educación primaria son las siguientes: 1996 – Evaluación censal; 1999 2002 y 2005- Evaluaciones muestrales.

Tabla 2.3
**Los niveles de desempeño en la prueba de Matemática
 aplicada a 6º año de Educación Primaria. 2005**

Nivel	Competencia matemática	% alumnos
IV 577 puntos o más	<p>En este nivel los alumnos pueden resolver problemas que integran conceptos geométricos y aritméticos complejos. Además, resuelven actividades que requieren de la búsqueda de regularidades y de la aplicación de estrategias no preestablecidas. Muestran habilidad en el manejo de lenguajes matemáticos. Por ejemplo, son capaces de identificar una expresión gráfica a partir de otra dada en lenguaje natural y de reconocer la expresión literal de un perímetro. En la interpretación de un gráfico cartesiano reconocen la tabla de valores que le dio origen. Son capaces de ordenar cuatro fracciones de uso habitual. En la aplicación de algoritmos, respetan el orden de prioridad de las operaciones en una operación combinada.</p>	23,5
III 506 a 576 puntos	<p>Los alumnos de este nivel son capaces de resolver situaciones complejas con estrategias de resolución en varios pasos. También son hábiles al manejar simultáneamente más de una condición entre los datos. Resuelven situaciones aplicando el cálculo de promedio y el de porcentajes (cálculo de IVA). Muestran ser competentes en la interpretación de gráficos que presentan datos complejos para la variable. En lo conceptual geométrico son capaces de reconocer la suma de los ángulos interiores de un triángulo, el desarrollo en el plano de un poliedro y figuras geométricas con igual área y perímetro. Reconocen regularidades en geometría. En aritmética, son capaces de aplicar la definición de división entera para el cálculo de uno de sus términos y, en el campo de los números naturales, de calcular operaciones combinadas con uso de paréntesis. Son competentes en la aplicación del algoritmo de la sustracción en todos los campos numéricos trabajados a este nivel. Representan una fracción en una situación de reparto y, además, pueden identificar el mayor de los números en un conjunto de decimales menores que uno.</p>	23,8
II 419 a 505 puntos	<p>Los alumnos en este nivel responden correctamente a situaciones que requieren seleccionar los datos a partir de la interpretación de tablas y/o gráficos. Son capaces de analizar la información relevante y descartar datos innecesarios en un problema. Resuelven situaciones de proporcionalidad directa que requieren el cálculo de porcentaje de una cantidad. Responden a situaciones problema sencillas que pueden ser resueltas siempre que se opere con números naturales que posibiliten el cálculo mental. Son capaces de identificar un valor solicitado en un gráfico de barras o continuo. En la aplicación de algoritmos, son capaces de resolver en forma correcta sustracciones en el campo de los números naturales y divisiones tanto en los naturales como en los decimales. Suman fracciones sencillas, de uso cotidiano, por asociación y son capaces de identificar la representación gráfica de una fracción. Resuelven operaciones combinadas que, por su presentación, no requieren la aplicación de las reglas de prioridad. Evidencian dominio en conceptos geométricos tales como ángulos complementarios y suplementarios. Reconocen el paralelismo y la perpendicularidad (entre los lados y diagonales de una figura) y las propiedades de los cuadriláteros en general.</p>	33,1
I 308 a 418 puntos	<p>En este nivel, los alumnos son competentes en la aplicación de algoritmos en el campo de los números naturales para las operaciones de multiplicación y división exacta con términos múltiplos de diez; en el campo de los números decimales para las operaciones de adición y multiplicación por un natural. Son capaces de responder a situaciones problema sencillas que pueden ser resueltas a partir de la visualización y en base a la interpretación de representaciones gráficas. Aplican el concepto de unidad de medida de capacidad. En lo conceptual geométrico, reconocen el concepto y las propiedades del rectángulo y visualizan figuras de tres dimensiones en un esquema de dos dimensiones.</p>	18,1
Debajo del Nivel I hasta 307 puntos	Alumnos que no son capaces de resolver las actividades más fáciles de la prueba.	1,5

2.7.1 Descripción de los niveles de desempeño a través de ítems que pertenecen a cada uno de los niveles descriptos

A continuación se presentan ítems propuestos en la prueba que son representativos de cada uno de estos niveles. Éstos han sido liberados en virtud de que no serán utilizados en futuras evaluaciones nacionales.

En los ítems de múltiple opción, al costado de cada respuesta, en recuadro grisado, se presenta información acerca del porcentaje de alumnos que la seleccionó. La clave está identificada con un asterisco y con el valor porcentual en negrita.

En los códigos correspondientes a las actividades abiertas de respuesta producida por el alumno, el crédito total corresponde a la respuesta correcta; el crédito parcial corresponde a aquellas respuestas que, si bien no son correctas, están cerca de serlo, o lo son parcialmente; y es inadecuada cuando el alumno da una respuesta totalmente fuera de contexto.

2.7.1.1 Ejemplo de ítem que se ubica por debajo del Nivel I

Contenido: Perímetro

De las siguientes figuras, ¿cuál es la que tiene mayor perímetro?

Opción	Figura	Porcentaje de alumnos
A)	Rectángulo (100m x 50m)	3 %
B)	Círculo (diámetro 100m)	9 %
C)	Cuadrado (100m x 100m)	77 % *
D)	Figura compuesta (rectángulo con semicírculos)	7 %

Puntaje del ítem: -288

Este ítem resultó ser el más fácil de la prueba. La propuesta supuso que el alumno podría realizar los cálculos pertinentes para seleccionar la figura de mayor perímetro. Es evidente que este no fue el proceso de resolución seguido por la mayoría de los alumnos, de lo contrario la dificultad sería mayor si lo hubieran resuelto realizando cálculos, que son propios del currículo del grado (longitud de circunferencia y figuras compuestas como en la opción "D"). Esto apoya la hipótesis de que la resolución fue visual. Es posible suponer que el hecho de que las figuras se presenten trazadas a escala facilita la percepción de la opción "C" como la que reúne las condiciones solicitadas por la consigna.

2.7.1.2 Ejemplos de ítems que se ubican en Nivel I*Contenido: Multiplicación*

En esta multiplicación, ¿qué factor corresponde en el casillero vacío?

$$\square \times 13 = 845$$

A) 60	3%
B) 65	83% *
C) 832	7%
D) 10985	3%

Puntaje del ítem: 308

Esta es una actividad típica de las que definen este nivel, en el que predominan fundamentalmente las habilidades de cálculo (multiplicación) en el campo de los números naturales.

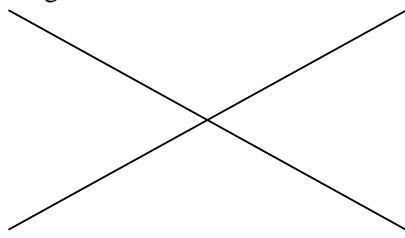
En la resolución de este ítem el alumno puede aplicar la operación inversa de la multiplicación: dividiendo el producto entre el factor conocido el cociente corresponderá al factor buscado. O bien, puede buscar por tanteo el factor pedido, para lo cual podrá recurrir a variadas estrategias como por ejemplo multiplicar cada uno de los valores que presentan las opciones por el factor dado hasta confirmar el producto. En este último caso, si el alumno aprovecha el formato de múltiple opción, podrá descartar las opciones C y D por ser factores con tres y cuatro cifras respectivamente, los cuales no corresponden a este producto de tres cifras. Luego puede seleccionar B en función que las unidades del producto indican que el factor debe terminar en 5.

Es muy probable que la estrategia de tanteo fuera la más utilizada en virtud de que la división ofrece mayor dificultad y que recién los alumnos que se ubican en el Nivel III manifiestan dominio del concepto y del algoritmo de división.

Contenido: Rectángulo, propiedades

Los segmentos iguales trazados, que se cortan en su punto medio, corresponden a las diagonales de un cuadrilátero.
El cuadrilátero determinado por estas diagonales es un

A) cuadrado	9 %
B) rombo	5%
C) romboide	5%
D) rectángulo	78%*



Puntaje del ítem: 321

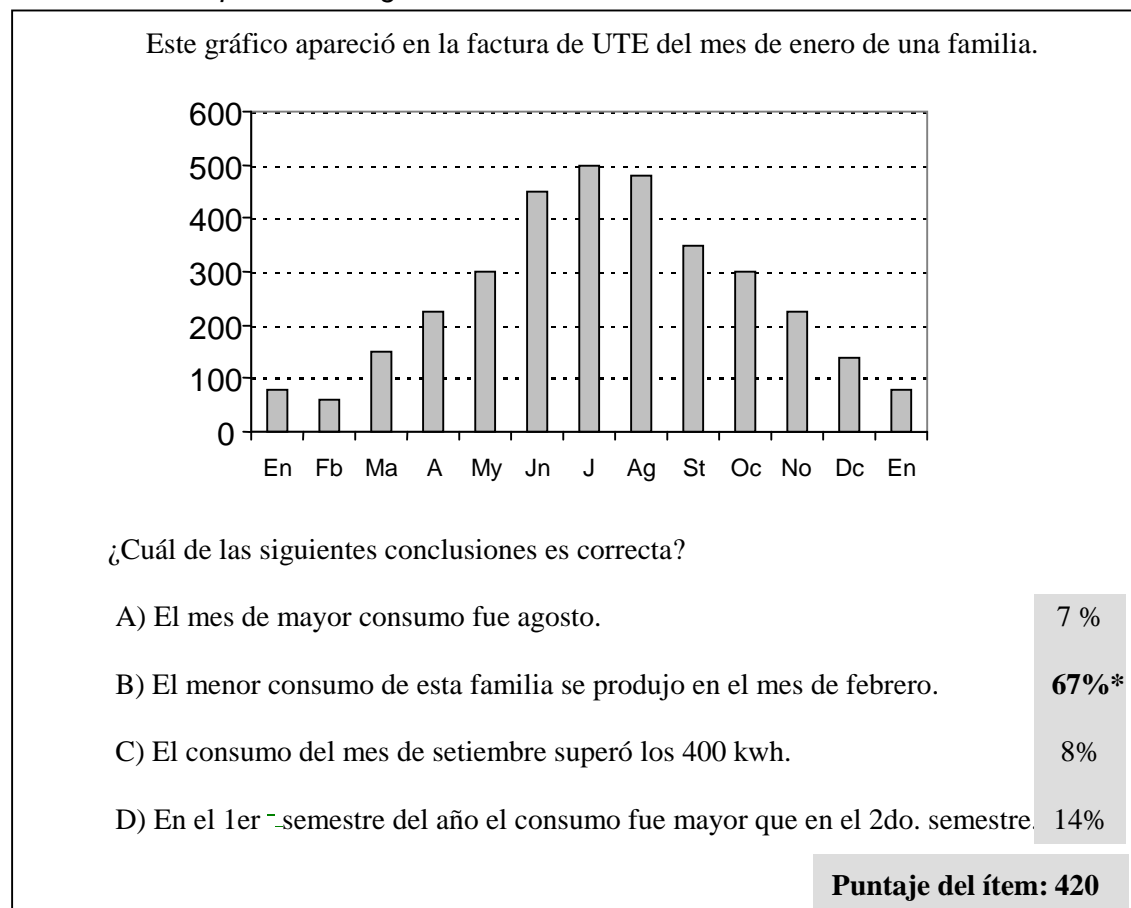
La presente actividad fue resuelta en forma correcta por aproximadamente 4/5 de los alumnos. Al igual que en el ítem anterior, la capacidad de interpretar representaciones

gráficas favorece su resolución. Un alumno con conocimientos de las propiedades de las diagonales de los cuadriláteros, leyendo la consigna y observando que la figura no presenta segmentos perpendiculares, descartará las opciones no clave (A, B y C).

Es posible también que muchos alumnos hayan reconocido la figura al identificar los extremos de los segmentos como los vértices de ésta e imaginar su trazado. En este caso no hay evidencia de que hayan aplicado las propiedades de los cuadriláteros.

2.7.1.2 Ejemplos de ítems que se ubican en el Nivel II

Contenido: Interpretación de gráficos



Se esperaba que para la resolución de esta actividad los alumnos hubieran activado habilidades de lectura de gráfico de barras e interpretación de textos de modo de analizar las conclusiones para seleccionar la opción correcta. Como puede observarse, tres de las opciones exigen del estudiante la capacidad para interpretar el valor de la variable en un solo momento del año. Para detectar la veracidad o falsedad de la afirmación respecto del gráfico basta con observar independientemente las longitudes de tres de las barras. La última opción difiere de las anteriores en tanto compara el consumo del primer semestre con el del segundo. Más de las dos terceras partes de los alumnos evaluados detectó la afirmación pedida.

Contenido: Análisis de datos

Solicito un préstamo de \$ 6500 que debo abonar en 15 cuotas de \$ 785.
Al saldar mi deuda habré pagado en total:

A) \$ 5275	10 %
B) \$ 7285	15%
C) \$ 7300	8%
D) \$ 11775	62%*

Puntaje del ítem: 446

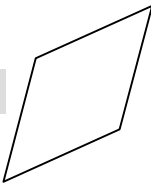
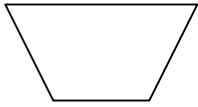
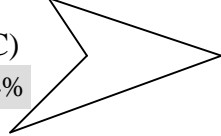
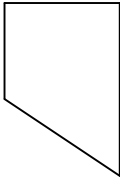
Esta actividad es una de las que define el Nivel 2 de desempeños. Presenta al alumno una situación problemática del área de la aritmética. Es posible resolverla con una sola operación: una multiplicación que calcula el total que se paga con las 15 cuotas (15×785). El dato que corresponde a la cantidad recibida como préstamo no sería utilizado al aplicar la estrategia anterior. También es posible que algún estudiante razonara de la siguiente forma: el resultado buscado es un número mayor que $10 \times 785 = 7850$, con lo que quedan descartadas las opciones no clave. En este caso, de igual modo, está presente el hecho de que el estudiante comprendió que el problema se resuelve mediante la aplicación de la multiplicación, aunque no la haya efectuado. Se evidencia la potencialidad de aprovechar el formato del ítem de múltiple opción.

El dato de monto pedido en préstamo fue utilizado para generar las opciones incorrectas a los efectos de detectar aquellos alumnos que no han sido capaces de resolver la actividad, por considerar datos no necesarios en el problema.

Contenido: Cuadriláteros, propiedades.

¿Cuál de estas figuras cumple todas las siguientes propiedades?

- Es un cuadrilátero.
- Tiene sólo dos lados paralelos.
- Tiene sólo un ángulo agudo.

<p>A) </p> <p>17%</p>	<p>B) </p> <p>18%</p>	<p>C) </p> <p>14%</p>	<p>D) </p> <p>47%*</p>
--	--	---	--

Puntaje del ítem: 501

Esta actividad propone al alumno identificar una figura geométrica a partir de su descripción. Para lograr resolver con éxito la actividad, el alumno podrá aplicar sus conocimientos sobre conceptos tales como: la clasificación de polígonos por el número de lados; el paralelismo entre rectas; la clasificación de ángulos por su amplitud. Esta actividad resulta de una complejidad mayor debido a que plantea la clasificación en base a propiedades más generales y que corresponden a cuadriláteros que no son tan conocidos para él como lo son el rectángulo y el cuadrado. Además, el estudiante se ve exigido al tener que manejar vocabulario específico de geometría.

Contenido: Sustracción

La diferencia entre dos números es 1348.
El sustraendo es 2769.
El minuendo es

-
<u>2769</u>
1348

A) 4117	54%*
B) 4107	10%
C) 4126	5%
D) 1421	26%

Puntaje del ítem: 468

Este ítem evalúa la resolución de sustracciones en el conjunto de los números naturales. La consigna de la actividad se presenta en dos lenguajes: natural y simbólico. Este diseño responde al objetivo de que el posible desconocimiento del significado de términos tales como minuendo, sustraendo y diferencia no interfiera en su resolución.

La actividad espera que el alumno conozca la definición de sustracción. No obstante, la actividad puede ser resuelta por ensayo y error, sustituyendo en el minuendo cada una de las opciones presentadas hasta comprobar la igualdad dada, resolviendo por suma.

Cerca de la mitad de los alumnos que lo respondieron fueron capaces de resolverlo en forma satisfactoria. Como puede apreciarse, más de la cuarta parte de los alumnos elige la opción D (es la opción de menor logro cognitivo) como la correcta. Esto puede dar evidencias de que algunos de estos alumnos no atendieron al concepto básico, de que la sustracción en el campo de los números naturales solo es posible si el minuendo es mayor que el sustraendo.

Contenido: Estrategias de resolución y verificación de resultados

Luis compra en la papelería 4 cuadernolas y 3 lápices. Paga \$ 120.
Si el precio de cada lápiz es de \$ 4, entonces el precio de cada cuadernola es

A) \$ 27	53%*
B) \$ 33	6%
C) \$ 40	6%
D) \$ 108	32%

Puntaje del ítem: 473

Esta situación problemática puede requerir una resolución en varios pasos; esta complejidad se atenúa porque las cantidades pueden facilitar la posible estrategia del cálculo mental.

Si se planteara en forma simbólica, la estrategia que la resuelve se corresponde a una ecuación de primer grado con una incógnita: $4 \times \$x + 3 \times \$4 = \$120$.

Una posible estrategia de resolución supone buscar el valor que verifica la igualdad sustituyendo la incógnita por los valores de las opciones. Esta última estrategia igualmente determina la realización de un mínimo de tres operaciones.

Es muy probable que alumnos de este nivel lo resuelvan calculando primeramente el precio de los lápices; luego restando al total este costo para finalmente dividir este resultado entre cuatro para calcular el precio de una cuadernola.

También, y de una forma más sencilla, el alumno puede aprovechar el formato de múltiple opción para resolver la situación. Es evidente que el precio de las cuatro cuadernolas debe ser inferior a 120. De esta forma una cuadernola debe valer menos de \$30. Solamente una opción cumple con esta condición.

De acuerdo a los resultados, algo más de la mitad de los estudiantes que la abordaron (53%) es capaz de llegar con éxito a este resultado. El 32% de los alumnos llega a un resultado parcial (precio de las cuatro cuadernolas).

2.7.1.3 Ejemplos de ítems que se ubican en el Nivel III

Contenido: *División entera*

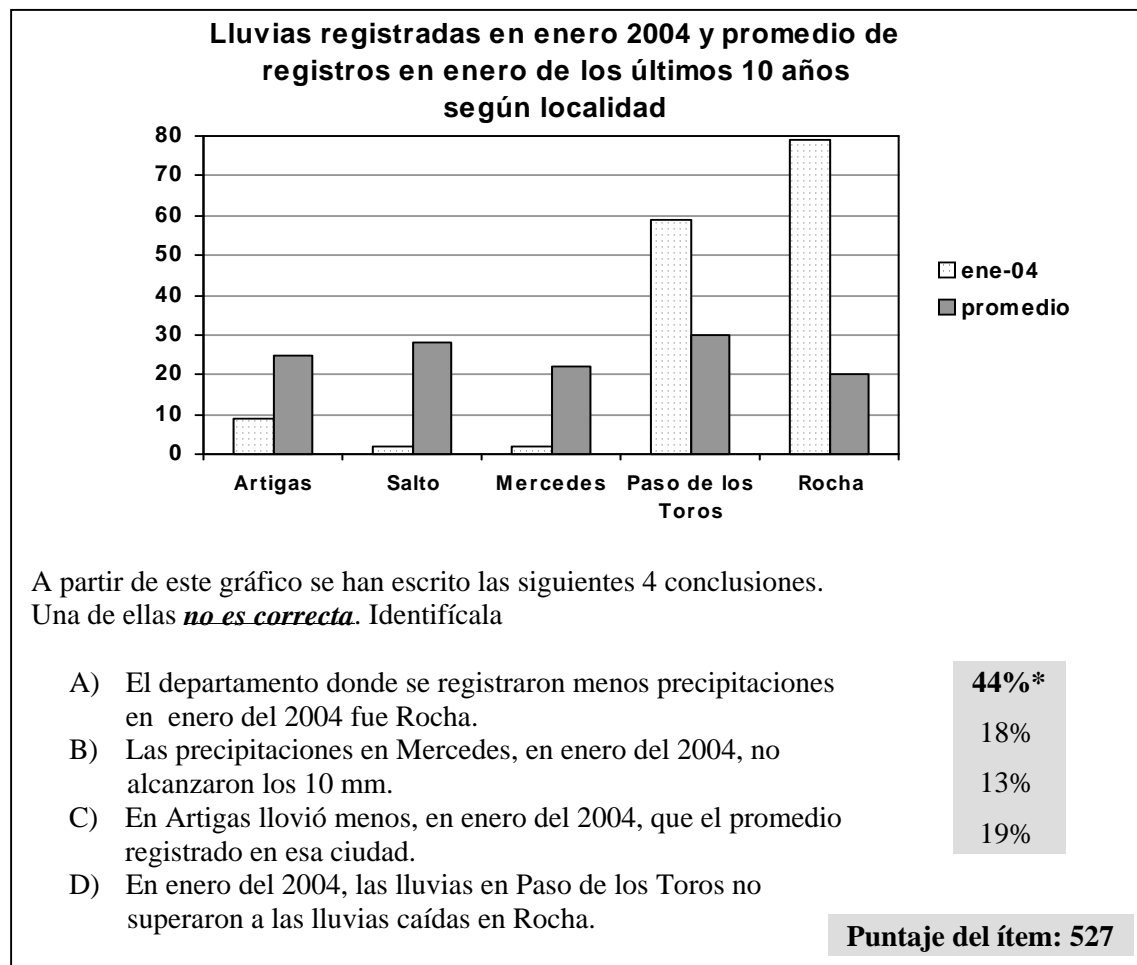
¿Cuál es el dividendo de esta división?	
$\begin{array}{r} \boxed{} \quad \quad 14 \\ \hline \text{Resto} = 11 \quad 26 \end{array}$	
A) 375	44%*
B) 364	28%
C) 353	12%
D) 300	10%
Puntaje del ítem: 524	

Esta actividad corresponde a un grado de conocimiento matemático medio, su puntaje se halla muy cercano al promedio (500 puntos).

La actividad pretende que el alumno aplique el concepto de división entera, no obstante, éste puede resolverla verificando la división para cada una de las opciones presentadas hasta encontrar la que corresponde. De todos modos esta estrategia siempre exige que el alumno considere el resto. La segunda opción detecta a algunos de los alumnos que dividiendo en forma correcta no tienen en cuenta el resto de la división. El mayor porcentaje de alumnos que la resuelve en forma incorrecta opta por esta opción (28%), que presenta el dividendo que corresponde a la división exacta (26 x 14).

Asimismo, un alumno competente puede aprovechar el formato de múltiple opción para seleccionar con facilidad la clave. Si considera solamente las unidades sabe que el dividendo es una cifra terminada en 5 (6 x 4 +1). Solamente una opción cumple con esta condición que es la clave

Contenido: Interpretación de gráficos



La resolución correcta de esta actividad requiere un mayor desempeño que el necesario para resolver la actividad asociada al gráfico de barras presentado como representativa del Nivel I.

En este caso el alumno deberá identificar la información solicitada teniendo en cuenta dos variables: las lluvias registradas en un mes determinado y el promedio correspondiente a ese mes tomado durante un lapso de 10 años. A esta complejidad se le suma la dificultad de comprensión de la consigna: se le solicita al alumno identificar una premisa que no es verdadera entre otras tres que sí lo son.

El porcentaje de alumnos que lo responde correctamente no alcanza al 50%, y las restantes opciones registran porcentajes similares de respuesta.

Contenido: Porcentaje

En una fiesta hay 410 invitados.
El 90 % de ellos consume refrescos.
La cantidad de personas que **no lo hacen** es:

- A) 400 9%
B) 369 30%
C) 320 15%
D) 41 **42%***

Puntaje del ítem: 544

En el Nivel III se ubican las actividades que requieren el cálculo de porcentajes. Corresponden a este nivel las situaciones de descuento y de recargo con cálculo de porcentajes.

En la resolución de esta actividad la complejidad está relacionada con una premisa redactada en forma negativa. El alumno debe considerar el todo en el concepto de porcentaje (100%) y el porcentaje que corresponde a quienes no cumplen con la condición expuesta (10% del total). Estas características hacen que, pese a la simplicidad de los datos numéricos presentados y de los cálculos que la resuelven, sea una situación compleja con resolución en varios pasos.

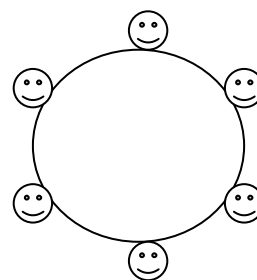
En los porcentajes de respuesta puede observarse que aproximadamente un tercio de los alumnos selecciona una opción que identifica a quienes son capaces de calcular correctamente el 90% de 410.

2.7.1.4 Ejemplos de ítems que se ubican en el Nivel IV

Contenido: Longitud de circunferencia

¿Cuánto debe medir aproximadamente, como mínimo, el diámetro de una mesa redonda, para que 6 personas sentadas a su alrededor dispongan de 0,60m cada una? Escribe todos los cálculos que hagas para llegar a la respuesta.

Código	Porcentaje
Crédito total	8
Crédito parcial	36
Sin crédito	15
Respuesta inadecuada	3
Omisión	39
Total	100



Puntaje del ítem: 663

La resolución de esta actividad abierta de respuesta producida por el alumno supone crear una estrategia no preestablecida que le permita, luego de ejecutarla, encontrar un resultado pertinente. Para ello es posible integrar contenidos geométricos abordados en años anteriores o en el programa en curso.

Esta actividad puede ser resuelta a través del cálculo de la longitud de una circunferencia o por triangulación, a partir del concepto de hexágono entre otras estrategias.

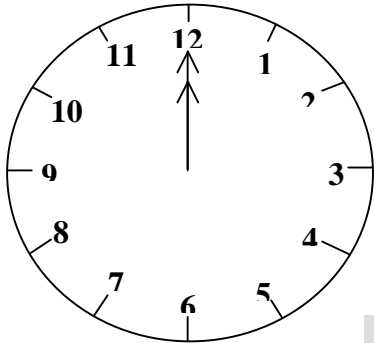
Al ser un ítem de respuesta abierta se le otorgó un crédito según la respuesta y la estrategia explicitada. Se consideraron correctos valores de respuesta comprendidos entre 1 m y 1,3 m.

Las respuestas que obtuvieron crédito parcial fueron aquellas que presentan en forma correcta alguna parte de la estrategia de solución elegida pero no llegan a explicitar el resultado esperado.

Contenido: Análisis de datos, estrategias de resolución y verificación

¿Cuántos días pasarán antes de que un reloj de agujas, que atrasa 30 minutos al día, muestre la hora correcta otra vez?

A)	6 días	38%
B)	12 días	27%
C)	24 días	19%*
D)	48 días	9%



Puntaje del ítem: 975

Resolver esta actividad correctamente representa un desafío cognitivo muy grande para un alumno de sexto año. Su puntaje en TRI la ubica como una de las dos actividades más difíciles de la prueba. El porcentaje de respuesta correcta también indica que son muy pocos los alumnos que llegan a este nivel de competencia.

La dificultad de la propuesta quizá radique en la complejidad de los aspectos que se debe considerar. Efectivamente, en los datos presentados se deben tener en cuenta aspectos tales como: el atraso diario, el hecho de que en un día las agujas barren dos veces la esfera del reloj, el que las agujas deben llegar a la posición inicial indicando la hora correcta y finalmente las unidades utilizadas minutos, horas, días. Esta complejidad de los datos en juego conspira con una posible solución sencilla. Un posible razonamiento puede ser el siguiente: las agujas se volverán a encontrar luego de los días necesarios para que el reloj muestre 12 horas de atraso; si el atraso es de media hora por día se necesitarán 24 días para que vuelvan a coincidir.

El mayor porcentaje de alumnos (38%) considera una opción que rastrea un error grueso (no tiene en cuenta los datos ni logra establecer una relación adecuada entre ellos), un porcentaje menor, pero aún mayor que los que la resolvieron correctamente, selecciona la opción que rastrea un error más fino (no consideran el número de veces que el horario recorre la esfera para coincidir con el minuterero, de acuerdo al atraso de media hora diaria, en la posición inicial).

Contenido: Múltiplos y divisores

¿En cuál de las siguientes opciones ubicarías en el recuadro en blanco la expresión “es múltiplo de” para generar un enunciado correcto?

A)	38	<input type="text"/>	8	8%
B)	30	<input type="text"/>	30	24%*
C)	12	<input type="text"/>	24	55%
D)	34	<input type="text"/>	4	6%

Puntaje del ítem: 1004

Esta actividad resultó ser la de mayor dificultad de la prueba. Solamente el 31% de los alumnos que se ubican en el Nivel IV de desempeño la resuelven en forma correcta.

Si bien la actividad considera un contenido que corresponde al currículo del grado, evidencia que es de difícil apropiamiento ya sea porque ha tenido pocas instancias de abordaje o porque posee una dificultad intrínseca. La experiencia docente coincide en que conceptos tales como: “todo número es múltiplo de sí mismo”, “todo número es múltiplo de 1” y que “el 0 es múltiplo de todos los números” resultan de difícil comprensión hasta en el nivel secundario.

Puede observarse que el mayor porcentaje de respuesta se da en la opción C, lo cual es comprensible, ya que los alumnos pueden tener la idea de que si un número es múltiplo de otro debe ser necesariamente uno mayor y el otro menor y que el mayor es divisible entre el menor (24 es múltiplo de 12 y 12 es divisor de 24). En esta opción la relación que corresponde es: “es divisor de”, pero parece ser que la concepción del alumno es tan fuerte que la invierte para seleccionar el par de números coincidente con su idea.

2.8 Los niveles de desempeño en la Prueba de Lenguaje

Son múltiples los factores que inciden en la comprensión de un texto. Algunos de estos factores se refieren al sujeto que lee (la memoria, la atención, la significatividad del tema para el sujeto, sus conocimientos previos sobre el tema), otros al texto que se lee (la organización de la información, su sintaxis, su vocabulario, etc.).

Con respecto a los factores relacionados con el lector, todos los modelos del proceso de la lectura concuerdan en que lo decisivo es lo que el sujeto ya sabe (“información no visual”, “conocimientos previos”, “esquemas conceptuales”, según los distintos autores). Cuanto más se sabe sobre el tema y cuanto más se haya frecuentado el tipo de texto que se va a leer, más posibilidades de comprender, es decir, de anticipar, de formular hipótesis y de relacionar con otros textos y otras situaciones tendrá el lector.

Ningún texto es absolutamente explícito. Todo texto requiere que el lector realice algún tipo de inferencias para poder construir su coherencia. Esa reposición de información implícita que el lector hace la realiza a partir de lo que ya sabe. Por lo tanto, uno de los obstáculos más frecuentes que el lector enfrenta para dar sentido a un texto es la falta de

conocimientos que le permitan inferir lo que el texto no dice explícitamente. Investigaciones en psicología cognitiva plantean que los conocimientos se organizan básicamente de dos maneras: como esquemas situacionales y como esquemas conceptuales. Los primeros se forman por la experiencia cotidiana, son los conocimientos que se deducen de las situaciones concretas, y de ellos depende, en buena medida, la comprensión de los textos narrativos. Los segundos son las organizaciones de los conceptos abstractos, que se realizan según un orden lógico y son, en gran parte, los que se ponen en juego en la comprensión de los textos explicativos y argumentativos. Estos esquemas se forman, en su mayor parte, en el ámbito de la escuela, al punto que se podría decir que una de las funciones de la escuela es colaborar en la formación de una red conceptual sólida que permita a los estudiantes seguir aprendiendo.

Los textos narrativos y los explicativos responden, en la concepción de Bruner, a dos modalidades de pensamiento diferentes: la narrativa, que es la modalidad típica en que se transmite la experiencia, y la paradigmática, que se desarrolla predominantemente sobre el eje de la lógica. La elección de textos narrativos y textos expositivos en esta evaluación responde a esta manera de mirar los textos desde el punto de vista de su exigencia cognitiva.

Tabla 2.2
Los niveles de desempeño en la prueba de Lenguaje
aplicada a 6º año de Educación Primaria 2005

Nivel	Competencia	% alumnos
IV 600 puntos y más	<p>Los estudiantes en este nivel manifiestan una comprensión global de los textos, de manera que logran abstraer el tema, relacionando los aspectos parciales.</p> <p>Demuestran un manejo de la clasificación de los textos no solo por su uso social, sino también por su estructura y el uso del lenguaje.</p> <p>Interpretan el motivo por el cual un personaje actúa de cierta manera, para lo cual deben relacionar información que se ubica a lo largo del texto. Para descartar las demás alternativas propuestas, deben haber comprendido globalmente cómo funciona el mundo creado por el texto.</p> <p>Ubican cuál es la causa que provoca una situación planteada en el texto narrativo. Para hacerlo deben relacionar e inferir información. Además deben comprender el ordenamiento cronológico de los hechos narrados.</p> <p>Reconocen la formulación más precisa del tema de un texto explicativo. En el texto se utilizan términos específicos de las ciencias y enunciados extensos. Logran descartar las opciones que enuncian aspectos parciales, para lo cual deben generalizar a partir de pistas lingüísticas.</p> <p>Identifican la clase de texto de un artículo explicativo. Para hacerlo deben reconocer la estructura, la intención y el estilo.</p>	19,3
III 520 a 599 puntos	<p>Los estudiantes en este nivel manifiestan una comprensión global tanto del texto narrativo como del explicativo. Por lo tanto, reconocen el tema y descartan opciones que enuncian aspectos parciales. Además, logran inferir los motivos que impulsan los actos de los personajes de una narración.</p> <p>En este nivel comienzan a manejar clasificaciones de los textos basadas en el uso social.</p> <p>Reconocen por qué los personajes actúan de cierta manera. La información es ofrecida directamente por uno de los personajes, de forma casi literal. Dificulta su identificación el hecho de que el dato sea requerido mediante expresiones que significan negación ("no hacer", "evitar"). También identifican por qué un personaje actúa de cierta manera, cuando la información no es preguntada literalmente sino mediante una paráfrasis. En ambos casos, la información aparece sobre el final del cuento.</p>	24,4
II 440 a 519 puntos	<p>En este nivel, los estudiantes resuelven actividades que implican la inferencia de información implícita, para lo cual relacionan datos que se ofrecen a lo largo del texto, en lugares destacados. También logran inferir el significado de términos que no son de uso cotidiano, a partir de pistas lingüísticas.</p> <p>Identifican qué hace un personaje, cuyo acto es la resolución de la narración. Para hacerlo deben asociar datos que aparecen explícitos en distintos fragmentos del texto. La información que se les pide aparece de forma casi literal en la última línea del cuento, y cae fuera de las expectativas del lector, es decir, que deben haberlo leído todo para poder identificarla.</p> <p>En un artículo de difusión científica, de estructura explicativa que carece de imágenes que faciliten la comprensión, identifica cuál es la definición del objeto de que trata el texto. La información aparece explícita de manera parcialmente literal en un espacio destacado del artículo, y reiterada literalmente en dos oportunidades.</p>	28,1
I 439 a 350 puntos	<p>Los estudiantes en este nivel son capaces de localizar en el texto la información que se les requiere de manera literal tanto en un texto narrativo como en uno explicativo. Descartan soluciones que se encuentran conceptualmente alejadas de lo solicitado. Además, infieren información implícita cuando las pistas lingüísticas son claras y bien localizadas.</p> <p>Reconocen qué hace o qué le pasa a un personaje de un texto narrativo si la información aparece literalmente en el texto y se reitera en un fragmento breve del cuento. Identifican la opción correcta entre otras alejadas conceptualmente.</p> <p>En una secuencia de diálogo entre dos personajes de un texto narrativo compuesta por breves intercambios acerca de la situación planteada en el texto, reconocen lo que motiva la respuesta de un personaje, para lo cual deben poner en relación los dichos de ambos. Se valen de las pistas lingüísticas que el texto presenta para inferir información implícita.</p> <p>Identifican información que aparece explícita en un texto explicativo. El dato requerido es un nombre propio que aparece en un lugar destacado del artículo, de forma cercana a lo literal y se repite en otra parte del texto.</p>	23,3
Debajo del Nivel I hasta 349 puntos	<p>Alumnos que no son capaces de resolver las actividades más fáciles de la prueba.</p>	4,9

A continuación, se presentan preguntas que resultaron típicas de cada nivel. Estas preguntas, así como los textos a los cuales se asocian, son dados a conocer en este informe, por lo cual no serán utilizados en futuras evaluaciones. Presentamos en primer lugar un texto narrativo y luego uno explicativo. Para cada texto presentamos ítems que corresponden a diferentes niveles.

2.8.1 Descripción de los niveles de desempeño a través de ítems representativos

"EL país sin punta"

Juanito trotamundos era un gran viajero. Una vez llegó a una ciudad donde las esquinas de las casas eran redondas y los tejados no terminaban en punta, sino en una especie de joroba suave. En la calle había un rosal y Juanito arrancó una rosa para ponérsela en el ojal de la chaqueta. Mientras la sacaba se dio cuenta de que las espinas no pinchaban, no tenían puntas y parecían de goma y hacían cosquillas en las manos.

De pronto apareció un guardia municipal y le dijo sonriendo:

—¿No sabía que está prohibido arrancar rosas?

—¡Lo siento, no había pensado en ello!

—En este caso solo pagará la mitad de la multa —dijo el guardián sonriendo.

Juanito observó que escribía la multa con un lápiz sin punta, y le dijo:

—¿Me permite ver su espada?

—Con mucho gusto —contestó el guardia.

Y, naturalmente, la espada tampoco tenía punta.

—¿Pero qué país es éste? —preguntó Juanito.

—El país sin punta. Y ahora, por favor, deme dos bofetadas —dijo el guardia.

Juanito se quedó de piedra. Y respondió:

—¡Por el amor de Dios, no quiero ir a la cárcel por maltrato a un oficial! Las dos bofetadas, en todo caso, debería recibirlas yo.

—Pero aquí se hace así —explicó gentilmente el guardia—. Por una multa entera, cuatro bofetadas, por media multa, solo dos.

—¿Al guardia?

—Al guardia.

—¡Pero es injusto! ¡Es terrible!

—Oh, ¡claro que es injusto! —dijo el guardia—. La cosa es tan odiosa que la gente, para no verse obligada a abofetear a unos pobres inocentes, se cuida de no hacer nada contra la ley. Venga, deme esas dos bofetadas y otra vez ponga más atención a lo que hace.

—Pero yo no quiero dárselas; sí acaso una caricia.

—Si es así, —concluyó el guardián— lo tendré que acompañar a la frontera.

Y Juanito, avergonzado, se vio obligado a abandonar el país sin punta; pero aún hoy sueña con volver.

"El país sin punta"
del libro "Cuentos por teléfono"
Gianni Rodari
(adaptación)

2.8.1.1 Ejemplos de ítems que describen los niveles de desempeño en la comprensión de un texto narrativo

Nivel I

Objetivo del ítem: identificar información explícita

El guardia municipal se dirigió a Juanito		
A)	agresivamente.	4.4%
B)	malhumorado.	4.1%
C)	sonriendo.	85.9 % *
D)	autoritariamente.	5.7%
		Puntaje del ítem: 312

El alumno debe ubicar la información que se encuentra explícita en el texto. El dato se encuentra en el segundo párrafo, cuando el personaje aparece por primera vez, y se reitera como aclaración del narrador a la segunda intervención del personaje. El estudiante debe atenerse a la información que el texto le ofrece y descartar las alternativas que están conceptualmente alejadas, pero que podrían ser las actitudes esperables en el mundo "real" ante la falta que comete el protagonista. Este ítem resultó muy fácil, el 74,4 % de los estudiantes cuyos puntajes se ubican en el primer nivel acertaron con la clave.

Seguramente quienes eligieron las demás opciones se guiaron por sus ideas previas a la lectura del texto, que no fueron modificadas por la comprensión de las características del personaje ni por la "lógica" de la historia.

Solo el 15,5% de los alumnos que se ubicaron en el Bajo I la resolvió.

Nivel III

Objetivo del ítem: reconocimiento de información explícita

Los habitantes de "El país sin punta" tratan de no hacer nada contra la ley para evitar		
A)	abofetear a los guardias.	54.2 % *
B)	recibir bofetadas.	12.8 %
C)	tener que irse del país.	18 %
D)	ir a la cárcel.	15%
		Puntaje del ítem: 530

Sobre el final del cuento, en palabras del guardia, se encuentra la explicación por la que se pregunta: "...la gente, para no verse obligada a abofetear a unos pobres inocentes, se cuida de no hacer nada contra la ley".

Observemos que esta actividad se asocia al mismo texto que la anterior, y tiene el mismo objetivo, que es el reconocimiento de información que aparece explícita en el cuento. Sin embargo, la dificultad es mucho mayor al punto que la anterior se ubica en el Nivel I, y esta en el Nivel III.

Podemos suponer algunos factores que obraron en el sentido de aumentar la dificultad de la pregunta:

- la mayor extensión y complejidad sintáctica del enunciado que responde a la pregunta,
- su ubicación al final del cuento,
- la presencia de expresiones de valor negativo en la formulación de la pregunta ("*no hacer nada*", "*evitar*").

El 75,7% de los estudiantes cuyos puntajes se ubican en el tercer nivel acertó la clave. Solamente el 21,3% de los estudiantes del Nivel I la contesta correctamente. El 30,5% de los estudiantes de este nivel prefiere el distractor "C". Esta elección posiblemente se deba a que el alumno se centra en lo que pasó finalmente al personaje por no cumplir la pena (es decir, que se queda en el plano de los hechos) y no llega a realizar la generalización que establece la norma para todos, que es por lo que se le pregunta en este ítem.

Nivel IV

Objetivo del ítem: evaluar la comprensión global y realizar una interpretación de la resolución de la historia.

Juanito tuvo que abandonar " <i>El país sin punta</i> " porque	
A) le dio dos bofetadas al guardia.	11.4%
B) el guardia lo quería llevar a la cárcel.	8.7 %
C) no quiso cumplir con la multa que el guardia le aplicó.	38.3 %*
D) el guardia lo acompañó a la frontera.	41.7 %
Puntaje del ítem: 636	

Identificar la causa de que el protagonista haya tenido que abandonar el país implica haber interpretado, a partir del diálogo entre éste y el guardia, que la negativa de Juanito, que resulta, sin embargo, muy amable (habla de una caricia), implica una violación de las normas, y que el guardia actúa según esas normas y no según la otra "lógica". Un 67,3% de los estudiantes del Nivel IV acertó la clave. En los demás tramos de puntaje se destaca la preferencia que los estudiantes tuvieron por el distractor "D": 39,7% de los del Nivel I, 48% de los del Nivel II y el 47,1% de los del Nivel III. En los tres casos, este distractor resultó más atractivo que la clave. Tal vez estos alumnos hayan interpretado "el guardia lo acompañó a la frontera" como "lo echó". De todos modos, no establecen correctamente la causa de los hechos.

EL QUETZAL,

un símbolo de prosperidad del pasado y del presente

Texto por EUGENIA PALLARES/SIERRA MADRE

A partir de una preocupante problemática ambiental mundial, la humanidad trata hoy de retomar su antigua relación con la naturaleza y redefine sus ideas sobre el desarrollo a partir de una idea básica en los programas de conservación de la naturaleza: "desarrollo sustentable".

Y es en estos programas donde especies como el quetzal adquieren de nuevo el valor de símbolo de poder y de prosperidad que tuvieron para las culturas que se asentaron en un ambiente natural. Hoy, con el auge de una nueva cultura ambientalista, la humanidad reencuentra el valor de esta ave-símbolo.

Los bosques de niebla que constituyen el hábitat del quetzal son los más amenazados de la Tierra.

El Triunfo es un bosque de niebla extraordinario, que suma a la belleza de un paisaje envuelto en una incesante bruma, una enorme variedad de especies de plantas y animales.

El bosque de niebla de El Triunfo, localizado en la Sierra Madre de Chiapas, es uno de los últimos bosques de México en donde aún vuela el mágico quetzal.

El quetzal puede ser hoy, como antes, un emblema en la preservación de la riqueza natural de estos bosques. Pero conservar estas especies-símbolos significa en la actualidad algo más: salvar toda la flora y la fauna que son el hábitat donde estas especies coexisten.

Fuente: National Geographic, junio 1998

2.8.1.2 Ejemplos de ítems que describen los niveles de desempeño en la comprensión de un texto expositivo

Nivel II

Objetivo del ítem: reconocimiento de información explícita

El quetzal es		
A)	un ave – símbolo.	74.9 %*
B)	una leyenda.	6.9 %
C)	un lugar histórico.	5.3 %
D)	un bosque de niebla.	12.9 %
		Puntaje del ítem: 470

En el segundo párrafo del texto aparece la información requerida en forma explícita, y en el título aparece una aposición de Quetzal, que indica la respuesta. A lo largo del texto se reitera la idea de símbolo, pero para quien no sabe antes de leer el texto que el quetzal es un ave, esto se aclara recién al final del segundo párrafo ("*esta ave símbolo*"). El hecho de tratarse de un artículo en el que abunda el léxico específico de las ciencias tal vez haya sido el factor que imprimió la dificultad de la tarea.

Nivel III

El texto: " <i>EL QUETZAL, un símbolo de prosperidad del pasado y del presente</i> ", es		
A)	una narración fantástica sobre un ave legendaria.	24.4 %
B)	un artículo ambientalista por la preservación del hábitat del quetzal.	52.5 % *
C)	un aviso publicitario sobre el hábitat del quetzal.	12.3 %
D)	un reglamento para evitar la extinción del quetzal.	10.8 %
		Puntaje del ítem: 556

Quienes eligieron la opción correcta (el 61,6% de los del Nivel III) seguramente hayan captado que este texto carece de las categorías de la narración. Justamente, la opción por el distractor "A" ("*una narración fantástica sobre un ave legendaria*") fue mayoritaria entre los estudiantes del Nivel I (37,1%), y recogió un 28,8% de las preferencias entre los del Nivel II.

La clasificación suele ser una operación cognitivamente exigente, sobre todo cuando entran en juego varios factores que inciden en la caracterización de las clases (lenguaje, tema, intención, fuente, estructura). Recién en el Nivel III los estudiantes parecen lograr clasificar los textos.

Capítulo 3

Resultados de los aprendizajes en Matemática y Lenguaje por Contexto sociocultural de las escuelas

3.1 Principales características de la muestra

Para la presente evaluación se realizó un muestreo probabilístico estratificado según contexto sociocultural, en el que el peso de cada estrato en la muestra fue proporcional al peso de los estratos en el universo (según el número de alumnos en 6º año). Para seleccionar la muestra se procedió de la siguiente forma: en la primera etapa, se sortearon escuelas con probabilidad proporcional a la matrícula de 6º año; en la segunda etapa, en aquellas escuelas que tenían más de un grupo, se sorteó aleatoriamente solamente uno de ellos. Finalmente, cada uno de los grupos seleccionados fue censado.

En cuanto al tamaño de la muestra, en esta evaluación se seleccionaron 6000 casos, a los que se le aplicaron pruebas diseñadas con TRI sobre la base del piloto realizado en 2004. Es a partir de los resultados de esta muestra que se podrá comparar con futuras evaluaciones, conformando por lo tanto la nueva línea de base hacia el futuro. Sin embargo, para comparar con la evaluación de 2002 (donde se trabajó según TCT) se requirió que otro conjunto de alumnos contestara pruebas con ítems que hubieran sido aplicados en dicha oportunidad. Una vez consideradas distintas opciones sobre el tamaño de la muestra a utilizar para esta segunda comparación, se definió evaluar a 4800 niños con formas de prueba compuestas por ítems de 2002 e ítems de 2005 (los resultados de esta comparación se presentan en el siguiente capítulo).

Los dos subgrupos de casos (6000 en que se aplican pruebas para 2005 y 4800 en que se aplican pruebas que incluyen ítems de 2002) se sortearon conjuntamente, de manera de evitar sesgos diferenciales entre los grupos. Además de estos dos subconjuntos de alumnos que componen la muestra nacional, se evaluaron casos adicionales de Tiempo Completo (TC) y Contexto Sociocultural Crítico (CSCC). En el caso de TC se completó un censo de escuelas y en el caso de CSCC se incluyeron algunas escuelas de manera intencional.

La muestra nacional, por lo tanto, se compone de dos sub poblaciones: una que contesta cuadernillos a ser analizados según TRI y que conformará la línea de base para evaluaciones futuras, y otra que contesta cuadernillos a ser analizados con TCT y que permite comparar con la evaluación de 2002.

La cantidad de grupos y alumnos evaluados según tipo de muestra se puede observar en el siguiente cuadro:

Cuadro 3.1
Cantidad de grupos y de alumnos de la muestra

		Grupos	Alumnos
Muestra Nacional	Urbana común	180	5712
	TC	14	402
	TC y AI	2	50
	AI	76	2358
	CSCC y AI	2	58
	CSCC	37	1102
	Rural	3	29
	Autorizada	5	116
	Habilitada	38	948
	<i>Total Muestra Nacional</i>	<i>357</i>	<i>10775</i>
	Agregados	TC	101
TC (y AI)		6	183
CSCC (y AI)		2	57
CSCC		25	660
<i>Total agregados</i>		<i>134</i>	<i>3332</i>
Total		491	14107

Fuente: ANEP / Área de Evaluación de Aprendizajes – Evaluación 6to. año Educación Primaria 2005

Nota: TC= Tiempo Completo, CSCC= Contexto Sociocultural Crítico y AI= Áreas Integradas

Los siguientes dos cuadros de datos pretenden dar cuenta de la cantidad de alumnos que participaron efectivamente en las pruebas. En el primero de ellos, se presenta la cobertura lograda en las dos áreas de aprendizaje evaluadas según tipo de muestra, mientras que en el segundo se presenta la información referente a la respuesta lograda por las familias de los alumnos de sexto grado. Según se puede constatar a partir de la información incluida en los cuadros, las escuelas de Contexto Sociocultural Crítico que fueron agregadas a la muestra nacional mostraron una tasa de respuesta superior al resto, tanto en cuanto a la participación de los alumnos en las pruebas como en la respuesta dada por sus padres. Aún así, puede afirmarse que en el total nacional la cobertura de la información en ningún caso es inferior al 90% de los casos muestreados.

Cuadro 3.2
Porcentaje de respuesta según tipo de prueba y tipo de muestra

		Muestra			
		Nacional	Agregado TC	Agregado CSCC	Total
Lengua	TRI	92	91	98	92
	TCT	92	91		92
	Total	92	91	98	92
Matemática	TRI	90	90	98	91
	TCT	91	90		91
	Total	90	90	98	91

Fuente: ANEP / Área de Evaluación de Aprendizajes – Evaluación 6to. año Educación Primaria 2005

Cuadro 3.3
Porcentaje de respuesta a la encuesta a las familias

	Muestra		
	Nacional	Agregado	Total
Nacional	91	--	91
TC	86	96	94
CSCC	89	97	92
AI	94	--	95

Fuente: ANEP / Área de Evaluación de Aprendizajes – Evaluación 6to. año Educación Primaria 2005

3.2 Características sociales de la población evaluada

El Cuadro 3.4 presenta la evolución de un conjunto de indicadores sociales que pretenden ilustrar sobre la composición social de los alumnos de sexto grado, de acuerdo a los datos recabados en las cuatro instancias de evaluaciones de aprendizaje realizadas. Tomando como referencia a los indicadores involucrados en la elaboración del índice del contexto sociocultural, se presenta la evolución de los tramos considerados en cada uno de ellos: el porcentaje de madres con primaria completa como máximo nivel alcanzado y el porcentaje de madres con secundaria completa o más, la evolución del índice socioeducativo que resulta de la resta de los dos anteriores, el porcentaje de alumnos que viven en hogares con bajo equipamiento, con alto equipamiento, el índice socioeconómico que resulta de la resta de estos últimos dos, y finalmente el porcentaje de alumnos que viven en hogares con problemas de hacinamiento¹.

Cuadro 3.4
Evolución de algunos indicadores sociales seleccionados. Años 1996, 1999, 2002 y 2005

Porcentaje de:	TOTAL PAÍS			
	1996	1999	2002	2005
	Madres con primaria completa como máximo nivel educativo	38,5	36,0	35,9
Madres con educación secundaria completa o más	25,5	26,7	25,4	25,1
Valor del índice de saldo educativo	-13,0	-9,3	-10,5	-10,9
Alumnos en hogares con bajo equipamiento	32,4	18,5	19,1	36,0
Alumnos en hogares con alto equipamiento	22,0	29,1	32,7	21,6
Valor del índice de saldo socioeconómico	-10,4	10,6	13,6	-14,5
Alumnos en hogares con hacinamiento	17,3	16,0	16,0	16,5

Fuente: ANEP / Área de Evaluación de Aprendizajes – Evaluación 6to. año Educación Primaria 2005

Cuando se observa el perfil del alumnado de los sextos años de primaria según la educación de sus madres importa destacar que prácticamente no se observan cambios a lo largo de los años analizados. Si bien el porcentaje de madres con baja educación sufrió un leve

¹ Se considera que un hogar presenta esta característica cuando viven más de dos personas por habitación, sin considerar baños ni cocina.

descenso entre 1996 y 1999, a partir de entonces se ha mantenido prácticamente constante. Tampoco se observan cambios importantes en cuanto a las madres más educadas (a pesar del leve incremento registrado en el año 1999) y por lo tanto el índice socioeducativo se mantiene en los niveles anteriores.

En cuanto a la evolución de los indicadores de equipamiento de los hogares de los alumnos, es necesario hacer algunas precisiones. En primer término, recordar que en 2005 se amplió la cantidad de ítems que integran las categorías de referencia. En segundo lugar, es importante tener en cuenta que a partir del año 2002 los hogares de los alumnos pueden estar sufriendo los efectos de la crisis económica de dicho año, y por lo tanto haber disminuido el poder de recuperación o reposición de los elementos con los que contaban. Ambas precisiones se encuentran en la base de los cambios que se observan en el seguimiento de cada uno de los indicadores presentados². En efecto, luego de una marcada recuperación del índice socioeconómico en 1999, no se observan grandes cambios en el siguiente período y un gran deterioro en el último año analizado en el cuadro 3.4, probablemente relacionado con los cambios mencionados y con la actualización del índice socioeconómico, en particular sobre la conformación de las categorías de equipamiento.

Más allá de estos cambios mencionados, el nivel de hacinamiento de los hogares de los alumnos da cuenta de la poca variabilidad que ha tenido la población de sextos años a lo largo de la serie analizada. Al tratarse de un indicador que ha sido probado en diversos estudios, ya sea en su carácter predictivo como explicativo de ciertos comportamientos de riesgo, puede afirmarse que su estabilidad da cuenta de que a los sextos años de educación primaria continúan llegando, a lo largo de los últimos diez años, estudiantes similares en términos de su configuración social.

Si bien estos elementos dan cuenta de algunas características sociales de la población de sextos años, no informan cómo se distribuyen a lo largo del contexto sociocultural generado. Para ello, debemos remitirnos al Cuadro 3.5, que informa sobre el volumen de niños que pertenece a cada uno de los contextos y acerca del comportamiento de los indicadores seleccionados para cada una de las categorías sociales conformadas para el año 2005.

En lo que refiere al primer punto, se observa que aproximadamente un 40% de la población se ubica entre el contexto Muy desfavorable y Desfavorable, una cuarta parte en el contexto Medio, un quinto de ellos en el contexto Favorable y algo menos del 15% en la situación más favorable de la distribución.

En cuanto al comportamiento de las variables sociales relacionadas a cada uno de los contextos puede destacarse que todas ellas se ordenan de acuerdo a lo esperado: a medida que mejora el contexto sociocultural aumentan los porcentajes de alumnos que presentan situaciones favorables (por ejemplo, porcentaje de madres con alta educación), o visto desde los indicadores asociados a las “malas configuraciones” se observa un decrecimiento cuando mejora dicho contexto. En este sentido, lo más llamativo parece ser la gran “distancia” existente entre los valores de la categoría Muy favorable y el resto en los indicadores del primer tipo: madres altamente educadas y hogares con alto equipamiento.

² Ver Anexo 1

Cuadro 3.5
Algunos indicadores sociales seleccionados según contexto sociocultural

	Contexto sociocultural					Total
	Muy favorable	Favorable	Medio	Desfavorable	Muy desfavorable	
Cantidad	1572	2126	2786	1744	2614	10842
Porcentaje	14,5	19,6	25,7	16,1	24,1	100,0
% de madres con baja educación	4,6	19,8	36,1	44,8	62,5	36,1
% de madres con alta educación	69,9	35,5	17,2	10,5	5,4	24,5
Saldo de educación en el grupo	65,3	15,8	-18,8	-34,2	-57,1	-11,6
% de niños con bajo equipamiento	2,6	17,7	33,6	49,8	66,4	36,5
% de niños con alto equipamiento	66,4	28,3	14,5	7,1	2,8	20,7
Saldo de equipamiento en el grupo	63,8	10,6	-19,1	-42,7	-63,6	-15,8
% de niños con hacinamiento	2,5	9,2	14,8	23,4	28,0	16,5
% de hogares que o no tienen baño o lo comparten con otro hogar	0,4	2,0	3,5	5,8	6,9	3,9
% de familias inscritas y aceptadas en el Plan de Emergencias	0,8	3,3	5,7	9,8	12,5	6,8
% de familias que vive en un asentamiento	0,7	5,3	9,5	15,6	18,2	10,5

Fuente: ANEP / Área de Evaluación de Aprendizajes – Evaluación 6to. año Educación Primaria 2005

Finalmente, se presentan los indicadores que no están asociados a la creación del contexto sociocultural. En cuanto a ellos, que pretenden dar más luz a las características sociales del alumnado según el contexto en el cual se encuentren, se destaca el alto porcentaje de niños del contexto Muy Desfavorable que presentan problemas de hacinamiento o que viven en asentamientos irregulares³, aunque en el resto de los contextos también se observen situaciones preocupantes al punto de existir casos de familias inscritas y aceptadas en el Plan de Emergencia implementado por el actual gobierno.

3.3 Resultados académicos por contexto sociocultural

En esta sección se presentan los resultados nacionales obtenidos por los alumnos en Matemática y Lenguaje. A diferencia del tipo de evaluación que se venía realizando desde esta unidad, y tal como se aclarara en los capítulos precedentes, el modo en que hemos trabajado con la Teoría de respuesta al ítem permite hablar de *niveles de desempeño* y no de suficiencia, razón por la cual los resultados se presentan de acuerdo a ellos. Adicionalmente, y de acuerdo a la vasta literatura que demuestra el fuerte impacto del entorno social sobre el proceso de enseñanza-aprendizaje, los resultados también se presentan tomando en consideración el contexto sociocultural del grupo de referencia del alumno.

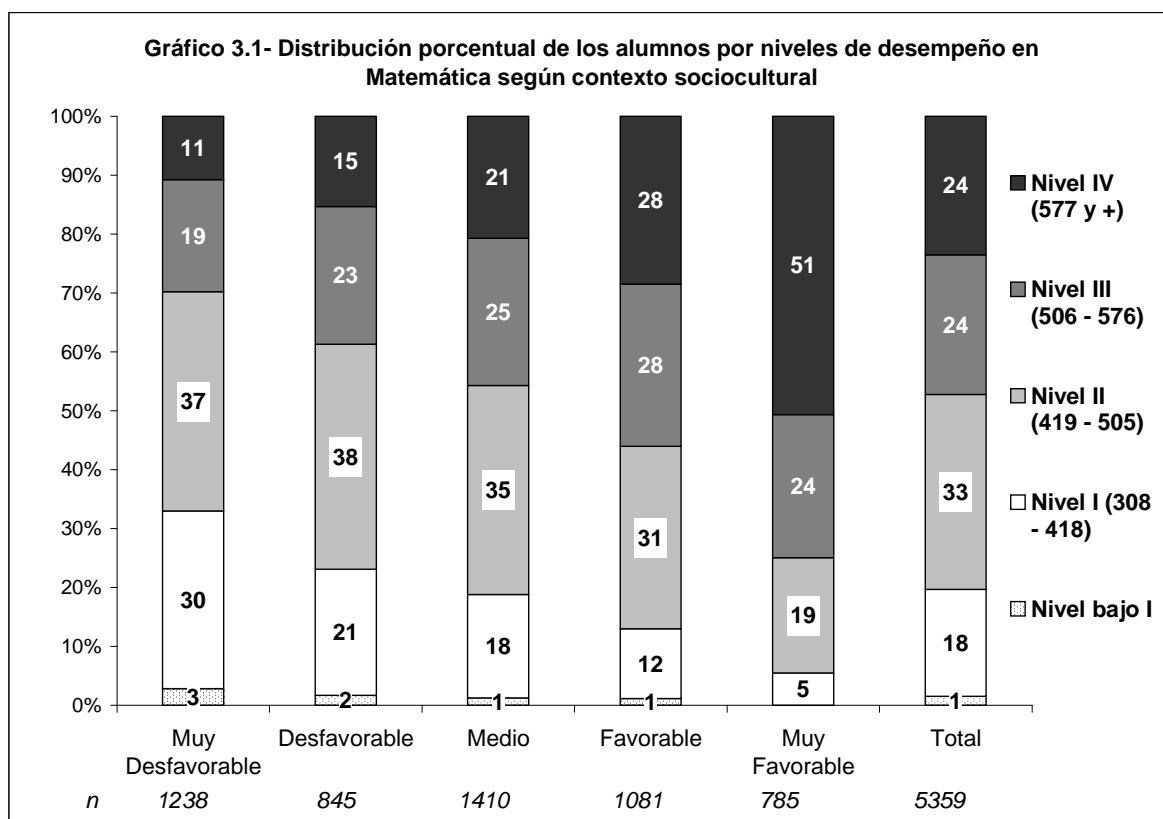
3.3.1 Matemática

En lo que refiere al área de Matemática, se observa que menos de la mitad de los alumnos (48%) ha logrado ubicarse en los dos tramos más altos de la escala (Niveles III y IV),

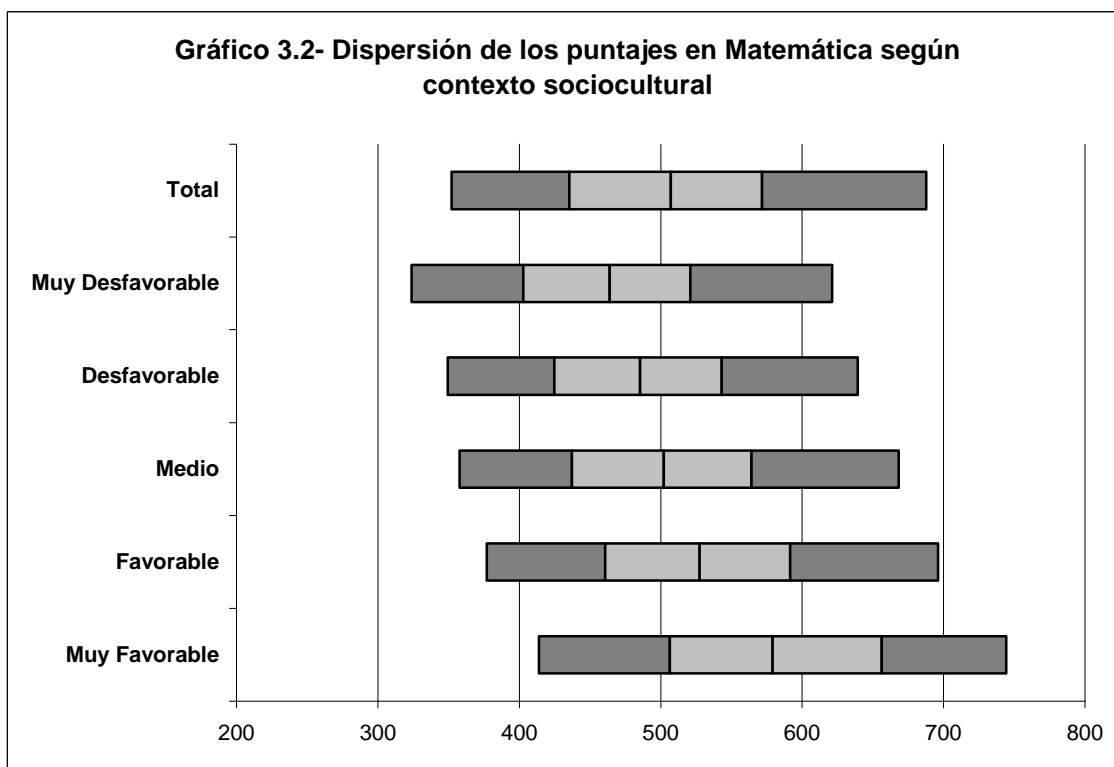
³ Según la declaración de los padres.

mientras que la tercera parte logra resultados equivalentes al Nivel II, un quinto de los alumnos al Nivel I y solamente el 1% no alcanza las habilidades descriptas para el Nivel I.

Al igual que en las anteriores evaluaciones de aprendizaje, los resultados varían de acuerdo al contexto sociocultural del grupo del alumno. Como se observa claramente, los alumnos del contexto Muy desfavorable que se ubican en el primer nivel de desempeño son seis veces más que los del Muy favorable en dicho nivel. En otras palabras, mientras una tercera parte de los que poseen peor situación social no logra superar el primer nivel de desempeño, tan solo el 5% de los alumnos más favorecidos logra un resultado similar. Inversamente, el porcentaje de alumnos del contexto Muy desfavorable que logra ubicarse en los Niveles III o IV de desempeño es del 30%, mientras que 3 de cada 4 alumnos del contexto Muy Favorable se ubica en estos niveles.



Fuente: ANEP / Área de Evaluación de Aprendizajes – Evaluación 6to. año Educación Primaria 2005



Fuente: ANEP / Área de Evaluación de Aprendizajes – Evaluación 6to. año Educación Primaria 2005

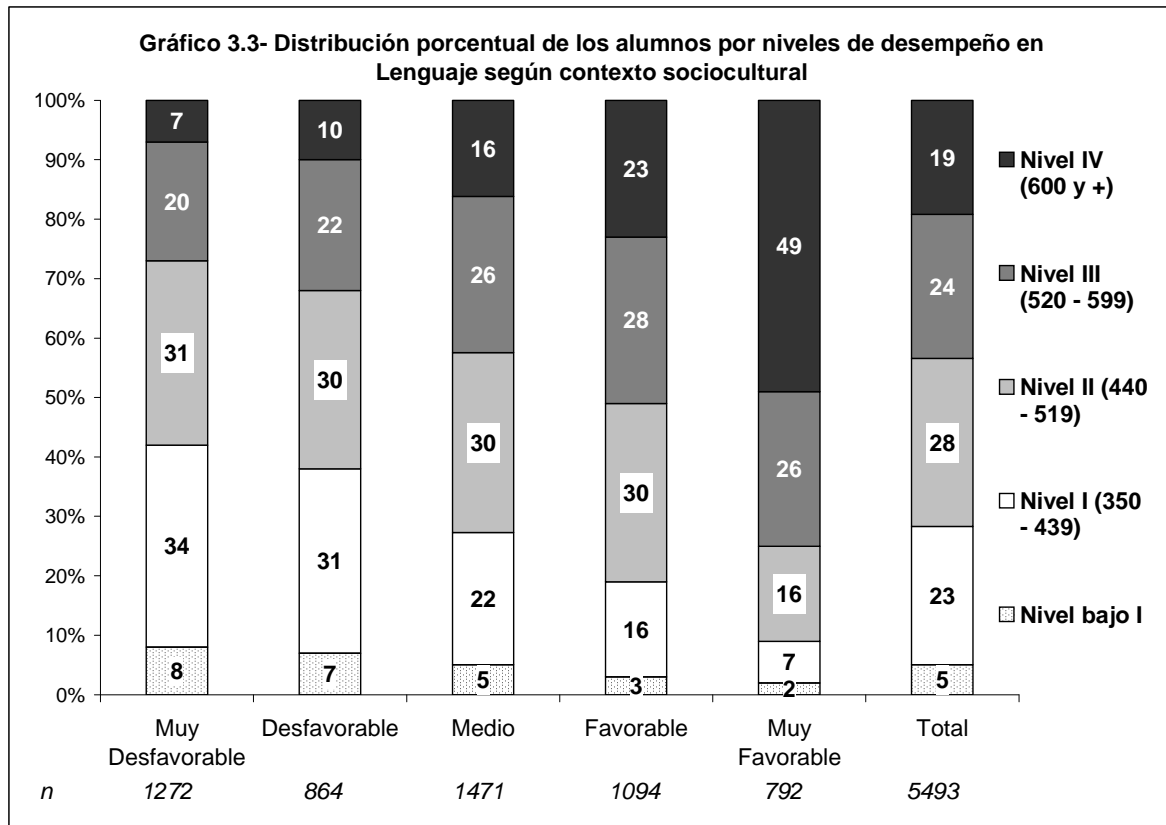
Más allá de la disparidad de resultados encontrada en los desempeños de los alumnos según pertenezcan a uno u otro contexto sociocultural, importa brindar una mirada adicional que observe la disparidad *al interior* de cada uno de ellos.

El Gráfico 3.2 pretende mostrar qué tan homogéneos o heterogéneos pueden ser los desempeños en cada uno de los contextos, utilizando en cada barra 5 valores: en el centro de ella el promedio del puntaje obtenido por los alumnos pertenecientes a ese contexto; a sus costados los puntajes obtenidos por el 25% de los alumnos de menor rendimiento (izquierda) y por el 25% con mejores resultados (derecha). En el extremo izquierdo el puntaje obtenido por el 5% de alumnos con peor desempeño y en el derecho el puntaje del 5% de mejor desempeño.

Al igual que se observó en el gráfico anterior, la evidencia indica que a medida que mejora el contexto los alumnos logran mejores desempeños. Cabe destacar que el puntaje obtenido por el 25% con peor desempeño en el contexto Muy desfavorable es aún algo menor al que obtiene el 5% del contexto Muy favorable. Finalmente se observa que, en términos generales, la desigualdad en los aprendizajes no parece ser mayor en un contexto que en otro.

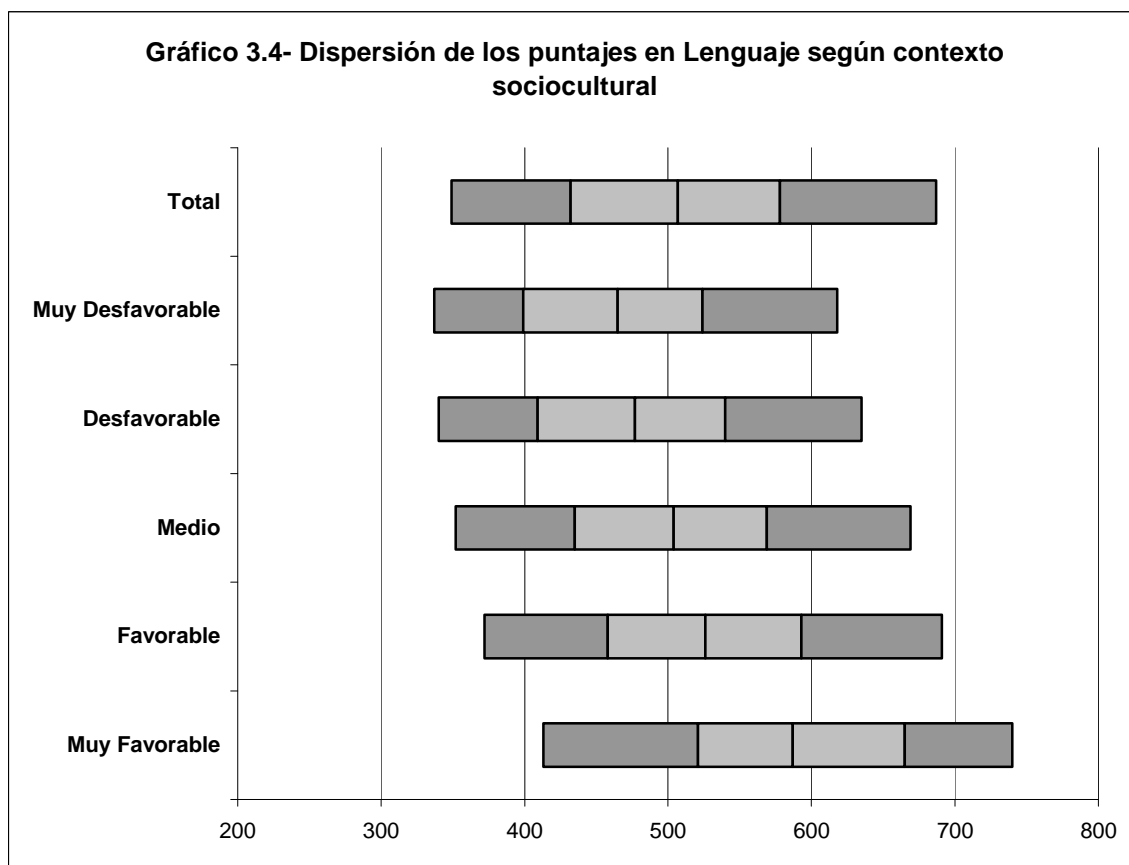
3.3.2 Lenguaje

En lo que refiere al área de Lenguaje, también se observa que menos de la mitad de los alumnos (43%) ha logrado ubicarse en los dos tramos más altos de la escala (Niveles III y IV), mientras que tan solo el 28% logra resultados equivalentes al segundo nivel, y uno de cada cuatro alumnos (25%) logra competencias que se ubican en el Nivel I. Sin embargo, es importante aclarar que el 5% de los alumnos no logra resolver las actividades que definen el nivel más básico de la prueba.



Fuente: ANEP / Área de Evaluación de Aprendizajes – Evaluación 6to. año Educación Primaria 2005

Al igual que en el caso de los resultados de Matemática, se observan diferencias en los desempeños de los estudiantes según el contexto sociocultural al que pertenezcan. En el caso del Lenguaje, la mitad de los alumnos del contexto Muy favorable se ubica en el nivel de desempeño más alto, mientras el mismo rendimiento lo logra tan solo el 7% de los alumnos de contextos más desfavorecidos. En el mismo sentido, se pueden agrupar y observar los dos niveles mejores de rendimiento: lo alcanza el 75% de los estudiantes del muy favorable contra tan solo el 27% del Muy desfavorable.



Fuente: ANEP / Área de Evaluación de Aprendizajes – Evaluación 6to. año Educación Primaria 2005

En cuanto a la disparidad de los desempeños estudiantiles según pertenezcan a uno u otro contexto sociocultural, el Gráfico 3.4 da cuenta de un leve aumento de la dispersión de resultados a medida que mejora la situación social de la población. En efecto, es en el contexto Muy favorable donde aparece la mayor longitud de la barra, y por tanto una desigualdad mayor en los resultados obtenidos por el conjunto de estudiantes de dicho contexto.

Por otra parte, el patrón que parece estar mostrando el gráfico es que si bien el 5% de estudiantes de peor rendimiento en cada contexto mejora levemente sus resultados a medida que mejora el contexto, al llegar a la mejor situación social ese “piso” del 5% logra incrementar bastante más sus resultados que lo que lo venían haciendo el resto. Al igual que en Matemática, vemos que el resultado de este 5% inferior del contexto Muy favorable es mayor al resultado del 25% inferior de los contextos Muy desfavorable y Desfavorable. En todo caso, la significación de las diferencias marcadas en este punto deberá ser analizada una vez que se calculen los márgenes de error de la muestra.

Capítulo 4

Comparación entre las muestras nacionales de 2005 y 2002

4.1 Diseño muestral y de pruebas

El diseño muestral y de pruebas utilizado en 2005 permite comparar los resultados de un subconjunto de alumnos con los obtenidos en la muestra nacional de 2002. Para ello, algunos cuadernillos de prueba incluyeron los mismos ítems que se habían evaluado en 2002. Dentro de cada uno de los grupos que conformaron la muestra, algunos alumnos, seleccionados aleatoriamente, contestaron cuadernillos con los mismos ítems que en 2002.

En el caso de Lengua, hubo tres cuadernillos que abordaban los ítems correspondientes al texto explicativo utilizado en 2002, y otros tres cuadernillos que presentaban el texto narrativo del año 2002 y los ítems asociados a él.

Cada uno de estos conjuntos de ítems se combinó en cada cuadernillo con uno de los bloques nuevos diseñados para la evaluación de 2005. En total, la prueba que contestaba cada niño tuvo 24 ítems, 12 del año 2002 y 12 del 2005. Los ítems del 2002 siempre se ubicaron en el mismo lugar del cuadernillo en el que estaban en la prueba de 2002.

En el caso de Matemática, se hizo una selección de 12 ítems del total de los 24 aplicados en 2002.

Se crearon cuatro cuadernillos que contenían este subconjunto de ítems. En dos de ellos estos ítems se encontraban en primer lugar y en los dos restantes una vez se ubicaron en segundo lugar y otra en tercero.

Estos 12 ítems fueron combinados con los ocho bloques de ítems nuevos que conformaron la prueba de 2005. En total, los niños que contestaron cuadernillos de matemática con ítems de 2002 respondieron 32 ítems, 12 de 2002 y 20 de la prueba de 2005 correspondientes a dos bloques.

Se consideró que 4800 casos era una cantidad razonablemente buena para establecer las comparaciones con márgenes de error aceptables. Como los textos de la prueba de Lenguaje fueron incluidos en cuadernillos diferentes, los resultados de cada uno de los textos se basan en un tamaño muestral de aproximadamente 2200 casos.

En el caso de Matemática, como existió un único conjunto de ítems, sus resultados se basan en un tamaño de muestra de aproximadamente 4800 casos.

4.2 Cómo interpretar los resultados

En el punto 4.3 se presentan los resultados para el texto explicativo de la prueba de lengua, en el 4.4 los del texto narrativo y en el 4.5 los correspondientes a los ítems de Matemática.

En las gráficas que muestran los resultados de cada área, en primer lugar presentamos el porcentaje promedio de respuestas correctas a todos los ítems para cada uno de los contextos y el total de la muestra. Luego se presenta el porcentaje de respuestas correctas a cada ítem, para el total de la muestra y para cada contexto sociocultural.

En todas las gráficas, la ubicación de cada punto en el eje de las “X” corresponde al porcentaje de respuestas correctas en 2002. Y la ubicación en el eje de las “Y” corresponde al porcentaje de respuestas correctas a los mismos ítems en el 2005.

Las gráficas incluyen también una línea que las atraviesa diagonalmente entre el vértice inferior izquierdo y el superior derecho. Ésta pretende operar como una ayuda para la lectura del gráfico. Si un punto queda ubicado sobre la línea, está indicando que se obtuvo exactamente el mismo porcentaje promedio de respuestas correctas en 2002 que en 2005.

Los puntos ubicados por debajo de la línea son aquellos en los que el porcentaje de respuestas correctas fue mayor en la evaluación de 2002 que en la de 2005. Mientras que si el punto se ubica por encima de la línea se debe a la situación contraria: resultados más altos en 2005 que en 2002. Cuanto más se alejan los puntos de esta línea diagonal, mayor es la diferencia entre los resultados de ambos años.

Cuando se presenta la información por contexto sociocultural debe considerarse que es el relativo a cada evaluación. Esto es, los datos de cada evaluación se presentan según la clasificación de contexto de cada medición. En el Anexo se presenta la comparación entre los resultados de ambas evaluaciones, pero clasificando a los grupos de 2005 según la definición de 2002¹.

Al momento de observar las diferencias entre 2002 y 2005 es importante tener en cuenta que, estadísticamente, hay ciertas diferencias que deben considerarse como resultados iguales, ya que el simple hecho de trabajar con muestras y no con censos añade a cada resultado un margen de error. Por lo tanto, cada punto de la gráfica tiene un intervalo de confianza, esto es, el porcentaje reportado varía entre +/- un cierto grado de error que está con relación al número de casos de la muestra.

Por este último aspecto, los resultados de 2002 (n=5327) tienen menos error que los de 2005 (n=4800). Y entre estos últimos el error de Matemática (n=4800) es menor que el de los textos de Lengua (n=2200 para cada uno).

Para esta primera publicación de resultados, como aún no se han calculado los errores del porcentaje de respuestas correctas para cada ítem en 2002 y en 2005, presentamos como referencia el error muestral de la evaluación 2002.

A pesar de ello, debe tenerse en cuenta que:

- a. éstos fueron calculados para el “porcentaje de alumnos suficientes” y que ahora estamos comparando ítem a ítem, por lo cual es probable que los errores varíen para el 2002.

¹ No es posible hacerlo a la inversa porque en 2005 se incluyeron ítems de equipamiento que no fueron relevados en 2002.

- b. los errores de 2005 probablemente serán más grandes que los de 2002, ya que los tamaños muestrales son más pequeños.

Tabla 4.1
Márgenes de error en 2002 y cantidad de casos en 2002 y 2005

	Error en 2002		Casos en Matemática		Casos en Lengua		
	Lengua	Matemática	2002	2005	2002	Texto explicativo en 2005	Texto narrativo en 2005
Total	3,0	3,4	5342	4333	5321	2210	2193
Muy favorable	5,0	8,6	786	648	780	312	329
Favorable	6,2	8,1	918	882	922	449	435
Medio	6,7	7,3	925	1107	924	586	547
Desfavorable	7,0	7,4	1140	685	1120	348	338
Muy desfavorable	6,9	6,2	1570	1011	1575	515	517

Como puede observarse en la Tabla 4.1, solo en los contextos Muy favorable, Favorable y Medio, para la prueba de Matemática, el tamaño muestral de 2005 es similar al de 2002. En el resto de las situaciones, el número de casos es claramente inferior en 2005 que en 2002, por lo que los errores muestrales serán mayores.

Los comentarios sobre cómo interpretar las gráficas son válidos para todas las que se presentan a continuación.

4.3 Comparación entre 2002 y 2005 del porcentaje de respuestas correctas a cada ítem del texto explicativo de la prueba de 2002

La Gráfica 4.1 presenta el porcentaje promedio de respuestas correctas a los ítems del texto explicativo para cada contexto sociocultural y el total de la población.

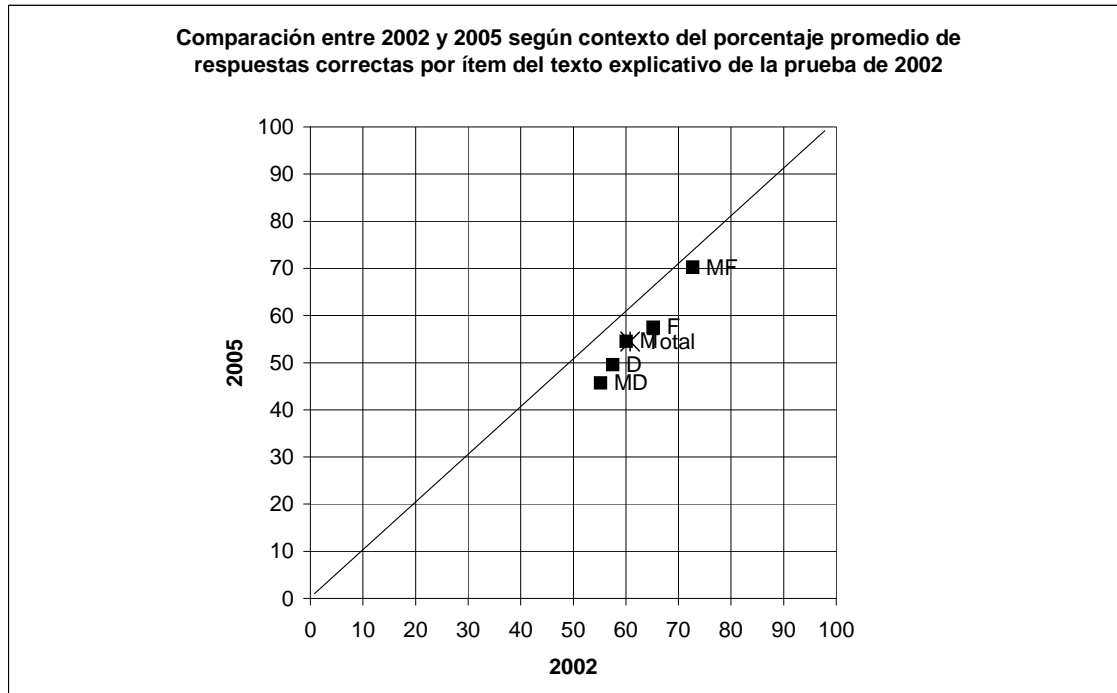
Cada uno de los puntos surge de promediar, para cada contexto, el porcentaje de respuestas correctas a cada ítem.

Allí observamos que los resultados obtenidos para el texto explicativo fueron algo inferiores en 2005 que en 2002, aunque debemos analizar en el siguiente informe si estas diferencias son significativas. También se observa cómo el porcentaje promedio de respuestas correctas aumenta con el contexto sociocultural.

Profundizando en las diferencias entre ambas evaluaciones, podemos ver que en el contexto Muy favorable se trata de una pequeña diferencia (3%), mientras que en los contextos Favorable (8%), Medio (6%), Desfavorable (8%) y Muy desfavorable (9%), la magnitud de la diferencia tiende a crecer².

² Es importante señalar que cuando se hace referencia a diferencias entre las evaluaciones no estamos hablando de diferencias estadísticamente significativas, ya que aún no se han calculado los márgenes de error. Estos se incluirán en una futura publicación.

Gráfico 4.1

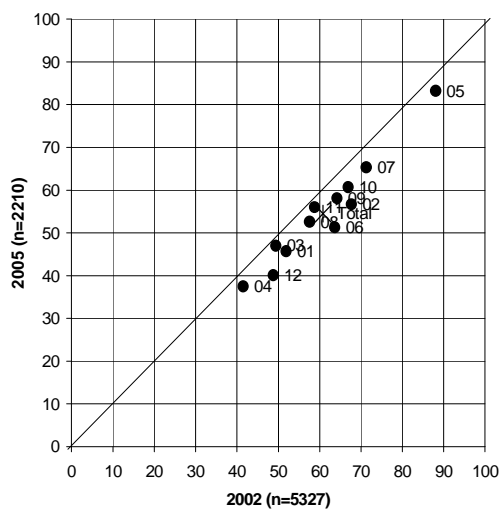


En el Gráfico 4.2, en el primer panel cada punto representa las respuestas correctas a cada ítem en 2002 y 2005. Además, con una cruz se señala el porcentaje promedio de respuestas correctas a estos ítems. Este total es el mismo que el de la gráfica anterior.

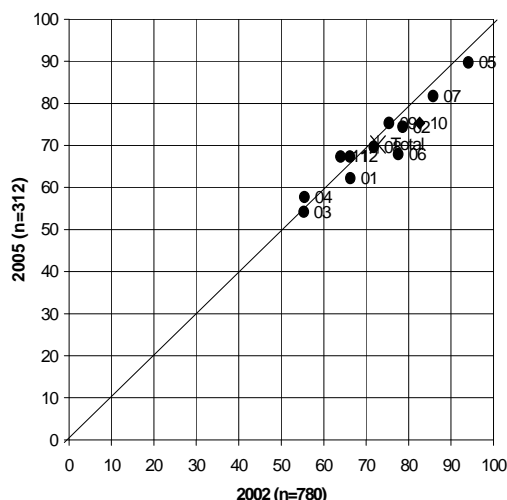
Como podemos ver, si bien todos los ítems se encuentran por debajo de la diagonal, esto es, sus resultados son algo inferiores en 2005 que en 2002, hay variaciones en las respuestas correctas a los ítems de ambas evaluaciones.

Los cinco gráficos muestran las diferencias en los porcentajes de respuestas correctas a los ítems comparables en ambas evaluaciones, por contexto sociocultural. La tendencia general muestra que los porcentajes de respuestas correctas fueron similares o levemente inferiores a los del 2002: casi iguales entre los contextos Medio, Favorable y Muy favorable, y levemente inferiores en los dos contextos desfavorables. En promedio, las diferencias en estos dos contextos son 8 y 9 por ciento inferiores en el 2005 frente al 2002. Cabe nuevamente hacer la salvedad de que, hasta que no se estimen los márgenes de error, no se puede valorar estadísticamente si las diferencias son significativas.

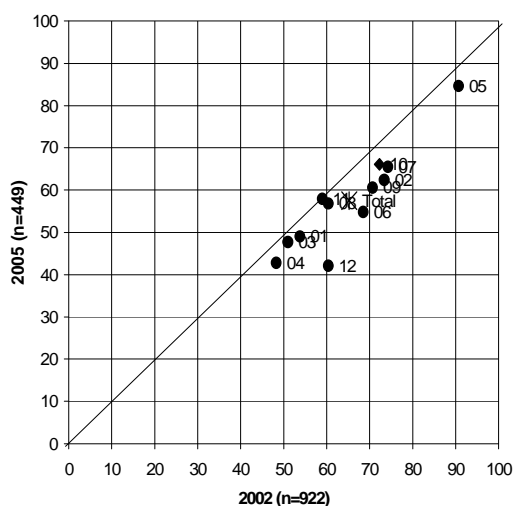
Gráfico 4.2
**Comparación (2002-2005) del
 porcentaje de respuestas correctas a los ítems de Texto Explicativo**



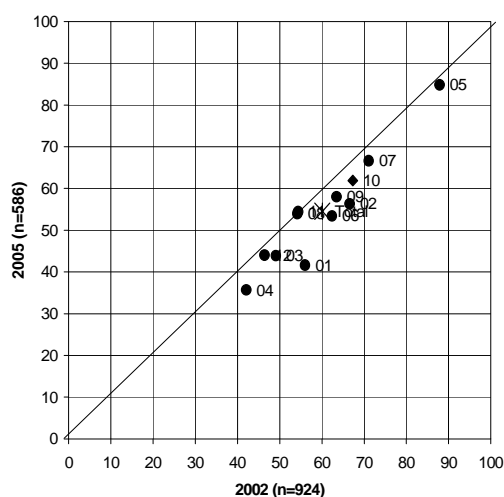
Muestra Nacional



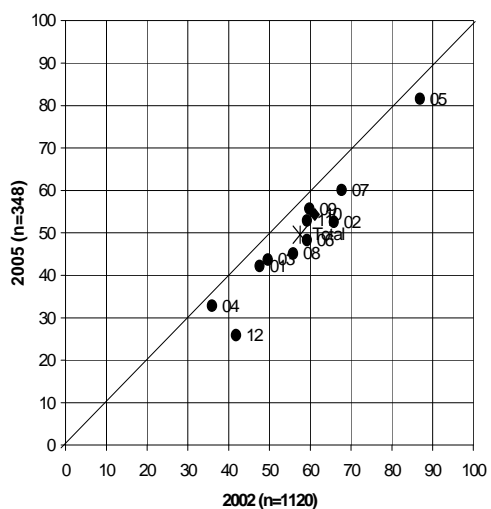
Contexto Muy Favorable



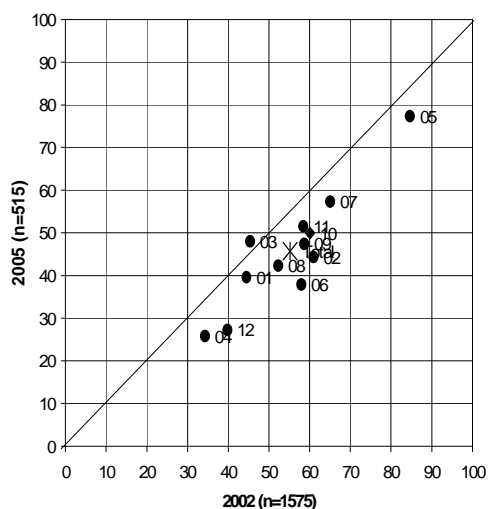
Contexto Favorable



Contexto Medio



Contexto Desfavorable



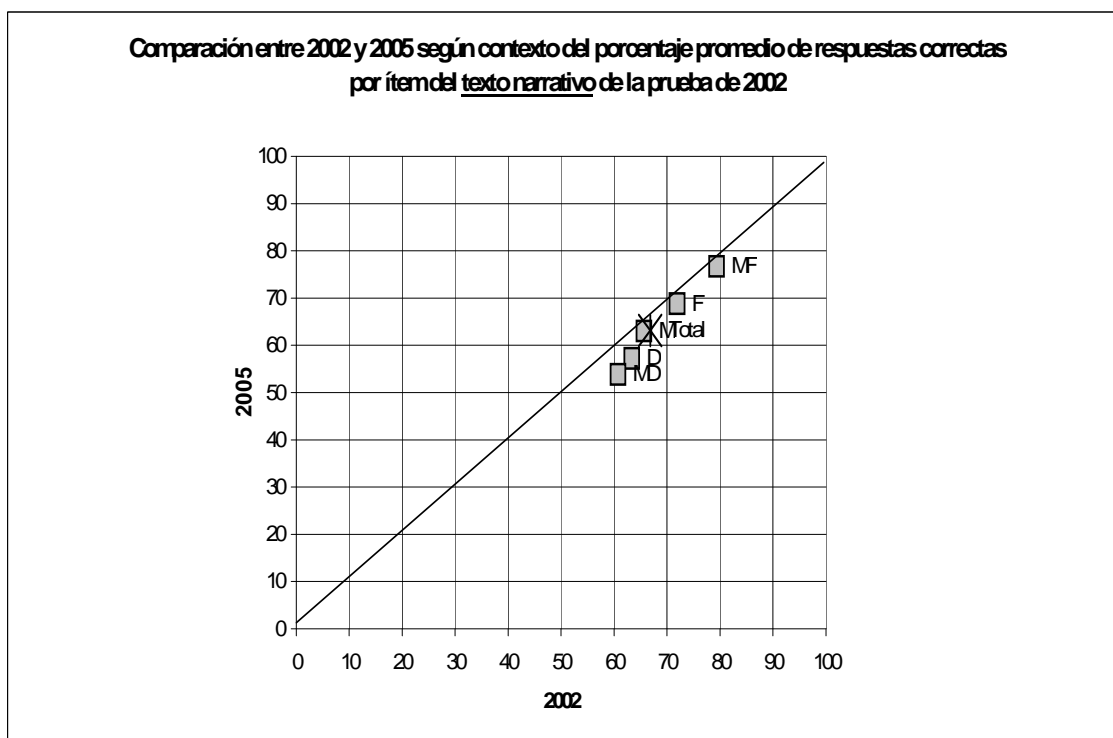
Contexto Muy Desfavorable

4.4 Comparación entre 2002 y 2005 del porcentaje de respuestas correctas a cada ítem del texto narrativo de la prueba de 2002

En la Gráfica 4.3 se presenta el porcentaje promedio de respuestas correctas a los ítems del texto narrativo para cada contexto sociocultural y el total de la muestra.

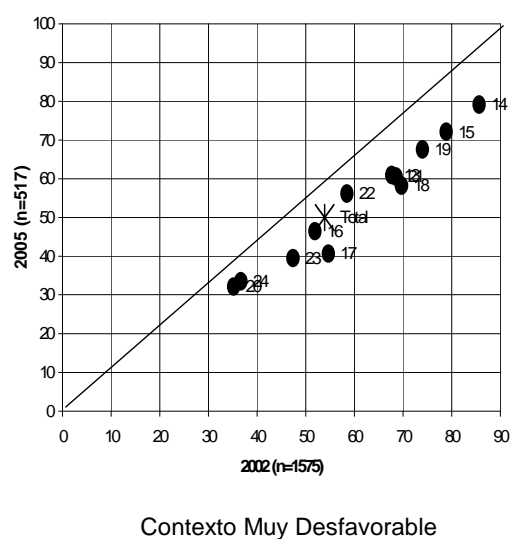
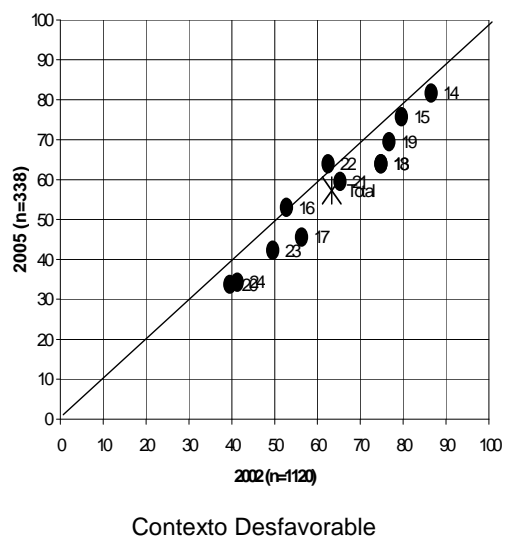
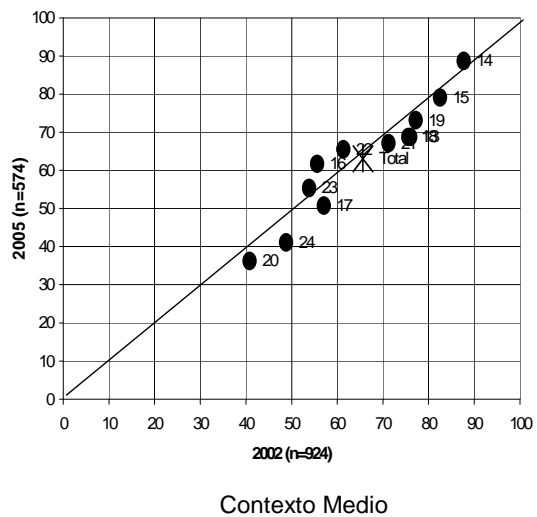
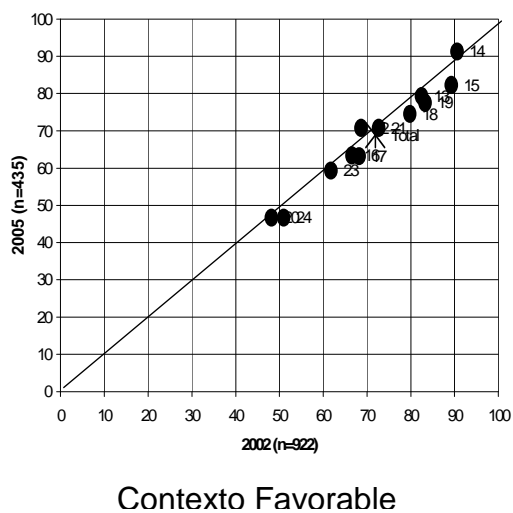
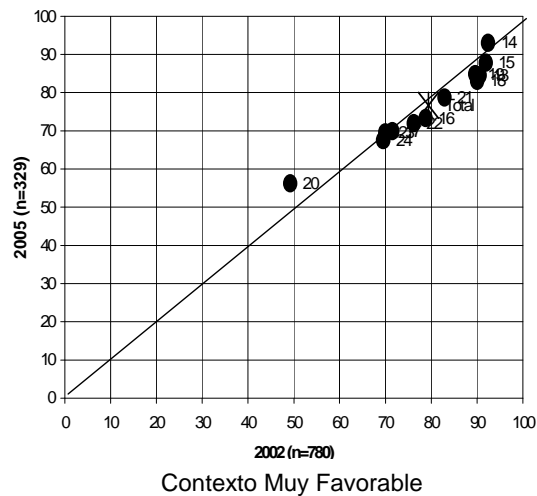
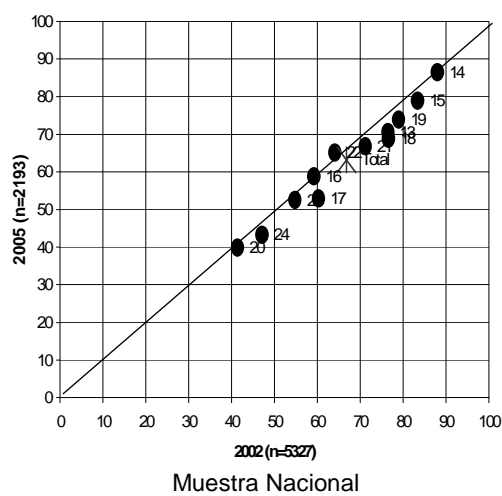
Al igual que en el caso del texto explicativo, en este caso se observa que los resultados de los alumnos de los contextos Muy favorable, Favorable y Medio son muy similares entre los dos años evaluados. Asimismo, en el caso del contexto Desfavorable y Muy desfavorable aumentan levemente las diferencias entre 2005 y 2002 (6% y 7% respectivamente), posiblemente aún dentro de los márgenes de error de la muestra.

Gráfico 4.3



Al considerar el porcentaje de respuestas correctas a cada ítem del texto narrativo para toda la muestra encontramos que si bien los resultados de 2005 son algo inferiores a los de 2002, las diferencias entre ellos son muy pequeñas. Solamente en tres ítems la diferencia porcentual es mayor a 5% (primer panel del Gráfico 4.4).

Gráfico 4.4
Comparación (2002-2005) del
porcentaje de respuestas correctas a los ítems de Texto Narrativo



4.5 Comparación entre 2002 y 2005 del porcentaje de respuestas correctas a algunos ítems de Matemática de la prueba de 2002

Los resultados de los doce ítems de 2002 que se incluyeron en la prueba de Matemática del 2005 muestran que casi no hay diferencia en el porcentaje promedio de respuestas correctas para los alumnos de los contextos Muy favorable (2%) y Favorable (3%), pero sí aumenta la diferencia en el caso de los contextos Medio (6%), Desfavorable (5%) y Muy desfavorable (7%).

Gráfico 4.5

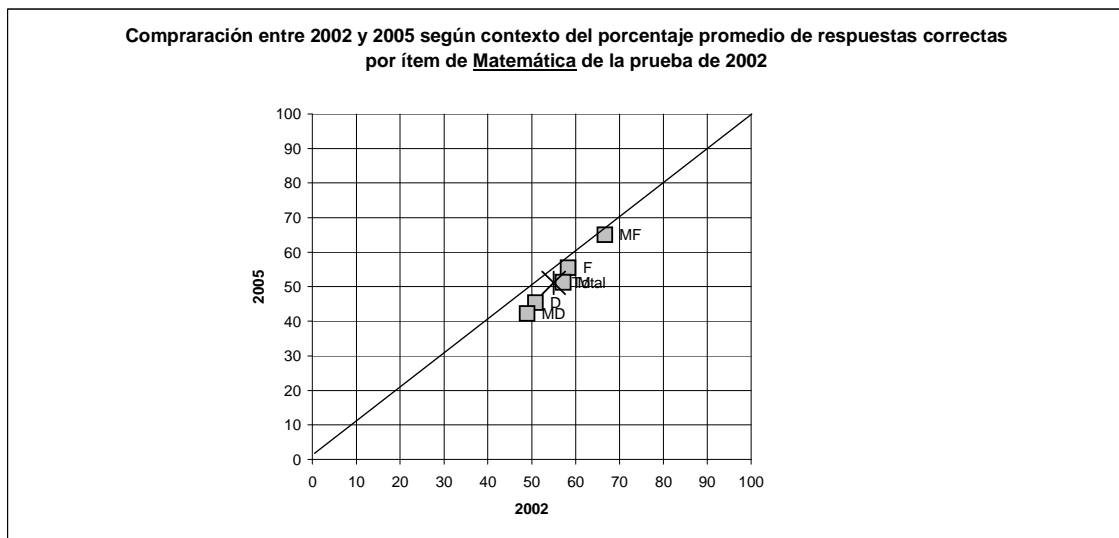
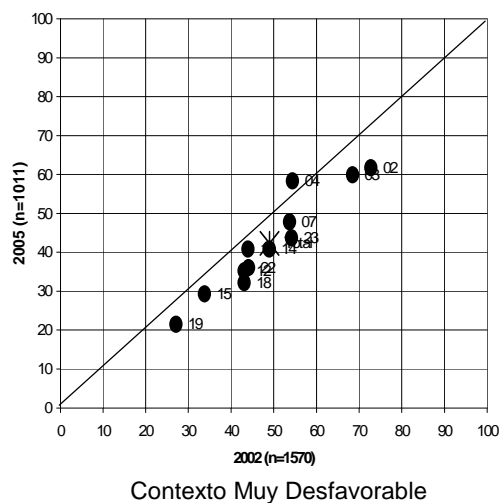
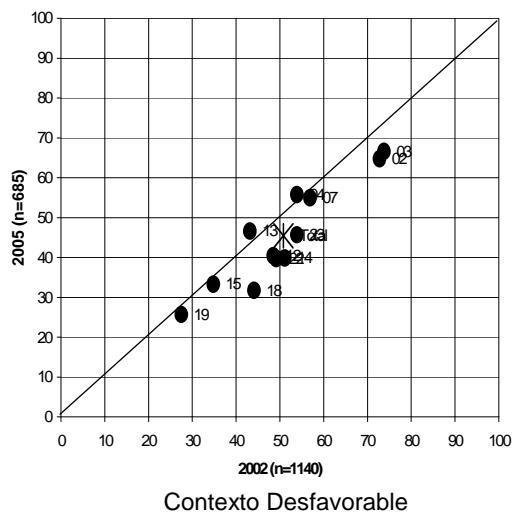
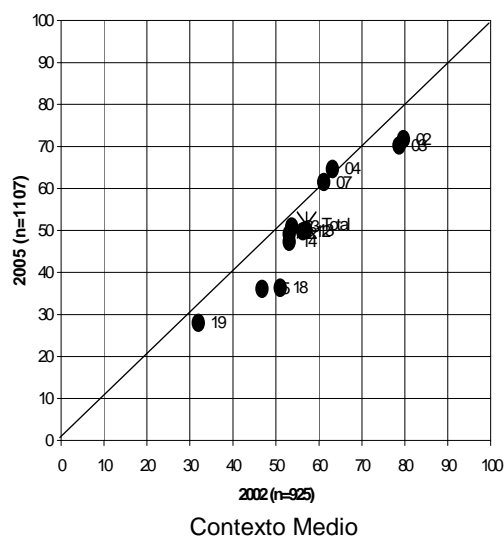
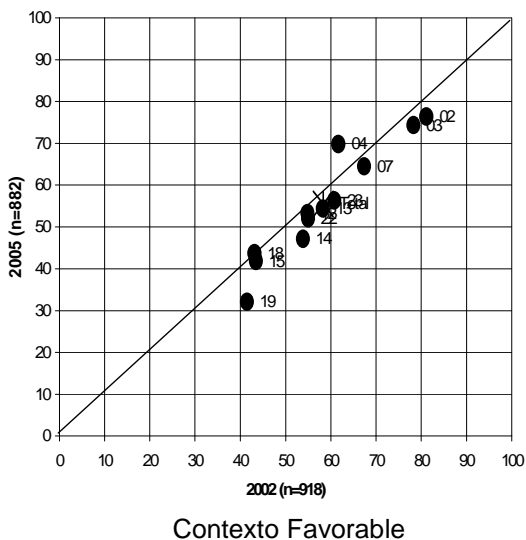
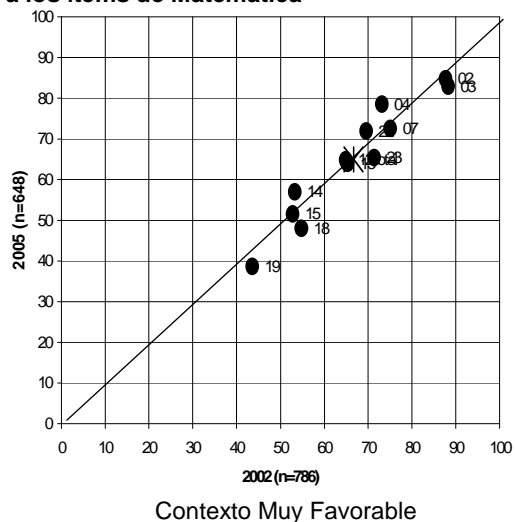
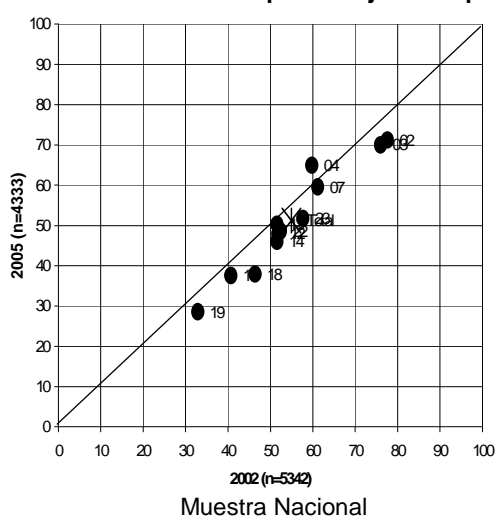


Gráfico 4.6
 Comparación (2002-2005) del
 porcentaje de respuestas correctas a los ítems de Matemática



Una posible explicación a las diferencias entre las evaluaciones (que cabe para la prueba de Matemática pero no para la de Lengua) es la diferencia en el largo de la prueba (24 a 32 ítems). Un indicador de si este factor afectó o no los resultados de 2005 es ver si el porcentaje de respuestas correctas a los ítems varía según su posición en el cuadernillo.

Si el porcentaje de respuestas correctas desciende significativamente a medida que el bloque se encuentra más al final del cuadernillo, podremos decir que este aspecto representa una diferencia relevante con la prueba de 2002 y que probablemente explique la diferencia en los resultados de las evaluaciones.

En la Tabla 4.2 presentamos el porcentaje de respuestas correctas a cada ítem de 2002 en 2005, según su lugar en el cuadernillo. En las dos últimas columnas se indica si la prueba de hipótesis χ^2 es significativa o no³. En caso de que resulte significativa podremos decir que los resultados varían según la posición del ítem en la prueba.

Tabla 4.2
Porcentaje de respuestas correctas en 2005 a cada ítem de 2002,
según la posición del bloque en el cuadernillo

	Posición del bloque con ítems de 2002			Total	χ^2	
	1°	2°	3°		Sin dicotomizar	Dicotomizado
It 1	72%	75%	71%	73%	--	
It 2	72%	74%	70%	72%	--	
It 3	68%	67%	62%	65%	sig.*	sig.*
It 4	63%	66%	61%	63%	sig.*	sig.*
It 5	48%	48%	48%	48%	--	
It 6	54%	50%	48%	51%	sig.*	sig.*
It 7	46%	49%	47%	47%	sig.*	--
It 8	38%	37%	40%	38%	--	
It 9	42%	41%	39%	40%	--	
It 10	32%	32%	30%	32%	--	
It 11	48%	50%	47%	48%	--	
It 12	55%	55%	53%	54%	--	

* La diferencia según el orden en el cuadernillo es significativa al 95% de confianza

Como podemos ver, en un tercio del total de ítems del 2002 su posición en el cuadernillo afectó los resultados. En dos de ellos (ítem 3 e ítem 6) se observa que el porcentaje de respuestas correctas desciende a medida que el ítem se aproxima al final de la prueba.

Sin embargo, en otros dos el mayor porcentaje de respuestas correctas no se encuentra cuando el bloque del ítem está en primer lugar, sino cuando está en segundo lugar (ítem 4 e ítem 7).

Cabe destacar que en uno de estos dos ítems (7), además de lo recién señalado, el porcentaje de respuestas correctas es mayor cuando se encuentra al final del cuadernillo que cuando se ubica al principio. Para este ítem la diferencia deja de ser significativa cuando las respuestas se dicotomizan en "correctas" e "incorrectas".

³ La diferencia entre ambas, es que en la última la prueba de hipótesis se realizó considerando a los distractores por separado, mientras que en la penúltima se agruparon en una única categoría.

Por lo tanto, solamente en tres de los doce ítems de la prueba de 2002 de Matemática aplicada en 2005 encontramos que el porcentaje de respuestas disminuye cuando el ítem se encuentra en el último bloque del cuadernillo.

De acuerdo a esta evidencia no podríamos afirmar que el diseño de la prueba haya afectado los resultados de 2005.

Otro de los aspectos que podría estar influyendo en la comparación de 2002 y 2005 es la utilización de criterios diferentes para definir los contextos socioculturales. Para controlar este aspecto, los grupos evaluados en 2005 fueron clasificados según los criterios usados en 2002.

Al comparar los resultados de 2002 y los del 2005 categorizados según este último criterio no encontramos modificaciones relevantes en lo presentado hasta ahora. Esta información se presenta en anexo.

En resumen, tanto para Lengua como para Matemática, podemos decir que si bien se observa una tendencia que muestra resultados algo inferiores en 2005 que en 2002, si se consideran los márgenes de error de 2002 y se tiene en cuenta que en 2005 serán mayores que aquellos, probablemente sean pocas las diferencias estadísticamente significativas.

Anexo 1

CONTEXTO SOCIOCULTURAL

Desde la primera experiencia de evaluación nacional de aprendizajes en 6to. año (1996) se puso un fuerte énfasis en el análisis socialmente contextualizado de los resultados académicos, en virtud de los profusos antecedentes de investigación en el ámbito nacional e internacional acerca de la incidencia de las características socioculturales de las familias en el desarrollo de las competencias de los niños.

En el año 1996 se realizó un importante esfuerzo de clasificación de las escuelas primarias en función de su “contexto sociocultural”. Ello se hizo a partir de dos variables fundamentales: el nivel educativo de las madres de los niños y el nivel de equipamiento de los hogares de éstos. El nivel educativo materno fue utilizado bajo la forma de un “índice socioeducativo” calculado como la diferencia o saldo entre el porcentaje de madres con estudios secundarios completos (educación alta) y el porcentaje de madres con educación primaria como máximo (educación baja). Del mismo modo, el nivel de equipamiento de los hogares fue utilizado bajo la forma de un “índice socioeconómico” calculado como la diferencia o saldo entre el porcentaje de niños en hogares altamente equipados y el porcentaje de niños en hogares con bajo equipamiento. A la vez, las categorías de equipamiento alto y bajo fueron definidas a partir de una lista de 11 ítems de uso en los hogares¹. De esta manera, y luego de un arduo proceso de análisis de datos y verificación con informantes calificados con conocimiento directo de las escuelas, se logró una clasificación en cinco categorías: Muy favorable, Favorable, Medio, Desfavorable y Muy desfavorable. El resto de los indicadores sociales y los resultados de aprendizaje se ordenaron en términos perfectamente razonables a partir de esta clasificación, la que, por otra parte, ha sido utilizada en múltiples decisiones administrativas y de política educativa a lo largo de los siguientes seis años².

Naturalmente, a nueve años de haber elaborado el mencionado índice surge la necesidad de actualizarlo, tanto en cuanto a su concepción como a la forma de construirlo.

Actualización del contexto socioeducativo

Una de las razones fundamentales que justifican la actualización del índice es, precisamente, el lapso de tiempo que pasó desde su construcción y la actualidad. El hecho de que en 1996 haya sido una evaluación de carácter censal para los sextos años de educación primaria permitió reutilizar dicho índice en las evaluaciones de 1999 y 2002, de carácter muestral, incluso manteniendo los mismos criterios de construcción y puntos de corte de las variables constitutivas.

¹ Obviamente ambos índices correlacionan muy fuertemente entre sí, dado que las familias con más alto nivel educativo son al mismo tiempo las que tienen mejores ingresos. Si bien entonces los índices en cierto modo son redundantes, el uso combinado de ambos tiene la ventaja de corregir o suplir deficiencias en la información aportada por las familias. De todos modos, desde el punto de vista conceptual, se entiende que los factores con incidencia más significativa en los aprendizajes son los de carácter cultural, reflejados en el índice socioeducativo.

² Por más detalles acerca de los fundamentos y procedimientos seguidos para la definición de los contextos socioculturales véase el *Segundo Informe de Difusión Pública de Resultados-Evaluación Nacional de Aprendizajes en Lengua Materna y Matemática 1996* (ANEP/UMRE, 1997) y el informe de investigación *Estudio de los Factores Institucionales y Pedagógicos que inciden en los Aprendizajes en Escuelas Primarias de Contextos Sociales Desfavorecidos en el Uruguay* (ANEP/UMRE, 1999).

No obstante ello, tanto la nueva realidad socioeconómica como el tiempo transcurrido desde la primera elaboración del índice obligaron a replantear y actualizar su construcción. Así, luego de varias comparaciones con otros indicadores sociales que fueron utilizados para validar la construcción del índice se llega a uno nuevo que, si bien mantiene las variables que lo componen desde sus inicios³, actualiza los puntos de corte haciéndolos relativos a la nueva población estudiantil de sextos años.

Dicho en otros términos, se mantiene la lógica utilizada en el año 1996 pero con dos grandes novedades: se integran nuevos elementos de equipamiento de los hogares y se actualizan los criterios que conformarán las categorías del nuevo contexto. En cuanto al primer punto, se utilizó una lista de 15 ítems de equipamiento y no los 11 utilizados en aquel año, estableciendo que aquellos hogares con entre 0 y 4 elementos fueran considerados de *bajo equipamiento*, y los de 10 a 15 de *alto equipamiento*⁴. Seguidamente, se calculó el “saldo de equipamiento” mediante la resta del porcentaje de alumnos con alto equipamiento menos el porcentaje de alumnos con bajo equipamiento.

Una vez que se obtiene el “saldo educativo” de la escuela (al igual que en el caso anterior surge de la resta del porcentaje de madres con alto nivel de estudios menos el porcentaje de madres con baja escolarización⁵) se conforman cinco grupos con la misma cantidad de escuelas para establecer un orden en función de dicha caracterización, por lo que se llega a conformar de esta manera los *quintiles del saldo educativo de las escuelas*. De igual modo se procede con la variable “saldo de equipamiento” de las escuelas: se conforman los *quintiles del saldo de equipamiento de las escuelas*. Finalmente, y una vez que se combinan ambas medidas (los quintiles), se establece una tipología de 25 cuadrantes, desde los cuales se determinan las actuales cinco categorías del **nuevo contexto sociocultural** de las escuelas: Muy Favorable, Favorable, Medio, Desfavorable y Muy Desfavorable.

Antes de tomar decisiones sobre la modificación del índice de equipamiento se analizó la evolución de cada uno de los ítems desde el año 1996 hasta el 2005. Para ello se utilizó la información resultante de las anteriores mediciones de resultados, complementándose con información proveniente de la Encuesta Continua de Hogares (ECH) del Instituto Nacional de Estadística (INE) del año 2004 para una selección de población de similares características que la de las evaluaciones.

³ Las variables continúan siendo la educación materna del alumnado que concurre a la escuela y el nivel de equipamiento de sus hogares.

⁴ Anteriormente, si bien “bajo equipamiento” contemplaba la misma cantidad de elementos, la categoría de “alto equipamiento” significaba la tenencia de entre 9 y 11 ítems.

⁵ Para establecer Baja y Alta educación se mantienen los criterios utilizados anteriormente: hasta primaria completa en el primer caso y con secundaria o más en el segundo.

Equipamiento de los hogares de los alumnos de sextos años. 1996 - 2005

	1996	1999	2002	2005	ECH - 2004
Calefón o calentador de agua	72.8	80.4	81.4	67.1	66.1 (1)
Refrigerador sin freezer	66.9	63.7	53.7	48.4	92.6 (2)
Refrigerador con freezer	26.4	35.7	46.9	42.9	
Lavarropa	48.2	62.0	70.0	61.8	60.1
Teléfono	46.9	67.3	67.5	68.4	63.5
TV color	88.0	94.3	95.5	93.2	90.1
Video casete	44.4	44.8	42.4	32.6	31.2
Video game	50.7	48.7	39.6	33.9	
Computadora	11.0	19.3	27.8	29.2	22.5
Auto	33.2	37.3	36.2	30.7	28.8
Microondas		25.9	40.4	37.9	34.3
Lavavajillas		4.8	5.8	5.1	4.1
Secarropa		10.1	13.6		
Teléfono celular		22.7	33.5		
TV CABLE			40.6	37.1	
TV SATELITAL			3.9		
DVD			2.8	21.7	
Internet			17.6	15.5	12.7

(1) 2005 - pregunta solamente x calefón, y ECH - pregunta x calefón o termofón y en pregunta separada x calentador instantáneo de agua (17,3).

(2) ECH - pregunta solo por refrigerador.

La población seleccionada en la ECH – 2004 fue: a) entre 11 y 14 años de edad; b) encuestada entre agosto y noviembre; c) asistentes a 6º de primaria.

Según el cuadro se puede observar que el calefón, la heladera (con y sin freezer), el lavarropa y el auto son los ítems que más descienden al final de la serie; también lo hacen el video casete y el video game, aunque estos parecen compensarse con el aumento de la computadora y del DVD. Inversamente, la TV, microondas, TV cable, lavavajillas e Internet disminuyen levemente.

Por otro lado, las leves diferencias encontradas entre la medición del año 2005 y la información recogida por la ECH del 2004 parecen confirmar la necesidad de actualizar el índice de equipamiento y redefinir los puntos de corte para obtener una medida del contexto sociocultural de las escuelas actualizadas a la realidad del 2005.

El siguiente cuadro da cuenta de cómo ha evolucionado la distribución de los alumnos de sexto según el contexto sociocultural de las escuelas. Sin embargo, debe tenerse en cuenta que hasta el año 2002 la metodología utilizada para el cálculo del contexto fue la misma, mientras que para el 2005 se actualizó tal como fue detallado anteriormente. La última columna corresponde también a la información del año 2005 pero utilizando la metodología de cálculo que se usó hasta el año 2002: con los 11 ítems de equipamiento del hogar (y no los 15 mencionados) y respetando los cortes de los quintiles de los saldos educativos y de equipamiento trabajados hasta esa fecha.

Distribución de alumnos de sexto según contexto sociocultural. 1996 - 2005

	1996	1999	2002	2005	2005 (*)
Muy favorable	15,2	14,4	14,1	14,5	11,7
Favorable	18,5	18,4	16,8	19,6	15,3
Medio	19,2	18,4	17,3	25,7	20,6
Desfavorable	20,4	20,6	20,7	16,1	24,5
Muy desfavorable	26,7	28,1	29,9	24,1	27,9
Sin Dato			1,2		
Total	100,0	100,0	100,0	100,0	100,0

(*) Metodología del año 1996: mismos cortes de los quintiles que en 1996 y con 11 ítems de equipamiento

Anexo 2

Cuadro
Porcentaje promedio de respuestas correctas en 2002 y 2005 según el contexto sociocultural de cada evaluación y en 2005 clasificando el contexto según los criterios de 2002

		Resultados		
		Año 2002, según contexto 2002	Año 2005, según contexto 2005	Año 2005, según contexto 2002
Muy favorable	Texto explicativo	73	70	72
	Texto narrativo	79	77	79
	Matemática	67	65	67
Favorable	Texto explicativo	65	58	58
	Texto narrativo	72	69	70
	Matemática	58	55	57
Medio	Texto explicativo	60	55	58
	Texto narrativo	66	63	65
	Matemática	57	51	52
Desfavorable	Texto explicativo	58	50	51
	Texto narrativo	63	57	59
	Matemática	51	45	48
Muy desfavorable	Texto explicativo	55	46	47
	Texto narrativo	61	54	55
	Matemática	49	42	42