

# Rendimiento en lectura y matemática

de estudiantes de primero y tercero primaria, un  
estudio muestral a nivel nacional

M.Sc. Francisco  
José Ureta,  
M.Sc. Álvaro  
Mauricio Fortín,  
Dr. Roberto  
Molina Cruz

## Resumen

La evaluación fue realizada por el Centro de Investigaciones Educativas (CIE), del Instituto de Investigaciones de la Universidad del Valle de Guatemala, ejecutado por el Programa Nacional de Evaluación del Rendimiento Escolar (PRONERE), a requerimiento del Ministerio de Educación de Guatemala (MINEDUC). PRONERE realizó durante 2004, una evaluación en 823 escuelas públicas sobre el rendimiento en lectura y matemática de estudiantes de primero y tercero primaria. En primer grado se utilizaron instrumentos referidos a “criterio” y en tercer grado se utilizaron instrumentos referidos a la “norma”. Se aplicaron entrevistas y cuestionarios para alumnos(as), docentes y directores(as). En primer grado se aplicaron pruebas en español, kaqchikel y k’iche’; en tercero primaria además de las ya citadas, en mam y q’eqchi’.

La muestra se elaboró en paneles y para aplicaciones multianuales, es representativa a nivel nacional y departamental. Consideró las variables de ubicación geográfica (urbana-rural) y el tamaño de la escuela (número de estudiantes que atiende). Al estimar los resultados se consideró el peso que cada estrato tiene con respecto a la población de todo el país. En primer grado se evaluaron 19,530 alumnos(as) y en tercero a 16,770.

En primer grado el 48% de niños(as) lograron el criterio de lectura y 27.5% en matemática. Tienen mejor rendimiento los alumnos(as) de las escuelas urbanas, no indígenas y de varones. En tercero primaria las puntuaciones se estandarizaron a una media arbitraria de 100 puntos como promedio nacional, y tienen mejor rendimiento los alumnos(as) de las escuelas urbanas y no indígenas; en género no hay diferencias. El rendimiento en los dos grados es deficiente en términos generales. Como variables asociadas estudiadas influyen en el rendimiento: el desayuno el día de la evaluación, la inseguridad, el agrado por la escuela, el ambiente letrado en la escuela y la casa, la experiencia docente e infraestructura de la escuela (agua entubada y electricidad).

## Información general acerca de las pruebas

Se decidió evaluar las áreas de matemática y lectura de primero y tercero grados en el año 2004. Se seleccionaron estas materias ya que con base a la lectura y las destrezas básicas en matemática, es posible desarrollar habilidades y lograr aprendizajes que son gradualmente más complejos. La selección de los grados tuvo dos razones. Primero es un grado que no ha sido evaluado con anterioridad y que marca el punto de inicio de los estudiantes en su carrera primaria. Por ello es de relevancia conocer el grado de adquisición que han logrado de las destrezas básicas sobre las cuales se espera que construyan los conocimientos de primaria. Por otra parte, el tercer grado es importante en Guatemala ya que marca un punto en el cual cerca de la mitad de los estudiantes desertan del sistema.

Con respecto a los instrumentos que se utilizaron, existen dos corrientes de reciente interés acerca de la evaluación con propósitos educativos; evaluación que está relacionada en buena medida en la selección del tipo de prueba que se utiliza para la medición del logro de los estudiantes, así como la corriente predominante en Guatemala que es la de utilizar pruebas con referencia a la norma. Estos son instrumentos que toman un amplio espectro de áreas y sub-áreas relacionados con la materia que está siendo medida. Es un intento de muestrear el constructo bajo evaluación, asimismo, esta muestra de contenidos y habilidades evaluadas cubre un rango relativamente amplio de niveles de desempeño. Se busca que el techo y piso de la prueba sean capaces de reportar el desempeño de la gran mayoría de los estudiantes.

Su énfasis se encuentra en la posición que el estudiante toma dentro del grupo de estudiantes que son evaluados para establecer su grado de desempeño (1-3).

Otro tipo de pruebas son las de criterio, estas se elaboran en base a un nivel de desempeño claramente definido, pretenden descubrir cómo se encuentran los estudiantes respecto a ese criterio. Al igual que las referidas a la norma, describen el desempeño en las diferentes sub-áreas que se relacionan con el progreso de los estudiantes hacia ese mismo criterio (1-3).

Actualmente, en el área educativa la tendencia mundial se orienta principalmente hacia las pruebas de criterio, ya que ofrecen información que es más fácilmente convertida en insumos para retroalimentar al desarrollo curricular. Las pruebas referidas a la norma siguen siendo instrumentos valiosos, pero la tendencia actual es de utilizarlas para medir constructos cuya distribución se asume normal dentro de la población y, son características de la población que debe alcanzar los estándares y criterios establecidos por el currículo.

En Guatemala se ha tendido a utilizar principalmente pruebas referidas a la norma, pero se ha evidenciado la necesidad de proporcionar información más puntual acerca de destrezas concretas que son susceptibles de ser enfocadas por el currículo. Por ello se ha iniciado la transición hacia las

evaluaciones referidas a criterio. En el caso de la evaluación ejecutada por el PRONERE en el año 2004 y presentada en este trabajo, el esfuerzo se inició con la elaboración de pruebas referidas a criterio para el primer grado, un grado que no había sido evaluado en el pasado. Las evaluaciones del tercer grado se basaron en ítems de pruebas aplicadas de años anteriores, las cuales son pruebas normadas.

- Características de las pruebas del primer grado

Se utilizó un instrumento para matemática y uno para lectura, los cuales fueron diseñados y piloteados en su versión en español por el Proyecto MEDIR / Juárez & Asociados. Con base a estos instrumentos se realizaron traducciones y adaptaciones para crear instrumentos en los idiomas k'iche' y kaqchikel. Ambos instrumentos, el de lectura y el de matemática, fueron diseñados como instrumentos de criterio, estableciéndose éste por el logro en ciertos ítems. Después de la aplicación de los instrumentos se realizó el análisis de ítems. En la Tabla 1 se muestra el valor del  $\alpha$  de Cronbach de cada uno de los instrumentos utilizados, los cuales se consideran adecuados (de 0.80 a 1 se considera que la prueba tiene una buena confiabilidad y consistencia interna (4)).

- Cuestionarios adicionales de primero

Para indagar sobre los factores asociados al rendimiento de los alumnos se utilizaron cuestionarios y entrevistas, los instrumentos utilizados fueron: a) guía de entrevista a dos niños y a dos niñas de cada aula evaluada, b) cuestionario del docente, y c) cuestionario del director(a).

**TABLA 1**

Valores del  $\alpha$  de Cronbach para los instrumentos del primer grado de primaria

| Instrumento          | $\alpha$ de Cronbach | No. de ítems | No. de casos |
|----------------------|----------------------|--------------|--------------|
| Lectura español      | 0.88                 | 20           | 19,530       |
| Matemática español   | 0.86                 | 25           | 19,525       |
| Lectura K'iche'      | 0.88                 | 20           | 703          |
| Matemática K'iche'   | 0.87                 | 25           | 696          |
| Lectura Kaqchikel    | 0.76                 | 20           | 266          |
| Matemática Kaqchikel | 0.86                 | 25           | 266          |

- Características de las pruebas de tercero

PRONERE había evaluado desde el año 1997 a estudiantes del tercer grado de primaria, este bagaje de información fue aprovechada en la evaluación realizada en el año 2004, al utilizar ítems que ya habían sido probados en años anteriores y aprovechar los datos obtenidos en el pasado. Como en el pasado PRONERE había realizado sus evaluaciones con base a instrumentos normativos, en el 2004 la evaluación del tercer grado se basó en estos instrumentos ya validados y piloteados.<sup>1</sup> Los nuevos instrumentos, ensamblados a partir de ítems que ya habían sido probados con anterioridad, fueron contruidos a partir de la misma tabla de especificaciones que PRONERE utilizó desde sus inicios. Para la evaluación del desempeño en español se crearon cuatro instrumentos diferentes, todos aplicados en áreas urbana y rural, conteniendo ítems de las diferentes formas creadas durante el año 2000. Se seleccionó este año debido a la calidad psicométrica de los instrumentos, el número de casos con los cuales se podían hacer comparaciones y la distancia temporal. En dicho año se habían creado cuatro formas diferentes, dos de ellas diseñadas para ser aplicadas en el área rural y dos de ellas en la urbana. Las cuatro pruebas contaban con 40 ítems, 20 de estos eran comunes entre formas. Para la evaluación del desempeño en los idiomas k'iche, kakchikel, q'eqchi y mam, se utilizaron las mismas pruebas que PRONERE diseñó y aplicó en el año 1999, por las mismas razones mencionadas anteriormente.

Todas las formas creadas para la evaluación realizada se basaron en la misma tabla de especificaciones y en las mismas especificaciones de ítems, como se ilustra en detalle en *Especificación de ítems de las pruebas de tercer grado primaria* (6).

Después de la aplicación de los instrumentos se realizó el análisis de ítems de los mismos. En la Tabla 2 se muestran los índices Alfa de Cronbach, los cuales se refieren a la confiabilidad de las diferentes pruebas de tercer grado de primaria, estos índices se consideran adecuadas (de 0.80 a 1 se considera que la prueba tiene una buena confiabilidad y consistencia interna).

<sup>1</sup> Detalles pueden ser consultados en *Informe Nacional de la Evaluación del Rendimiento en Lectura y Matemática de Estudiantes del Tercer Grado de la Primaria de Instituciones Oficiales de Guatemala* (5).

**TABLA 2**Valores del  $\alpha$  de Cronbach para los instrumentos del tercer grado de primaria

| Instrumento  | $\alpha$ de Cronbach | No. de ítems | No. de casos |
|--|----------------------|--------------|--------------|
| Lectura Español<br>(todas las formas en conjunto)    | 0.86                 | 40           | 18,269       |
| Matemática Español<br>(todas las formas en conjunto) | 0.85                 | 40           | 18,268       |
| Lectura Español (forma A)                            | 0.87                 | 40           | 4,317        |
| Matemática Español (forma A)                         | 0.85                 | 40           | 73,340       |
| Lectura Español (forma B)                            | 0.88                 | 40           | 4,242        |
| Matemática Español (forma B)                         | 0.86                 | 40           | 71,264       |
| Lectura Español (forma C)                            | 0.85                 | 40           | 4,156        |
| Matemática Español (forma C)                         | 0.86                 | 40           | 69,736       |
| Lectura Español (forma D)                            | 0.88                 | 40           | 4,055        |
| Matemática Español (forma D)                         | 0.85                 | 40           | 67,719       |
| Lectura K'iche'                                      | 0.77                 | 40           | 651          |
| Matemática K'iche'                                   | 0.76                 | 40           | 653          |
| Lectura Kaqchikel                                    | 0.78                 | 40           | 249          |
| Matemática Kaqchikel                                 | 0.76                 | 40           | 202          |
| Lectura Q'eqchi'                                     | 0.82                 | 40           | 190          |
| Matemática Q'eqchi'                                  | 0.80                 | 40           | 189          |
| Lectura Mam  | 0.55                 | 40           | 205          |
| Matemática Mam                                       | 0.76                 | 40           | 202          |

Se comparó el funcionamiento de las diferentes formas en relación de una con las otras. Como primer instancia, se utilizó la  $t$  de medias relacionadas, ya que cada niño tomó ambas formas comparadas. Esto se realizó con el objetivo de determinar la equivalencia de medias y desviaciones estándar. También, se realizó una prueba de homogeneidad entre formas para definir si ambas formas producían distribuciones equivalentes. Todos estos análisis se realizaron para decidir si sería necesario realizar una conversión de los datos o los puntajes brutos podrían ser directamente comparados. Se encontraron diferencias entre las medias de varias de las parejas.

Sin embargo, sólo en dos casos se detectó heterogeneidad. Con base a ello se decidió equiparar las puntuaciones. Dado que en general las formas mostraron distribuciones similares, se optó por estimar para cada forma el puntaje estándar  $z$  y convertir las puntuaciones de acuerdo a la siguiente fórmula:

$$Y = m + k * z$$

en donde:  
 $m$  = media arbitraria  
 $k$  = desviación típica arbitraria  
 $z$  = valor  $z$ .

Para la media nacional arbitraria se seleccionó 100 y para la desviación típica nacional arbitraria 20. De esta manera las formas serían comparables. Al ajustar la distribución de los casos a la distribución normal, se hacía posible utilizar las pruebas estadísticas que exigen esto como requisito (por ejemplo, las comparaciones  $t$  de medias).

- Cuestionarios adicionales de tercero

También en tercer grado se indagó sobre algunos factores asociados al rendimiento de los alumnos, se utilizaron cuestionarios y entrevistas dedicados a recolectar información acerca de factores que podrían estar relacionados con el rendimiento de los estudiantes. Los instrumentos utilizados fueron: a) cuestionario al estudiante, b) cuestionario del docente y c) cuestionario del director(a).

## Descripción del muestreo realizado

- El diseño de la muestra

El marco de muestreo fue elaborado a partir de la base de datos del MINEDUC del 2003, con una actualización del status de los establecimientos en el 2004. Se diseñó un procedimiento de selección de los establecimientos y alumnos, tomando en cuenta que una evaluación similar podría estarse realizando cada año, por lo que se incorporó la metodología estándar del manejo de aplicaciones periódicas. A cada alumno examinado se le asignó un factor de expansión empleando un ajuste por post estratificación, para ello se consideró el género del alumno, el departamento y área en que se localiza su establecimiento. Por último, se calcularon calificaciones estandarizadas para los resultados de las dos pruebas –lectura y matemática- que fueron suministrados a cada estudiante.

- Las poblaciones objetivo y en estudio

Entre los objetivos del estudio quedó establecido que se debían considerar a todos los alumnos de primer y tercer grados de primaria de los establecimientos del sector oficial, con jornadas matutina y doble del plan regular (diario) de todo el país. No se consideraron las escuelas con jornada vespertina ya que solo hay en el área urbana y no en la rural, para que la muestra mantuviera su estructura no se incluyeron estas escuelas urbanas. Para la elaboración del marco de muestreo, se utilizó como base una lista actualizada de todos los establecimientos oficiales de primaria del país. En esta lista se observaron 14,678 establecimientos, de los cuales 74 pertenecían al sector municipal, 861 eran establecimientos con jornada vespertina, y 2 eran del plan sabatino o fin de semana. Se redujo esta primera lista a 13,741 establecimientos con las características deseadas.

No se tomó en cuenta los establecimientos con una matrícula menor a 10 alumnos en primer grado o menor a 5 alumnos en tercero. En la lista de los 13,741 establecimientos habían 2,453 (17.9%) establecimientos sin la matrícula requerida. Estos establecimientos correspondían en mayor porcentaje (18.8%) al área rural que al área urbana (4.3%). Los siguientes departamentos tuvieron más escuelas en el área rural: El Progreso (33.5%), Chiquimula (31.7%), Zacapa (27.8%), Baja Verapaz (25.9%), Alta Verapaz (24.0%), Petén (23.9%), Izabal (23.0%), Quiché (22.2%), Jalapa (20.9%) y Huehuetenango (20.6%).

La población objetivo consistió en 521,490 estudiantes en primer grado y 307,932 en tercero. Los alumnos que asistían a los establecimientos pequeños (menos de 10 alumnos en alguno de los grados evaluados) no considerados en la evaluación fueron 31,162 (6.0%) en primer grado y 10,326 (2.4%) en tercero. Hubo porcentajes importantes de alumnos que asistían a estas escuelas pequeñas del área rural: 30,222 (6.9% del total) en primer grado y 10,268 (4.2%) en tercero. Estos alumnos fueron más en el área rural de los siguientes departamentos: Chiquimula, Alta Verapaz, Baja Verapaz, Izabal, Petén, Zacapa, El Progreso y Quiché.

- El objetivo considerado

El objetivo principal de la evaluación fue obtener estimaciones del rendimiento de los alumnos(as) de cada departamento, la ciudad capital y al nivel nacional. Por lo que la muestra se seleccionó para obtener estimaciones con las mejores precisiones posibles al nivel departamental y nacional. También se requirió que las estimaciones fueran comparables entre dos años diferentes, considerando la posibilidad que una evaluación similar se esté realizando cada año; por ello se aplicó la metodología de evaluaciones periódicas, con un diseño que consideró el uso de paneles de establecimientos y sus posibles formas de rotación.

- El marco de muestreo

Se consideró a las escuelas como las unidades primarias de muestreo y las secciones de cada grado como las unidades secundarias. Esto es un diseño de dos etapas, con los establecimientos y las secciones como conglomerados. Los establecimientos fueron estratificados por medio de las variables siguientes: a) departamento, por ser los dominios de estudio, b) área: urbano y rural, como primera variable de estratificación, y c) el tamaño de los establecimientos, entre 2 y 6 niveles de tamaño, como segunda variable de estratificación.

Para un adecuado manejo de las posibles evaluaciones anuales similares (aplicación secuencial), se consideró la población de establecimientos formada por paneles o grupos de establecimientos de un mismo estrato. El número de paneles en cada estrato quedó determinado por lo siguiente: a) el número de establecimientos en el estrato, b) el tipo de rotación que tendrán los paneles entre las muestras de cada dos años consecutivos, y c) el número de establecimientos de cada estrato que se decidan incluir en la muestra de cada año.

Para establecer el tipo de rotación de los paneles, se consideró que el proceso educativo no sufre de cambios significativos tan rápido, por lo que es más conveniente regresar a visitar las mismas escuela después de un período más largo de tiempo, para poder evaluar los posibles cambios ocurridos durante esos períodos. Por ello se incluirá en la muestra de cada año un solo panel por cada estrato, los paneles de la muestra tienen una rotación completa (100%) cada año.

- La distribución de la muestra

Se determinó que la muestra debería contar con un máximo de 820 escuelas; de las cuales 20 fueron seleccionadas en forma independiente por otros investigadores entre las escuelas del programa PAEBI (se seleccionó ese número de escuelas debido al factor financiero, cantidad de escuelas que podían financiar para que fueran evaluadas), por lo que se estableció un procedimiento de selección de 800 escuelas, dividido en las dos etapas siguientes: a) la selección aleatoria de aproximadamente 770 escuelas de todos los estratos, y b) la posible selección no aleatoria de 30 escuelas de los otros tres programas considerados: PRONADE, DIGEBI y CETT, para asegurar un número adecuado de escuelas de estos programas en la muestra.

Los detalles de la distribución de los 770 establecimientos, la definición y la estimación de los factores de expansión y la estandarización de las calificaciones pueden encontrarse en *Diseño de la encuesta y la obtención de la muestra 2004* (7) y en *Evaluación de la muestra realizada y cálculo de los expansores* (8).

## Resultados de primero primaria

- Resultados de la prueba de Lectura

El número total de niños y niñas que tomó la evaluación de lectura a nivel nacional fue de 19,548. Los resultados de aquellos que alcanzaron el criterio establecido se presentan en la Tabla 3, por estratos de área, etnia y género.

### TABLA 3

Porcentajes de logro en la prueba de primer grado primaria por área (urbano-rural), etnia (no indígena-indígena) y género (masculino-femenino)

| Área               | %                   | Etnia       | %                  | Género    | %      |
|--------------------|---------------------|-------------|--------------------|-----------|--------|
| Urbana<br>n = 5599 | 71.33%<br>n = 14123 | No indígena | 72.31%<br>n = 9650 | Masculino | 49.41% |
| Rural<br>n = 13931 | 28.67%<br>n = 5407  | Indígena    | 27.69%<br>n = 9880 | Femenino  | 50.59% |

Existe una diferencia estadísticamente significativa ( $p < 0.05$ , z de proporciones) entre los resultados del área urbana y rural. En el área urbana un mayor porcentaje de estudiantes de primer grado alcanzó el criterio de la prueba de lectura, esto se debe a que en las áreas urbanas hay más ambientes letrados, acceso a material escrito y numérico para los alumnos y, las escuelas en general, tienen mejores condiciones de infraestructura y materiales de apoyo que las rurales. En relación con la etnia, el porcentaje de estudiantes que fue descrito por los docentes como indígena y que alcanzó el criterio es menor al porcentaje de estudiantes que no fue descrito como indígena y que alcanzó el criterio. Esta diferencia es estadísticamente significativa ( $p < 0.05$ , z de proporciones). Para la interpretación apropiada de este hallazgo, hay que tener presente que un alto porcentaje de estudiantes denominados como indígenas viven en el área rural, por lo que podría ser la ruralidad y no la etnia la verdadera razón de esta diferencia. El porcentaje de estudiantes mujeres que alcanzó el criterio es similar al porcentaje de estudiantes de género masculino que lo logró. Aunque hubo una diferencia estadísticamente significativa entre proporciones (la proporción no aproximada para las niñas

fue de 47.7%), la variable no es significativa en un modelo que explique las diferencias de logro.<sup>2</sup>

La Tabla 4 que se presenta a continuación resume los porcentajes de los(as) estudiantes del país que fueron exitosos en cada sección de la prueba. Debe notarse la disminución del éxito conforme aumenta la dificultad de las secciones, en general solo se puede considerar apropiados los porcentajes de logro en las primeras dos secciones, el resto están alrededor o por debajo del 50% de logro.

Se consideró exitoso el logro de cada sección cuando los niños respondieron de forma correcta por lo menos a dos ítems de las secciones uno, dos y, las cuales estaban conformadas por un mínimo de dos ítems y un máximo de tres.

La sección cuatro es el criterio de logro o no logro de la prueba, los estudiantes debieron responder correctamente los 3 de esta sección.

### TABLA 4

Porcentaje de estudiantes del país que obtuvo logro en cada sección de la prueba de lectura de primer grado primaria

| Sección | Aspecto evaluado  | %  |
|---------|---|----|
| 1       | Asociar un sonido o una letra del alfabeto (fonema-grafema)                     | 79 |
| 2       | Asociar una palabra escrita con una palabra expresada oralmente                 | 67 |
| 3       | Asociar una palabra escrita a una imagen que la representa                      | 47 |
| 4       | Comprender instrucciones que se dan en oraciones formadas por cuatro palabras   | 51 |
| 5       | Asociar frases de dos palabras o imágenes                                       | 48 |
| 6       | Leer una oración y responder una pregunta relacionada con ella                  | 41 |
| 7       | Completar una frase con la información que obtiene de una primera oración       | 32 |
| 8       | Leer un párrafo de tres oraciones y responder una pregunta relacionada con éste | 25 |

<sup>2</sup> Existen tres tipos de significancia que los investigadores en ciencias sociales utilizan para dar sentido a los resultados que obtienen:

a) **Estadística:** Evalúa la probabilidad de que los resultados en la muestra sean verdaderos tomando en cuenta el tamaño de ésta y asumiendo que proviene de una población en la cual, de verificarse así en la muestra, la hipótesis nula es exactamente verdadera. Este es el caso de la significancia que se encontró entre la proporción de estudiantes del género masculino y femenino que obtuvieron el logro del criterio.

b) **Práctica:** Cuantifica el grado hasta el cual los resultados de la muestra difieren de la hipótesis nula. Esto también se denomina "tamaño del efecto", lo cual involucra la confiabilidad con la cual los resultados muestran el efecto de la variable independiente.

c) **Clinica:** Explora si el efecto reportado indica un cambio cualitativo. Por ejemplo, si el efecto de una intervención provoca un cambio en puntuaciones de rendimiento, se exploraría si el cambio es suficiente para que los y las estudiantes hayan logrado un criterio.

En el caso aquí expuesto, la diferencia de logro entre niños y niñas es de 0.3%. Por ello, aunque se encuentra una significancia estadística utilizando la prueba z de proporciones, esta puede haber sido un resultado del gran tamaño de la muestra. Un análisis posterior realizado con regresión logística indicó que la variable de género no es relevante para explicar la diferencia en el logro.

En la Tabla 5 se muestran los resultados para el grupo de estudiantes que tomó tanto la prueba en español como en k'iche' o kaqchikel, de acuerdo al idioma hablado por los niños y niñas.

**TABLA 5**

Logro en las pruebas de lectura de primer grado en k'iche' o kaqchikel y español

| Lectura                                  | K'iche'<br>n = 703 | % ajustado<br>al diseño de<br>la muestra | Kaqchikel<br>n = 266 | % ajustado<br>al diseño de<br>la muestra |
|--|--------------------|--|----------------------|--|
| Logro en k'iche' o kaqchikel             | 372                | 56                                       | 138                  | 57                                       |
| Logro en español                         | 340                | 62                                       | 160                  | 55                                       |
| Logro en ambos idiomas                   | 236                | 38                                       | 99                   | 39                                       |
| Logro en uno o los dos idiomas evaluados | 476                | 79                                       | 199                  | 73                                       |

Para interpretar apropiadamente los datos relacionados con pruebas en idiomas mayas, es importante mencionar lo siguiente:

- Cada uno de los casos incluidos fue un o una estudiante que completó la prueba tanto en español como en k'iche' o kaqchikel.
- La primera fila, *Logro en k'iche' o kaqchikel*, indica la frecuencia real y el porcentaje ajustado a la muestra de niños que tomaron los instrumentos tanto en español como en k'iche' o kaqchikel y que llenaron el criterio de logro de estas últimas.
- La segunda fila, *Logro en español*, indica la frecuencia real y el porcentaje ajustado a la muestra de niños que tomaron los instrumentos tanto en español como en k'iche' o kaqchikel y que llenaron el criterio de logro en la prueba en español.
- La tercera fila, *Logro en ambos idiomas*, indica la frecuencia real y el porcentaje ajustado a la muestra de niños y niñas que tomaron los instrumentos tanto en español como en k'iche' o kaqchikel y que llenaron el criterio de logro en español y el idioma Maya correspondiente.

- Finalmente la cuarta fila, *Logro en uno o los dos idiomas evaluados*, indica la frecuencia real y el porcentaje ajustado al diseño de la muestra de niños que llenaron el criterio de logro en español, más los que lo llenaron en el idioma Maya correspondiente, más los que lo llenaron en ambos. En el caso de la columna que hace referencia al idioma k'iche' se cuantifica el número de estudiantes que alcanzó el logro en a) español o b) k'iche' o c) español y k'iche', igualmente para la columna del idioma kaqchikel.

- **Resultados de la prueba de Matemática**

El número total de niños que tomó la evaluación de matemática a nivel nacional fue de 19,525. El criterio se estableció por la ejecución de dos tipos diferentes de tareas y cada criterio se analizó por separado y en conjunto. Del total de estudiantes que tomó la evaluación de matemática, el 45.3% alcanzó el criterio establecido para sumas y el 38.6% alcanzó el criterio establecido para restas.

El 27.5% alcanzó los criterios de suma y resta combinados. Los resultados desagregados se presentan en la Tabla 6.

**TABLA 6**

Porcentaje de estudiantes que alcanzó los criterios

|           |                          | Logro de criterio |       |              |
|-----------|--------------------------|-------------------|-------|--------------|
|           |                          | Suma              | Resta | Suma y Resta |
| Área      | Urbana<br>n = 5598       | 56%               | 50%   | 36%          |
|           | Rural<br>n = 13927       | 43%               | 36%   | 26%          |
| Etnicidad | Indígena<br>n = 5541     | 45%               | 35%   | 26%          |
|           | No indígena<br>n = 13984 | 46%               | 40%   | 28%          |
| Género    | Femenino<br>n = 9876     | 45%               | 38%   | 27%          |
|           | Masculino<br>n = 9649    | 46%               | 39%   | 28%          |

Se realizaron análisis por separado para comparar las variables área, etnia y género. Se observó una diferencia estadísticamente significativa entre el porcentaje de niños y niñas de las áreas urbana y rural que alcanzaron el criterio ( $p < 0.05$ ,  $z$  de proporciones). El área urbana alcanzó un desempeño

superior al del área rural. El porcentaje de estudiantes que alcanzó los criterios y que fue descrito por los docentes como indígena, es menor que el porcentaje de estudiantes que no es descrito como indígena y que alcanza el criterio. Esta diferencia es significativa ( $p < 0.05$ ,  $z$  de proporciones). Las razones explicadas en los resultados de lectura para las diferencias de ubicación geográfica y etnicidad también se aplican a los resultados de matemática. El porcentaje de estudiantes mujeres que alcanzó los criterios es menor que el porcentaje de estudiantes varones que lo alcanzó. Sin embargo, la variable género no alcanzó el nivel de significancia para formar parte de un modelo que prediga el logro entre estudiantes de ambos sexos.

La prueba de matemática estuvo compuesta por ocho diferentes secciones, cada una de las cuales evaluó un aspecto diferente. Se analizaron los datos por cada sección a nivel nacional y de cada departamento. Se consideró exitoso el logro de cada sección cuando los niños respondieron de forma correcta por lo menos a dos ítems de las secciones uno, dos, tres, cuatro, siete y ocho, las cuales estaban conformadas por un mínimo de dos ítems y un máximo de tres. En las secciones cinco y seis, que son el criterio de la prueba en su totalidad, los estudiantes debieron marcar correctamente la respuesta de los ítems que se establecieron como criterio. En la Tabla 7 se indican los porcentajes de los(as) estudiantes del país que fueron exitosas(os) en cada sección. Solamente las secciones 2, 3 y 4 tuvieron un logro adecuado, el resto se ubican por debajo del 50% de logro.

**TABLA 7**

Porcentaje de estudiantes que obtuvieron logro en cada sección de la prueba de matemática de primer grado primaria

| Sección | Criterio evaluado  | %  |
|---------|--|----|
| 1       | Reconocer los números mayas.   | 24 |
| 2       | Reconocer cantidades en ilustraciones identificando "más" y "menos".                       | 80 |
| 3       | Identificar el número que expresa la cantidad de figuras en un grupo.                      | 80 |
| 4       | Reconocer la posición de una figura dentro de un grupo, identificando "antes" y "después". | 75 |
| 5       | Hacer sumas de dos o tres sumandos sin llevar.   | 47 |
| 6       | Hacer restas con el minuendo de un dígito y el sustraendo de un dígito sin prestar.        | 41 |
| 7       | Identificar el lugar que ocupan las figuras de una serie (primero, segundo, tercero).      | 27 |
| 8       | Completar series de números.   | 13 |

En la Tabla 8 se muestran los resultados para el grupo de estudiantes que tomó tanto la prueba en español como en k'iche' o kaqchikel (de acuerdo al idioma utilizado en la escuela). El criterio en este caso es el logro combinado de sumas y restas. Las recomendaciones para interpretar los datos se mencionaron anteriormente.

**TABLA 8**

Logro en la prueba de matemática de primer grado en K'iche' y Kaqchikel

| Matemática                               | K'iche'<br>n = 696 | % ajustado<br>al diseño de<br>la muestra | Kaqchikel<br>n = 266 | % ajustado<br>al diseño de<br>la muestra |
|--|--------------------|--|----------------------|--|
| Logro en K'iche' o Kaqchikel             | 165                | 28                                       | 72                   | 26                                       |
| Logro en español                         | 149                | 23                                       | 74                   | 25                                       |
| Logro en ambos idiomas                   | 95                 | 16                                       | 45                   | 16                                       |
| Logro en uno o los dos idiomas evaluados | 219                | 35                                       | 101                  | 36                                       |

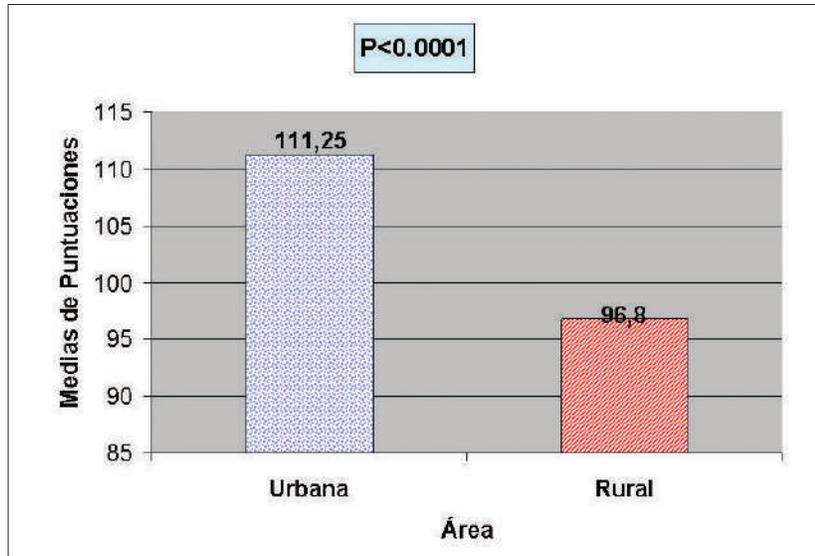
## Resultados de tercero primaria

- Resultados de la prueba de Lectura

Los resultados se presentan como promedios de las puntuaciones obtenidas por los estudiantes de cada departamento. Tanto en vocabulario como en comprensión, la puntuación global se estandarizó de manera independiente, con una media arbitraria de 100 y una desviación estándar de 20 para los resultados nacionales. Los resultados mostrados en las Gráficas 1 a la 3 indican lo siguiente: a) hay diferencia estadísticamente significativa entre el desempeño de estudiantes del área urbana y rural, la cual indica que el rendimiento es menor en el área rural, b) en lo que respecta a etnia, existió una diferencia significativa en los estudiantes que no se definieron como pertenecientes a alguna etnia, quienes obtuvieron mejores resultados, (las razones argumentadas anteriormente en la interpretación de estas variables también se aplican en este caso) y c) al comparar por género no se encontró una diferencia estadísticamente significativa entre hombres y mujeres.

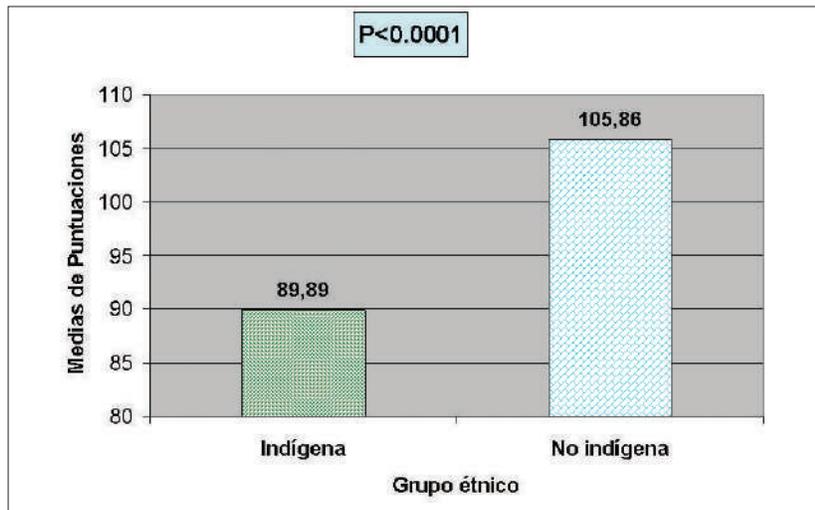
### GRÁFICA 1

Medias de lectura en español de estudiantes de tercer grado por área (urbana o rural)



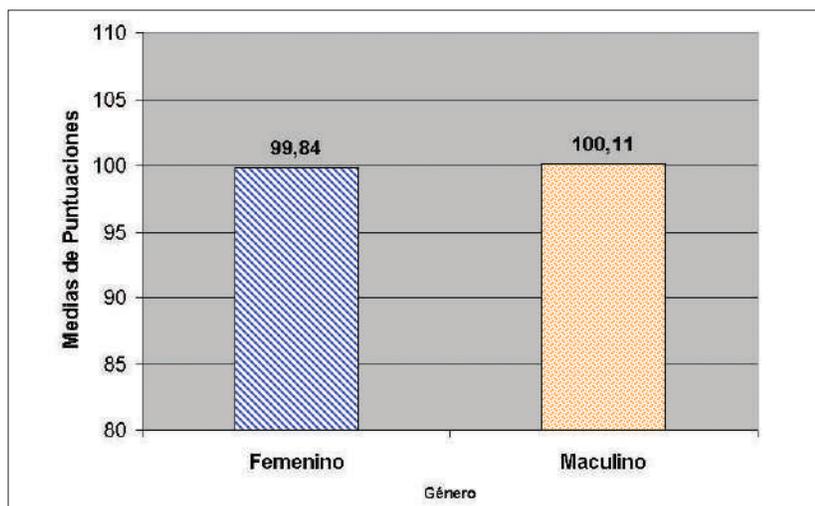
### GRÁFICA 2

Medias de lectura en español de estudiantes de tercer grado por etnia (indígena o no indígena)



### GRÁFICA 3

Medias de lectura en español de estudiantes de tercer grado por género (masculino o femenino)



Es necesario llamar la atención al lector sobre la dispersión de los datos de las puntuaciones en la muestra de estudiantes de este tamaño, considerada en la prueba estadística. Las gráficas permiten apreciar entre lo que es estadísticamente significativo a lo que es igual o diferente en la realidad.

La prueba de lectura de tercer grado estuvo constituida por las áreas de vocabulario y comprensión lectora. Las puntuaciones de cada una de estas áreas fueron estandarizadas de manera independiente utilizando una media de 100 y una desviación típica de 20, ambas a nivel nacional. Debido a ello, el promedio o la suma de los resultados por sección no son equivalentes a la puntuación total nacional. En la Tabla 9 se presentan los resultados con base en área, etnia y género. Las diferencias a favor del área urbana y etnia no indígena vuelven a aparecer en estos resultados.

### TABLA 9

Comparación de medias de las sub-secciones de la prueba de lectura en español por área (urbana o rural), etnia (indígena o no indígena) y género (masculino o femenino)

| Sección     | Comparación | Categoría   | n     | Media  | Error típico | Desviación típica |
|-------------|-------------|-------------|-------|--------|--------------|-------------------|
| Vocabulario | Área        | Urbana      | 5596  | 110.15 | 0.63         | 17.16             |
|             |             | Rural       | 11174 | 97.13  | 0.53         | 19.81             |
|             | Etnia       | Indígena    | 5367  | 90.81  | 0.66         | 19.15             |
|             |             | No indígena | 11403 | 105.56 | 0.45         | 17.83             |
|             | Género      | Femenino    | 8552  | 99.77  | 0.52         | 20.00             |
|             |             | Masculino   | 8218  | 100.22 | 0.50         | 20.00             |
| Comprensión | Área        | Urbano      | 5596  | 110.64 | 0.66         | 18.04             |
|             |             | Rural       | 11174 | 96.97  | 0.51         | 19.49             |
|             | Etnia       | Indígena    | 5367  | 90.55  | 0.63         | 18.56             |
|             |             | No indígena | 11403 | 105.26 | 0.44         | 18.08             |
|             | Género      | Femenino    | 8552  | 99.97  | 0.52         | 20.34             |
|             |             | Masculino   | 8218  | 100.03 | 0.48         | 19.68             |

A continuación se presentan dos tablas; la Tabla 10 muestra los resultados en las evaluaciones en cada uno de los idiomas mayas por género. Mientras que en el idioma k'iche' se mostró ventaja para las niñas, en kaqchikel, se obtuvo para los niños. En q'eqchi' y en mam los resultados son similares. Tal como en el caso de las evaluaciones en idioma español, los resultados fueron estandarizados utilizando una media arbitraria de 100 y una desviación típica de 20 y en los datos debe apreciarse entre lo que es estadísticamente significativo a lo que es igual o diferente en la realidad. La Tabla 11 muestra el resultado de una correlación entre los resultados en español con el idioma materno (k'iche', kaqchikel, q'eqchi' o mam). Se observa una correlación estadísticamente significativa, en el nivel indicado, entre los resultados de lectura en español y k'iche', q'eqchi' y mam. Esta relación no fue significativa en el caso del kaqchikel. Las consideraciones para interpretar los datos fueron mencionadas anteriormente.

**TABLA 10**

Resultados prueba de lectura en idiomas K'iche', Kaqchikel, Q'eqchi' y Mam por género

| Idioma    | Género    | n   | Media  | Error típico | Desviación típica |
|-----------|-----------|-----|--------|--------------|-------------------|
| K'iche'   | Femenino  | 321 | 103.33 | 3.67         | 19.28             |
|           | Masculino | 317 | 97.03  | 3.80         | 20.17             |
| Kaqchikel | Femenino  | 116 | 98.28  | 5.01         | 20.99             |
|           | Masculino | 130 | 101.43 | 5.89         | 19.03             |
| Q'eqchi'  | Femenino  | 78  | 99.83  | 3.52         | 17.72             |
|           | Masculino | 110 | 100.13 | 4.30         | 21.63             |
| Mam       | Femenino  | 104 | 100.22 | 1.74         | 20.01             |
|           | Masculino | 85  | 99.76  | 3.83         | 19.99             |

**TABLA 11**

Correlación entre los resultados de lectura en idioma español y en los idiomas K'iche', Kaqchikel, Q'eqchi' y Mam

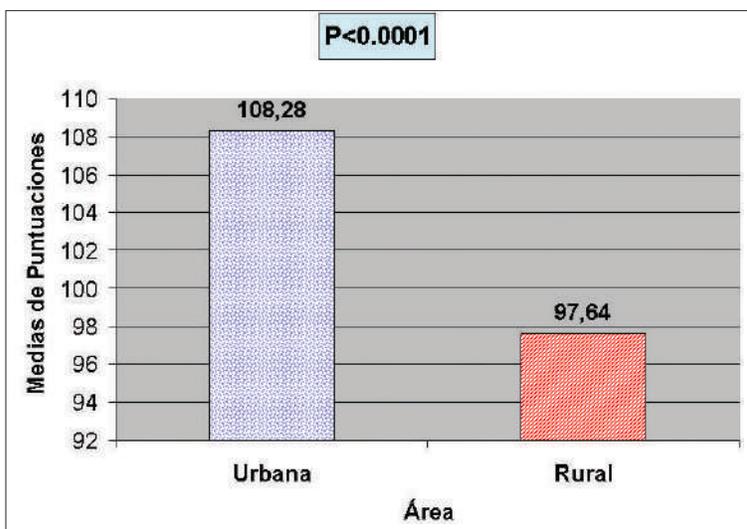
| Idioma    | R de Pearson | Significancia    |
|-----------|--------------|------------------|
| K'iche'   | 0.295        | 0.01             |
| Kaqchikel | 0.003        | No significativo |
| Q'eqchi'  | 0.425        | 0.01             |
| Mam       | 0.219        | 0.01             |

• Resultados de la prueba de Matemática

Tal como se indicó con antelación, los resultados se presentan como promedios de las puntuaciones obtenidas por los estudiantes de cada departamento. Para hacer más fácil la interpretación, los resultados fueron convertidos en puntuaciones estándar según las cuales una puntuación de 100 puntos es el promedio nacional, y la desviación típica es de 20 puntos.

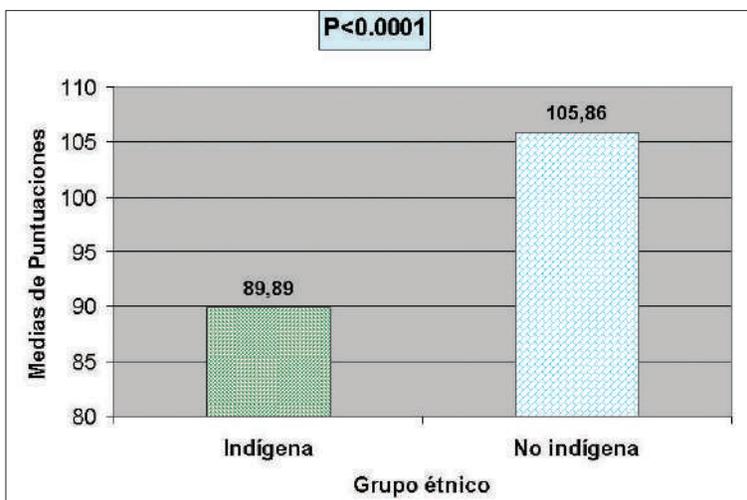
Los resultados que se muestran en las Gráficas 4, 5 y 6 indican lo siguiente:

- a) hay diferencia estadísticamente significativa entre el desempeño de estudiantes del área urbana y rural la cual indica que el rendimiento fue menor en el área rural, b) hay diferencia significativa en la comparación por etnia, en donde los estudiantes que se definieron como no indígenas obtuvieron mejores resultados, (las razones argumentadas anteriormente en la interpretación de estas variables también se aplican en este caso) y c) al comparar por género hay una diferencia significativa, los estudiantes varones obtuvieron resultados superiores a los obtenidos por las estudiantes de género femenino.



**GRÁFICA 4**

Comparación de medias de matemática español entre estudiantes de tercer grado por área (urbana o rural)

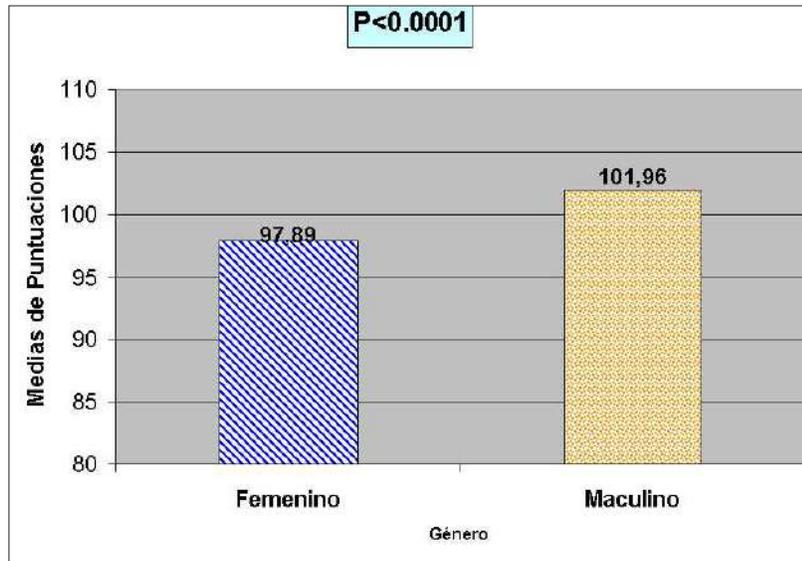


**GRÁFICA 5**

Comparación de medias de matemática español entre estudiantes de tercer grado por etnia (indígena o no indígena)

**GRÁFICA 6**

Comparación de medias de matemática español entre estudiantes de tercer grado por género (masculino o femenino)



Es necesario, de nuevo, llamar la atención al lector sobre la dispersión de los datos de las puntuaciones en una muestra de este tamaño, considerada en la prueba estadística. De nuevo las gráficas permiten apreciar entre lo que es estadísticamente significativo a lo que es igual o diferente en la realidad.

La prueba de matemática de tercer grado estuvo constituida por tres áreas: resolución de operaciones, conceptos aritméticos y resolución de problemas.

Las puntuaciones de cada una de estas áreas fueron estandarizadas de manera independiente utilizando una media de 100 y una desviación típica de 20. Debido a ello, el promedio o la suma de los resultados por sección no son equivalentes a la puntuación total nacional. La Tabla 12 presenta los resultados con base en área, etnia y género. Nuevamente aparecen las diferencias a favor del área urbana y etnia no indígena en estos resultados.

**TABLA 12**

Medias de las sub-secciones de la prueba de matemática en español entre estudiantes de tercer grado por área (urbana o rural), etnia (indígena o no indígena) y género (masculino o femenino)

| Sección                 | Comparación | Categoría   | n     | Media  | Error típico | Desviación típica |
|-------------------------|-------------|-------------|-------|--------|--------------|-------------------|
| Operaciones             | Área        | Urbana      | 5596  | 105.99 | 0.69         | 18.50             |
|                         |             | Rural       | 11174 | 98.40  | 0.57         | 20.07             |
|                         | Etnia       | Indígena    | 5367  | 96.01  | 0.70         | 19.90             |
|                         |             | No indígena | 11403 | 102.50 | 0.57         | 19.62             |
|                         | Género      | Femenino    | 8552  | 98.67  | 0.54         | 20.21             |
|                         |             | Masculino   | 8218  | 101.39 | 0.52         | 16.69             |
| Conceptos aritméticos   | Área        | Urbano      | 5596  | 108.72 | 0.66         | 18.71             |
|                         |             | Rural       | 11174 | 97.72  | 0.51         | 19.68             |
|                         | Etnia       | Indígena    | 5367  | 92.77  | 0.63         | 18.92             |
|                         |             | No indígena | 11403 | 104.15 | 0.47         | 19.10             |
|                         | Género      | Femenino    | 8552  | 98.43  | 0.50         | 19.78             |
|                         |             | Masculino   | 8218  | 101.78 | 0.49         | 20.02             |
| Resolución de Problemas | Área        | Urbano      | 5596  | 104.87 | 0.57         | 20.11             |
|                         |             | Rural       | 11174 | 98.70  | 0.44         | 19.77             |
|                         | Etnia       | Indígena    | 5367  | 95.99  | 0.56         | 19.23             |
|                         |             | No indígena | 11403 | 102.43 | 0.44         | 19.98             |
|                         | Género      | Femenino    | 8552  | 98.07  | 0.43         | 19.37             |
|                         |             | Masculino   | 8218  | 101.93 | 0.43         | 20.39             |

A continuación se presentan dos tablas; en la Tabla 13 aparecen los resultados obtenidos por género y en la Tabla 14 se muestra la correlación entre los resultados obtenidos en idiomas español y Maya. Tal como se hizo en las evaluaciones en idioma español, los resultados fueron estandarizados utilizando una media arbitraria de 100 y una desviación típica de 20. Las consideraciones para interpretar los datos fueron mencionadas anteriormente.

**TABLA 13**

Resultados en la prueba de matemática en idiomas K'iche, Kaqchikel, Q'eqchi' y Mam por género

| Idioma    | Género    | n   | Media  | Error típico | Desviación típica |
|-----------|-----------|-----|--------|--------------|-------------------|
| K'iche'   | Femenino  | 323 | 97.71  | 2.30         | 18.98             |
|           | Masculino | 317 | 102.10 | 3.72         | 20.57             |
| Kaqchikel | Femenino  | 116 | 98.04  | 4.86         | 19.91             |
|           | Masculino | 130 | 101.61 | 5.24         | 19.93             |
| Q'eqchi'  | Femenino  | 78  | 98.61  | 4.59         | 18.96             |
|           | Masculino | 109 | 101.10 | 5.67         | 20.73             |
| Mam       | Femenino  | 103 | 96.88  | 5.19         | 18.75             |
|           | Masculino | 83  | 103.66 | 6.48         | 20.79             |

**TABLA 14**

Correlación entre los resultados de matemática en idioma español y en los idiomas K'iche, Kaqchikel, Q'eqchi' y Mam

| Idioma    | R de Pearson | Significancia |
|-----------|--------------|---------------|
| K'iche'   | 0.467        | 0.01          |
| Kaqchikel | 0.736        | 0.01          |
| Q'eqchi'  | 0.604        | 0.01          |
| Mam       | 0.491        | 0.01          |

### Factores asociados al aprendizaje de las y los estudiantes de primero y tercero

Los resultados del rendimiento en lectura y matemática también se analizaron a la luz de información complementaria proporcionada por los directores, docentes y estudiantes. Los principales hallazgos se resumen así:

- **Nutrición y alimentación**

En primer grado se encontró que la ingesta de alimentos previo a asistir a la escuela es relevante en el logro alcanzado por los estudiantes. Esto se evidencia con particular claridad al contrastar los resultados obtenidos por

los estudiantes con el reporte que ellos mismos dan acerca de haber comido antes de llegar a la escuela. En este caso se encontraron diferencias estadísticamente significativas entre la proporción de estudiantes que ingirió alimentos y obtuvo el criterio de logro y la que obtuvo el criterio de logro y no se alimentó. Se aplica tanto en lectura (50.5% vs. 46.3%) como en matemática (29.4% vs. 27.1%).

En el caso del tercer grado se obtuvieron resultados similares, ya que la diferencia de la media de lectura de los estudiantes que reportaron haber comido antes de llegar a la escuela (media = 99.86; n no ponderada = 14543; DT = 20.20) y los que reportaron no haber comido (media = 104.46; n no ponderada = 14.12; DT = 19.71) es estadísticamente significativa ( $t = -4.77$ ;  $gl = 634$ ;  $p < 0.000$ ). En el caso de matemática la diferencia entre los niños que reportaron haberse alimentado (media = 100.05; n no ponderada = 14548; DT = 19.85) y los que reportaron no haberse alimentado (media = 102.97; n no ponderada = 1414; DT = 20.73) también demostró ser estadísticamente significativa ( $t = -3.23$ ;  $gl = 634$ ;  $p < 0.001$ ).

- Agrado hacia la escuela y participación de padres de familia

En el primer grado no se encontró diferencia en el logro de los grupos de estudiantes cuyos docentes reportan que los padres participan en actividades de la escuela (83.8% del total de docentes) y los grupos cuyos docentes no la reportan (11.4% del total de docentes). Por otra parte, sí se encontró una relación entre el logro y el agrado por la asistencia a la escuela reportada por los mismos estudiantes. Esto fue cierto en el caso de lectura cuando el 50.1% de los estudiantes que reportan que les agrada la escuela obtuvo el criterio de logro en comparación con el 42.1% de los que reportaron que no les agrada (Chi-cuadrado de Pearson = 12.566;  $gl = 2$ ;  $p < 0.002$ ). También en matemática, donde el 29.2% de los estudiantes que reportaron que les agradaba la escuela obtuvo el criterio de suma y resta combinada y el 21.6% de los que reportaron que no les agradaba la escuela lo obtuvo (Chi-cuadrado de Pearson = 246.623;  $gl = 2$ ;  $p < 0.000$ ).

En tercer grado sólo se indagó si la participación de los padres en el aula, reportada por los docentes, se relacionaba con el desempeño. No hubo una diferencia estadísticamente significativa en lectura ( $t = -0.984$ ;  $gl = 733$ ;  $p < 0.326$ ; se asumieron varianzas iguales) entre la media de las puntuaciones promedio de las aulas en las que los docentes reportaron participación de los padres de familia (media = 100.37; DT = 14.95) y donde no se reportó la participación de los padres de familia (media = 102.27; DT = 12.68). En matemática tampoco se encontró una diferencia estadísticamente significativa ( $t = -1.474$ ;  $gl = 733$ ;  $p < 0.141$ ; se asumieron varianzas iguales) entre la media de las puntuaciones promedio de las aulas en las que los docentes reportaron participación de los padres de familia (media = 99.65; DT = 13.37) y donde no se reportó la participación de los padres de familia (media = 102.22; DT = 12.81).

- Ambiente letrado y material

Un total de 670 directores (83.8% del total) reportó haber recibido libros de texto en el 2004. Para primer grado se comparó el promedio de los porcentajes de logro de las escuelas cuyos directores reportaron haber recibido textos (lectura: media = 49.74;  $n=670$ ;  $DT= 27.39$ / matemática: media=27.84;  $n= 670$ ;  $DT=21.59$ ) y para escuelas cuyos directores reportaron no haberlos recibido (lectura: media=45.98,  $n=107$ ,  $DT=27.61$ /matemática: media= 24.40;  $n= 107$ ;  $DT= 21.18$ ). No se encontró en este caso diferencias estadísticamente significativas para lectura (se asumieron varianzas iguales,  $t=1.317$ ,  $gl=775$ ,  $p < 0.188$ ) o matemática (se asumieron varianzas iguales,  $t=1.537$ ,  $gl=775$ ,  $p < 0.125$ ). Por otra parte, una Chi-cuadrada entre el logro alcanzado por los estudiantes y su respuesta a la pregunta de si hay libros en su casa (con base en datos ponderados), muestra relaciones significativas, aunque leves (Chi-cuadrado de Pearson = 1750.032;  $gl = 3$ ;  $p < 0.000$ ) entre contar con libros (55.3% alcanzó el criterio) y no contar con ellos (42% alcanzó el criterio). En matemática se encontró un resultado similar al tomar en cuenta el criterio de suma y resta combinado (Chi cuadrado de Pearson= 728.43;  $gl=3$ ;  $p < 0.000$  donde el porcentaje de los que reportan contar con libros en el hogar (32.4%) es superior al porcentaje de los que no lo reportan (24.4%).

En tercer grado se estimó el promedio de las puntuaciones obtenidas por los estudiantes para cada escuela. Se encontró una diferencia estadísticamente significativa en lectura ( $t = 2.040$ ;  $gl = 121.85$ ;  $p < 0.043$ ; se asumen varianzas desiguales) entre las escuelas en las cuales los directores reportaron haber recibido textos (media = 100.76;  $DT = 14.47$ ) y las escuelas en donde los directores reportaron no haberlos recibido (media = 97.04;  $DT = 17.45$ ). No se encontró una diferencia estadísticamente significativa en matemática ( $t = 0.948$ ;  $gl = 121.99$ ;  $p < 0.345$ ; se asumen varianzas desiguales) entre las escuelas en las cuales los directores reportaron haber recibido textos (media = 100.10;  $DT = 12.93$ ) y las escuelas en donde los directores no reportaron haber recibido textos (media = 98.56;  $DT = 15.54$ ).

En tercer grado, las medias de lectura de los estudiantes que reportaron que habían libros en su casa (media = 100.74;  $DT = 19.76$ ) y los que reportaron que no los había (media = 96.57;  $DT = 20.23$ ) fueron diferentes a un nivel estadísticamente significativo ( $t = 5.06$ ;  $gl = 619$ ;  $p < 0.000$ ). Las medias de lectura de los estudiantes que reportaron que alguien les leía en su casa (media = 99.29;  $DT = 19.30$ ) y los que reportaron que no les leían en su casa (media = 102.38;  $DT = 20.79$ ) fueron diferentes a un nivel estadísticamente significativo ( $t = -5.18$ ;  $gl = 611$ ;  $p < 0.000$ ).

En tercer grado, las medias de matemática de los estudiantes que reportaron que habían libros en su casa (media = 100.61;  $DT = 19.87$ ) y los que reportaron que no los había (media = 97.68;  $DT = 20.23$ ) fueron diferentes a un nivel estadísticamente significativo ( $t = 3.97$ ;  $gl = 619$ ;  $p < 0.000$ ).

Las medias de matemática de los estudiantes que reportaron que alguien les leía en su casa (media = 99.47; DT = 19.41) y los que reportaron que no les leían en su casa (media = 101.89; DT = 20.93) fueron diferentes a un nivel estadísticamente significativo ( $t = -4.16$ ;  $gl = 610$ ;  $p < 0.000$ ).

- Riesgos y/o violencia en la escuela

En primer grado se encontró una correlación estadísticamente significativa de sentido negativo entre las preguntas que hacen referencia a riesgo y violencia y el logro del criterio conjunto (suma y resta) de matemática (a mayor riesgo y violencia, menor desempeño) (R biserial puntual = -0.012;  $p < 0.000$ ). Esta relación no se observó en el caso de lectura (R biserial puntual = 0.002;  $p < 0.582$ ).

En el caso de tercer grado se descubrió una relación negativa estadísticamente significativa en lectura ( $r = -0.106$ ;  $p < 0.000$ ) y matemática ( $r = -0.094$ ;  $p < 0.000$ ) la cual indica que al incrementar la puntuación de eventos de riesgo o violencia, es menor el rendimiento de los estudiantes.

- Experiencia docente

Se observó que el promedio de años de experiencia de los docentes del área urbana es mayor que el de los docentes del área rural (14 y 7 años respectivamente). Por esta razón, no es apropiado concluir que la experiencia docente causa por sí sola un incremento en el desempeño de los estudiantes. Aún así es significativo que hubo una relación entre ambas variables.

En primer grado se observó que el logro en lectura se correlacionó de forma positiva y estadísticamente significativa con los años de experiencia del docente (R de Pearson = 0.229;  $p < 0.000$ ) y con los años de experiencia específica en primer grado primaria (R de Pearson = 0.115;  $p < 0.001$ ). Lo mismo se encontró en matemática, donde la relación del logro en la prueba de matemática se correlacionó de forma positiva y significativa con los años de experiencia docente (R de Pearson = 0.140;  $p < 0.000$ ) y los años de experiencia en primer grado (R de Pearson = 0.081;  $p < 0.021$ ).

En tercer grado también se encontró una correlación positiva y estadísticamente significativa entre el rendimiento promedio en las escuelas y el número de años de experiencia que el docente reporta. Esto fue así tanto en lectura ( $r = 0.341$ ;  $p < 0.000$ ) como en matemática ( $r = 0.296$ ;  $p < 0.000$ ). También se encontró una relación significativa, aunque menor que en el caso anterior, entre el número de años que el docente había trabajado en el tercer grado y el desempeño en lectura (0.087;  $p < 0.015$ ). No se encontró una relación estadísticamente significativa entre el número de años de experiencia en tercer grado reportados por el docente y el desempeño en matemática ( $r = 0.030$ ;  $p < 0.407$ ).

- Infraestructura

En primer grado se indagó en el cuestionario al director acerca de la presencia de agua entubada, electricidad y servicios sanitarios. Se encontró una correlación estadísticamente significativa entre el logro en lectura y la presencia de agua entubada ( $R$  biserial puntual = 0.200;  $p < 0.000$ ) y electricidad ( $R$  biserial puntual = 0.217;  $p < 0.000$ ). En el caso de sanitarios no se encontró relación ( $R$  biserial puntual = 0.068;  $p < 0.056$ ). También se encontró una correlación estadísticamente significativa entre el logro en matemática y la presencia de agua entubada ( $R$  biserial puntual = 0.146;  $p < 0.000$ ), electricidad ( $R$  biserial puntual = 0.171;  $p < 0.000$ ), pero igual que en el caso de lectura no se encontró con la presencia de sanitarios ( $R$  biserial puntual = 0.020;  $p < 0.582$ ).

En tercer grado los tres aspectos mostraron ser significativos. En lectura la puntuación se relacionó con la presencia de agua entubada ( $R$  biserial puntual = 0.174;  $p < 0.000$ ), electricidad ( $R$  biserial puntual = 0.290;  $p < 0.000$ ) y sanitarios ( $R$  biserial puntual = 0.080;  $p < 0.026$ ). En matemáticas también se relacionó el logro con la presencia de agua entubada ( $R$  biserial puntual = 0.128;  $p < 0.000$ ), electricidad ( $R$  biserial puntual = 0.260;  $p < 0.000$ ) y sanitarios ( $R$  biserial puntual = 0.092;  $p < 0.011$ ).

## Conclusiones

- a) El área urbana mostró una ventaja en el rendimiento sobre el área rural. La proporción de estudiantes del primer grado que alcanzaron el criterio en el área urbana fue en general superior a la proporción de estudiantes que lo alcanzaron en el área rural. Esto se encontró en lectura (61% vs 45%) como en matemática (36% vs 26%). También, se observó esta diferencia en tercer grado, tanto en lectura (media de 111.25 vs 96.80) como en matemática (media de 108.28 vs 97.64).
- b) La población no indígena mostró una ventaja sobre la población indígena. Las diferencias fueron estadísticamente significativas y se observó en primer grado en lectura (50% logra el criterio de la población no indígena vs el 43% de la población indígena) y matemática (28% de la población no indígena logró el criterio vs 26% de la población indígena). También, se observó en tercer grado en lectura (media de 105.86 para la población no indígena vs 89.89 para la población indígena) y matemática (media de 103.79 para la población no indígena vs 93.33 para la población indígena).
- c) Los niños mostraron una ventaja en rendimiento sobre las niñas. Esto se observó tanto en primer grado en lectura (48% de los niños logra el criterio vs 47.7% de las niñas) y matemática (28% de los niños logra el criterio vs

27% de las niñas), como en tercer grado en lectura (media de 100.11 para los niños vs 99.84 para las niñas) y matemática (media de 101.96 para los niños vs 97.89 para las niñas). Estas diferencias, sin embargo, no mostraron el grado de significancia necesario para formar parte de un modelo explicativo del desempeño de ambos sexos.

- d) El análisis de las sub-secciones de la prueba de lectura del primer grado de primaria indicaron que en las tareas que requirieron de destrezas de decodificación y asociación de fonema grafema (asociación sonido-letra), la proporción de estudiantes que logró el criterio (78.98% para asociación sonido-letra y 66.85 para asociación palabra escrita y dicha), fue superior a la proporción de estudiantes que obtuvieron el criterio en tareas que, además de la decodificación, requirieron de estrategias de pensamiento complejo (32.35% para completación de frases y 25.13% para responder preguntas con base a la información de un párrafo). El criterio, que requería seguimiento de instrucciones enunciadas en oraciones de cinco palabras fue alcanzado por la mitad de los estudiantes (50.96%).
- e) Al analizar las sub-secciones de la prueba de matemática del primer grado primaria se observó que tan sólo una cuarta parte de los estudiantes reconocieron los números Mayas (24.12%). El criterio, que requiere realizar sumas sin llevar y restas sin prestar, fue alcanzado por un porcentaje menor a la mitad de los estudiantes, aún cuando se analizaron sin tomar en cuenta el desempeño en los ítems anteriores (47.45% para sumas y 40.66% en restas). Cuando se requiere de estrategias de pensamiento más complejas, el porcentaje de estudiantes que logró los criterios se redujo dramáticamente (26.89% logró identificar el lugar que ocupa una figura en una serie y el 12.80% completó una serie de números).
- f) En el análisis de las sub-secciones de la prueba de lectura de tercer grado se observó que tanto en vocabulario como en comprensión, la población rural indígena y femenina obtuvo medias inferiores a las de la población urbana, no indígena y masculina (97.13, 90.81 y 99.77 vs 110.15, 105.56 y 100.22 en vocabulario; 96.97, 90.55 y 99.97 vs 110.64, 105.26 y 100.03 en comprensión).
- g) Al analizar las sub-secciones de la prueba de matemática de tercer grado se observó que en todas ellas (operaciones, conceptos aritméticos y resolución de problemas), las medias para operaciones fueron menores para la población rural, indígena y femenina que en el caso de la población urbana, no indígena y masculina (98.40, 96.01 y 98.67 vs 105.99, 102.50 y 101.39); para conceptos aritméticos 97.72, 92.77 y 98.43 vs 108.72, 104.15 y 101.78) y para resolución de problemas (98.70, 95.99 y 98.07 vs 104.87, 102.43 y 101.93).
- h) En las escuelas que sí ofrecen educación bilingüe en primer grado y fueron evaluadas en esta muestra, se observó que en lectura una mayor proporción de estudiantes logró los criterios en la prueba en español que en la prueba en k'iche' o kaqchikel (61.99% vs 55.69% para los que tomaron la prueba en k'iche' y 54.91% vs 56.95% para kaqchikel). Sin embargo, es importante notar que en ambos casos la proporción de estudiantes que logró el criterio

en la prueba en español fue superior a la proporción que para todo el país logró el criterio (61.99% y 54.91% vs el 48% de toda la muestra). Es importante tomar en cuenta que esto no refleja el estado general de todo el sistema de educación bilingüe. Sin embargo, sí apoya la hipótesis de que cuando efectivamente se ofrece educación bilingüe a estudiantes cuyo idioma materno no es el español, el aprendizaje de la lecto-escritura en español se ve reforzado.

- i) En primer grado también se aplicaron pruebas de matemática en español y también en k'iche o kaqchikel de acuerdo al idioma materno de los estudiantes. En este caso se observó un fenómeno similar al obtenido en la prueba de lectura, donde los resultados de la prueba de español para el grupo que tomó la prueba en ambos idiomas, fue superior al presentado en la prueba de lectura a nivel nacional. Sin embargo, las conclusiones deben ser tomadas con cautela, pues las pruebas en k'iche' y kaqchikel eran traducciones de la prueba en español. Considerando que la aplicación de los instrumentos en español se realizó durante el primer día, es de esperar que los resultados en las evaluaciones del segundo día fueran superiores a los del primer día ya que las operaciones, por ser expresada en número, no difirieron de las utilizadas en la evaluación del primer día.
- j) En el tercer grado hubo un grupo de estudiantes que tomó pruebas en idioma español y también en k'iche, kaqchikel, q'eqchi' o mam de acuerdo al idioma de la comunidad. En estos casos se realizaron comparaciones por género sin que se observara consistentemente una ventaja a favor del género masculino. En el caso de las evaluaciones en k'iche' (media masculina de 97.03 y femenina de 103.33) y mam (media masculina de 99.76 y femenina de 100.22) el resultado femenino fue superior al masculino. En el caso de matemática el resultado sí fue consistentemente superior en el caso de los niños. En todos los casos, menos en el de la prueba mam de lectura, se obtuvo una correlación estadísticamente significativa entre los resultados obtenidos por los estudiantes en las pruebas en español y en k'iche', kaqchikel, q'eqchi' o mam.
- k) Tanto en los instrumentos de lectura como de matemática se encontró que existen dificultades cuando los ítems requieren de estrategias complejas de pensamiento. Es necesario que estas áreas sean reforzadas y que se provea de lineamientos a los docentes para fortalecer destrezas de este tipo entre los estudiantes.
- l) A pesar de que existe un interés por el aprendizaje de aspectos de la cultura maya y se incluyen en el currículo del primer grado, tres cuartas partes de los estudiantes no fueron capaces de identificar la numeración maya. Si este objetivo es planteado en el currículo, los docentes deberán prestar particular atención a este aspecto.
- m) Los resultados para el primer grado muestran que la población evaluada que efectivamente se encuentra en programas de educación bilingüe, el porcentaje de estudiantes que logra el criterio en la prueba de español es superior que en la población general. Esta conclusión, por estar basada en una evaluación que se realizó por primera vez, debe ser tomada como generadora de hipótesis y no como un hecho constatado. Estos resultados

no se replican en tercer grado, por lo que también es necesario explorar con mayor detenimiento cómo se relacionan los beneficios de la inversión en la educación bilingüe con el progreso de los estudiantes a lo largo de los demás grados de la escuela primaria.

- n) Los resultados de esta evaluación no reflejan un punto final y definitivo. Debería, idealmente, considerarse como un punto de partida para el mejoramiento continuo. Es necesario que se realicen futuros esfuerzos de evaluación del rendimiento que tomen en cuenta los resultados obtenidos en el pasado y de esta manera obtener un gradual incremento del conocimiento del rendimiento de los estudiantes guatemaltecos, lo cual permitiría definir las políticas y estrategias que se requieran.
- ñ) La alimentación mostró relacionarse de forma significativa con el incremento en el desempeño. El porcentaje de estudiantes que alcanzó el criterio y que reportó haber tomado su desayuno antes de asistir a la escuela fue superior al porcentaje de estudiantes que reportó no haber desayunado. En primer grado se aplica tanto en lectura (50.5% vs 46.3%) como en matemática (29.4% vs 27.1%). De igual manera, se observó en tercero para lectura (105 vs 100) y matemática (103 vs 100).
- o) La aceptación y el agrado expresado por los estudiantes hacia la escuela fue un factor relevante en el logro del primer grado. Tanto para el primer grado en lectura (50.1% vs 42.1%), como en matemática (29.2% vs 21.6%). Un mayor porcentaje que indicó que le agradaba la escuela logró el criterio en comparación con el que reportó no sentir agrado hacia ésta. En tercer grado no se indagó esta variable.
- p) Contar con material de lectura fue significativo tanto para el primer grado como para el tercero, aunque parece ser más sólido para este último. En primero, el porcentaje de estudiantes que obtuvo el logro en lectura y matemática y reportó tener material de lectura en casa (53.3% y 32.4% respectivamente) fue superior al de estudiantes que reportaron no contar con éste (42.8% y 32.4%). En tercer grado, las medias de lectura y matemática cuando hubo material de lectura en casa (100.74 y 100.61) fueron superiores al caso en que no los hubo (96.57% y 97.68).
- q) La percepción de riesgo y violencia influyó en los resultados de matemática del primer grado y los de matemática y lectura en tercero. Se encontró en ambos casos correlaciones significativas.
- r) La experiencia docente influyó tanto en lectura como en matemática en primer y tercer grado. Se observaron correlaciones significativas entre años de experiencia y el desempeño en lectura y matemática de primer y tercer grado con un alto nivel de significancia.
- s) La presencia de agua entubada y electricidad mostró relación con el porcentaje de estudiantes que alcanzaron el criterio de logro en ambos grados. La presencia de sanitarios asociada al desempeño de estudiantes sólo mostró relación en tercer grado.



M.Sc. Francisco José Ureta  
fureta@pronere.uvg.edu.gt

Director del Centro de Investigaciones Educativas del Instituto de Investigaciones de la Universidad del Valle de Guatemala



M.Sc. Álvaro Mauricio Fortín  
alvarofortin@yahoo.com

Colaborador, Director del Sistema Nacional de Evaluación e Investigación Educativas, Ministerio de Educación



Dr. Roberto Molina Cruz  
rmolina@uvg.edu.gt

Catedrático del Departamento de Matemática de la Facultad de Ciencias y Humanidades de la Universidad del Valle de Guatemala

## Bibliografía

1. Crocker, L. & Algina, J. *Introduction to classical & modern test theory* Harcourt Brace Jovanovich College Publishers, New York, 1986
2. Gronlund, N.E. *How to make achievement tests and assessments* (5th ed) Allyn and Bacon. Boston, 1993
3. Nunnally, J.C. *Teoría psicométrica* McGraw-Hill, Mexico, 1995
4. Fortín, A. & Ureta, F. *Informe Nacional de la Evaluación del Rendimiento en Lectura y Matemática de Estudiantes del Primer Grado de la Primaria de Instituciones Oficiales de Guatemala* USAID, MEDIR, PRONERE, Universidad del Valle de Guatemala, 2005
- 5) Fortín, A. & Ureta, F. *Informe Nacional de la Evaluación del Rendimiento en Lectura y Matemática de Estudiantes del Tercer Grado de la Primaria de Instituciones Oficiales de Guatemala* USAID, MEDIR, PRONERE, Universidad del Valle de Guatemala, 2005
- 6) Fortín, A. & Ureta, F. *Especificación de ítems de las pruebas de tercer grado primaria Guatemala*, USAID, MEDIR, PRONERE, Universidad del Valle de Guatemala, 2005
- 7) Molina, R. et al. *Diseño de la encuesta y la obtención de la muestra 2004* USAID, MEDIR, PRONERE, Universidad del Valle de Guatemala, 2005a
- 8) Molina, R. et al. *Evaluación de la muestra realizada y cálculo de los expansores*. USAID, MEDIR, PRONERE, Universidad del Valle de Guatemala, 2005b