



Medición del efecto del programa Singapur en alumnos de 2° básico*
Julio de 2013

Contenidos

- Resumen
- Antecedentes
 - Antecedentes que originan el programa
 - Descripción y objetivos del programa
 - Selección de establecimientos del programa
- Servicio de medición de la efectividad en los aprendizajes del Programa Singapur
- Selección de establecimientos para la evaluación
- Efectividad en los aprendizajes del Programa Singapur
 - Descripción de la prueba aplicada: aspectos técnicos
 - Aplicación de la prueba: aspectos operativos
- Evaluación de impacto
 - Análisis del cumplimiento de condiciones para la evaluación de impacto
 - Estimación del impacto a nivel de alumnos
 - Estimación por MCO con datos agrupados (*clusters*)
 - Análisis Multinivel
 - Estimación del impacto a nivel de establecimientos
 - Estimaciones por grado de dificultad y eje temático de las preguntas
- Conclusiones
- Anexos
- Bibliografía

* Minuta Técnica elaborada por Mario Rivera Cayupi, Unidad de Evaluación, DIPLAP.

Resumen

El siguiente informe contiene los resultados de la medición del efecto del programa "Textos Singapur" en estudiantes de 2° básico de una muestra de 143 establecimientos seleccionados aleatoriamente de un total de 482, mediante la aplicación de una prueba estandarizada, piloteada y validada de matemáticas enmarcada en las Bases Curriculares de Educación Básica según el Decreto Supremo N° 439/2012. Se efectúan estimaciones para la muestra completa a nivel de estudiantes y de establecimientos, mediante regresiones con datos organizados en *clusters* y modelos multinivel y se efectúan estimaciones organizando los datos en distintas sub muestras de interés. Se encuentra que el programa tuvo un efecto estadísticamente significativo de entre 2,02% y 2,99% en el rendimiento en la prueba aplicada en favor de los niños y establecimientos acogidos al programa. Al considerar el grado de dificultad de las preguntas y los ejes temáticos a los cuales pertenecen, y analizar los resultados por género de los alumnos y la dependencia administrativa de los establecimientos, se observa un mayor efecto del programa en las respuestas a preguntas cuyo grado de dificultad es intermedio, de entre 2,6% y 4% a favor de los participantes en el programa y de 5,4% a nivel de establecimientos particulares subvencionados del programa. Por eje temático, el efecto del programa es diverso, siendo mayor en algunos ejes para alumnos de establecimientos participantes en el programa y en otros para los alumnos de establecimientos no participantes. Por género también se observa que el efecto en el porcentaje de respuestas correctas de las mujeres es mayor en algunos ejes y en otros no. Por dependencia, el efecto del programa en el porcentaje de respuestas correctas es mayor en cuatro de los seis ejes temáticos a favor de los alumnos de establecimientos particulares subvencionados participantes en el programa, con diferencias significativas que van desde 4,46% en patrones y álgebra hasta 10,44% en datos y probabilidad.

El informe está organizado de la siguiente manera. Primero se describen los antecedentes del programa. Luego se describen los antecedentes del servicio contratado por el Ministerio y la forma en que fueron seleccionados los establecimientos para la evaluación. En seguida se describen aspectos técnicos del instrumento aplicado y aspectos operativos de la evaluación ejecutada, del cual surgen los datos utilizados para la presente evaluación. Luego se aborda la medición misma del efecto del programa, describiendo tanto las metodologías de evaluación como los resultados obtenidos y finalmente se plantean las conclusiones a partir de los hallazgos más relevantes.

– Antecedentes

○ Antecedentes que originan el programa

Existe evidencia que da cuenta del insuficiente rendimiento de los estudiantes chilenos en matemática en comparación a niños de países de similar ingreso relativo y de la deficiente formación de los profesores en la materia.

Las estadísticas revelan que los estudiantes de Singapur se han ubicado sistemáticamente entre los primeros lugares del ranking de países que rinden la prueba TIMMS de matemáticas. En 1999 se ubicaron en el primer lugar en la prueba de 8° grado, en 2003 en el primer lugar en las pruebas de 4° y 8° grado, en 2007 en el segundo lugar en la prueba de 4° grado y en el tercer lugar en la de 8° grado y en

2011 en el primer lugar en la prueba de 4° grado y en el segundo lugar en la de 8° grado¹.

En el caso de Chile, en la prueba de matemáticas de 8° básico de 1999, el país se ubicó en el lugar 35 entre 38 evaluados, con un puntaje promedio de 392 puntos, mientras que el promedio internacional fue de 487 puntos (Mullis et al. 2000); en la medición de 2003 se ubicó en el lugar 39 entre 45 países evaluados, con un puntaje promedio de 387 puntos, siendo el promedio internacional de 467 puntos (Mullis et al. 2004) y en la evaluación de 2011 el puntaje aumentó 29 puntos en relación a 2003, alcanzando los 416 puntos (más de 0,25 desviaciones estándares), pero aun manteniéndose en los niveles de desempeño bajo². El resultado de los alumnos de 4° básico que rindieron la prueba TIMSS de matemáticas en 2011 fue de 462 puntos, puntaje que también corresponde a un nivel de desempeño bajo, a casi 40 puntos del centro de la escala TIMSS³.

Por su parte, la formación inicial de los profesores con mención en matemáticas es aún deficiente, lo que queda reflejado en la aplicación de la prueba Inicia que, en 2009, midió los conocimientos disciplinarios de educación básica de profesores con mención en matemáticas recién egresados en 3 ejes. En el eje 1 –orientación en el espacio y geometría–, un 91% de ellos tuvo un logro igual o inferior al 50% de la prueba; en el eje 2 –número y sus aplicaciones–, el 80% de los evaluados tuvo un logro de 50% o menos y en el eje 3 –números y sus operaciones–, el 83% de los egresados obtuvo un logro de un 50% o menos⁴.

o **Descripción y objetivos del programa**

Como una manera de enfrentar este problema, el Ministerio de Educación seleccionó y ofreció a 300 establecimientos municipalizados y particulares subvencionados del país la posibilidad de recibir la serie de libros de matemática “Pensar Sin Límites”, también conocidos como “Textos Singapur”, traducidos y adaptados al currículum nacional, con los materiales didácticos que los acompañan, en lugar de los textos tradicionales que entrega el Ministerio a todos los establecimientos subvencionados del país anualmente. Estos 300 establecimientos correspondían a aquellos con mejores puntajes promedio de las últimas tres evaluaciones SIMCE de matemática en 4° básico (entre 2007 y 2009), según región, dependencia y grupo socioeconómico y componían una matrícula total de 40.200 estudiantes en 1° y 2° básico.

¹ Cf. “*International mathematics report*”, años 1999, 2003, 2007 y 2011, publicados por TIMSS & PIRLS, International Study Center, Lynch School of Education, Boston College.

² Cf. “Resultados TIMSS 2011. Estudio Internacional de Tendencias en Matemáticas y Ciencias”. Informe preparado por la Agencia de Calidad del Ministerio de Educación y TIMSS PIRLS International Study Center, p. 11.

³ El centro de la escala TIMSS se define como el promedio internacional estandarizado del conjunto de países que participó en el primer ciclo de TIMSS de ciencias y matemáticas en 1995 y que, en una escala de 0 a 1.000 puntos para cada prueba, corresponde a 500 puntos, con una desviación estándar de 100 puntos.

⁴ Los datos pueden revisarse en la página web del Centro de Estudios de MINEDUC. En las bases de datos, ingresar a personal, donde se debe seleccionar la opción docentes y de allí los datos de la evaluación Inicia. Ver <http://centroestudios.mineduc.cl/index.php?t=96&i=2&cc=2036&tm=2>.

Si bien existió una preselección, la participación final de los establecimientos en el programa fue voluntaria y aquellos que decidieron no participar en el proyecto fueron sustituidos utilizando una muestra de reemplazo previamente seleccionada.

El programa entregó la serie de "Textos Singapur" a inicios del año escolar 2011, para los estudiantes de 1° y 2° básico⁵. Su diseño original contempla que el programa dure hasta el año 2013, año en el que se habrá completado la entrega de los textos escolares a los estudiantes de 1° a 4° básico de una misma escuela.

El Objetivo General del programa es mejorar el aprendizaje en matemáticas de los estudiantes mediante una estrategia de apoyo curricular y docente basado en el Sistema Singapur para la enseñanza de las matemáticas y sus Objetivos Específicos son:

- Introducir nuevas metodologías y estándares en la enseñanza de las matemáticas para la mejora en los aprendizajes.
- Apoyar a un conjunto de establecimientos para avanzar en el aprendizaje de matemáticas.
- Elevar el estándar de los textos de matemáticas que se utilizan y de las metodologías de enseñanza de las matemáticas.
- Estudiar la viabilidad de implementar textos más extensos y exigentes en el ritmo de aprendizaje de matemáticas.
- Obtener conclusiones relevantes respecto de la posibilidad de hacer extensivo el programa al resto de la población escolar.
- Extraer lecciones para el diseño y licitación de los futuros libros de textos entregados por el Mineduc.

- **Selección de establecimientos del programa**

Para identificar y posteriormente seleccionar los establecimientos elegibles para la entrega de los libros, el Ministerio procedió de acuerdo a los siguientes criterios:

- Los establecimientos debían ser municipales o particulares subvencionados.
- Debían contar con evaluaciones SIMCE en la prueba de Matemáticas de 4° básico en 2007, 2008 y 2009, alcanzando 250 puntos promedio o más en dichas evaluaciones.
- Debían contar con 20 alumnos o más en ese nivel.
- El conjunto de establecimientos seleccionados debía replicar la distribución nacional por grupo socioeconómico (GSE⁶) y dependencia.
- El total de alumnos, considerando todos los establecimientos seleccionados, debía ser de 20.000 por nivel.

⁵ A fines de 2011 y principios de 2012 se ejecutó una evaluación de implementación del programa, cuyos principales resultados se resumen en el Anexo N° 1.

⁶ El Ministerio de Educación clasifica a los establecimientos que rinden la prueba SIMCE según características socioeconómicas predominantes de sus alumnos, definidas como una combinación de los años de estudio promedio de los padres y de las madres, el ingreso familiar mensual y el índice de vulnerabilidad escolar del establecimiento calculado por la JUNAEB. Se utiliza la siguiente escala: A (bajo), B (medio bajo), C (medio), D (medio alto) y E (alto).

Antes de seleccionar los establecimientos, con la información disponible (Base matrícula 2010 y resultados SIMCE 2007, 2008 y 2009 para 4° básico) y los criterios descritos, fue necesario realizar las siguientes precisiones en algunas de las especificaciones:

- Los puntajes SIMCE considerados correspondían al subsector de Educación Matemática de las pruebas 2007, 2008 y 2009.
- Los establecimientos debían tener resultados en las 3 aplicaciones SIMCE mencionadas o, al menos, en las 2 últimas (2008 y 2009).
- Los establecimientos debían tener 20 alumnos o más en promedio en las 3 evaluaciones y en la matrícula de 1° básico 2010.
- Dados los criterios de selección no fue posible replicar la distribución nacional por GSE, aunque sí por dependencia (Municipalizado y Particular Subvencionado), por lo cual se respetó la distribución por dependencia y se verificó la existencia en la selección de establecimientos de todos los GSE.
- Considerando la posibilidad de que algunos establecimientos no participaran del proyecto, se incluyeron establecimientos de reemplazo (30% del total de la muestra).
- Por último se especificó que el 40% de los establecimientos seleccionados debía ubicarse en la región Metropolitana y el 60% restante debe distribuirse en el resto del país sin restricciones.

A partir de lo anterior es evidente que el conjunto de establecimientos elegibles para participar en el programa no correspondía a una muestra aleatoria, porque el programa mismo definía un conjunto de criterios que hicieron que los establecimientos no tuvieran igual probabilidad de participar en él. De hecho, el programa se focalizó en los establecimientos municipales y particulares subvencionados con mejores resultados en las pruebas SIMCE de matemáticas entre los años 2007 a 2009, los que fueron ordenados por su puntaje promedio, seleccionando, en cada estrato (grupo socio económico, dependencia y región) el número de establecimientos que permitiera cumplir los criterios antes mencionados.

La selección operó de la siguiente manera: habiendo identificado a los establecimientos que cumplían los requisitos básicos para ser elegibles (básicamente puntaje promedio SIMCE en las últimas evaluaciones de matemáticas sobre 250 puntos y tener más de 20 alumnos por establecimiento), se generó un marco muestral de más de 3.000 establecimientos. A ellos se les aplicaron los otros criterios (el 40% del total debía pertenecer a la Región Metropolitana, debían respetar la proporción de establecimientos municipales y particulares subvencionados, etc.) para determinar, finalmente, quiénes serían invitados al programa y quiénes serían considerados como reemplazos de aquellos que no pudieran o no quisieran participar en él.

La base de datos final contó con 482 establecimientos elegibles. Por razones presupuestarias, 296 de ellos fueron seleccionados aleatoriamente como establecimientos acogidos al programa (61%) y 186 fueron seleccionados aleatoriamente como establecimientos de reemplazo (39%). El Anexo 2 contiene tablas que permiten verificar el cumplimiento de los requisitos de selección para participar en el programa por parte de los establecimientos elegibles.

– **Servicio de medición de la efectividad en los aprendizajes del Programa Singapur**

En el marco anterior, el Ministerio de Educación requirió la contratación de un servicio consistente en la elaboración y posterior aplicación de una prueba estandarizada, piloteada y validada de matemáticas, dirigida a estudiantes de 2° Básico de una muestra aleatoria de establecimientos tomada de los 482 establecimientos elegibles, y en la construcción de una base de datos que incorporara la información recogida.

Para ello se solicitó, además, la elaboración de un informe de resultados por cada uno de los establecimientos sometidos a la evaluación, indicando promedios generales de la unidad educativa, la significancia estadística de los resultados y los niveles de logro para la prueba, desglosados por eje temático y habilidades de acuerdo al decreto 439 del año 2012. Como parte del mismo servicio se requirieron además, los resultados particulares de cada curso y los resultados comparativos con las demás unidades educativas que rindieron la prueba.

La ejecución de este servicio fue adjudicado al Centro de Investigación, Experimentación y Transferencia en Didáctica de la Matemática y la Ciencia, de la Universidad de Santiago, en lo sucesivo Centro Félix Klein o CFK.

Se debe tener presente que el servicio contratado corresponde a una evaluación de corto plazo, considerando el escaso tiempo transcurrido entre la aplicación del programa y el período de evaluación, sobre todo porque en el ámbito de la educación usualmente los procesos o programas toman algún tiempo en madurar y mostrar sus efectos. De hecho, según el Centro Félix Klein (2012), los textos del método Singapur de 1° y 2° básico no resultarían disruptivos respecto de la cultura escolar instalada en este sub-ciclo de la enseñanza, incorporando contenidos y modos de enseñanza relativamente cercanos para los profesores que hacen clases en estos niveles. Sin embargo, los textos de 3° y 4° básico incorporarían claras diferencias respecto de la práctica habitual de la enseñanza en estos cursos, resultando más complejos y distintos de lo que los profesores de estos cursos están acostumbrados a enseñar, tanto desde la perspectiva matemática como didáctica. De allí que la aplicación de la prueba SIMCE de matemáticas de 4° básico en el año 2014 constituirá una buena oportunidad de medir el efecto del programa en el aprendizaje de los niños, pero con más tiempo de maduración del mismo, determinando entonces si el efecto del programa es más nítido o no.

Además, se debe tener presente que es probable que el desarrollo de los ejes temáticos haya sido distinto entre los establecimientos que aplicaron el método Singapur y los demás establecimientos, lo cual puede haber incidido en el impacto del programa en los resultados de los alumnos según eje temático, cuestión que se evalúa más adelante.

– **Selección de establecimientos para la evaluación**

Para determinar la cantidad de establecimientos a evaluar y, a partir de ellos, de los alumnos a los que se debía aplicar la prueba, se trató el problema como un ensayo aleatorio por *clusters* en dos niveles. Como parámetros para el cálculo se consideró un nivel de significancia de 0,05 y potencias de 0,8 y de 0,9. Se supuso que los resultados promedio de las pruebas SIMCE anteriores de matemáticas a nivel de establecimiento,

deberían ser capaces de explicar (relacionarse estadísticamente) entre un 60% y un 70% con los resultados de la prueba estandarizada aplicada mediante esta evaluación. Se determinó la cantidad de establecimientos estimando el efecto tamaño del tratamiento en 0,15, 0,2 y 0,25 desviaciones estándares. El promedio de estudiantes de 2° básico en el conjunto de establecimientos es de 71 alumnos. Finalmente, se supuso que la correlación intra clase de los alumnos de 2° básico es igual a la de 4° básico (calculada utilizando los resultados de la prueba SIMCE de matemáticas de 4° básico de 2011) que es de 0,21, considerando los establecimientos tratados y de control⁷. Utilizando estos parámetros se estimó⁸ la cantidad de establecimientos que se requerirían en cada escenario. Los resultados se muestran en la Tabla 1:

Tabla 1: Estimaciones de establecimientos a evaluar según potencia y efecto tamaño (d)

Potencia	R ²	d=0,15	d=0,2	d=0,25
0,8	0,6	137	78	50
	0,7	107	61	39
0,9	0,6	179	102	66
	0,7	141	83	55

Finalmente, en función de estos parámetros y en virtud del presupuesto disponible, se seleccionaron 141 establecimientos del escenario con el efecto tamaño más conservador (0,15 desviaciones estándares), pero exigiendo un poder más alto (0,9), asumiendo una correlación 0,7 entre las pruebas. Se efectuó un ejercicio de aleatorización en base a las proporciones de establecimientos tratados (61%) versus controles (39%).

De ese modo, a CFK se le solicitó aplicar la prueba en, al menos, 141 establecimientos, aunque se le entregó un listado con más establecimientos, seleccionados aleatoriamente, para abordar el reemplazo por no respuesta, no contacto, etc. CFK aplicó la prueba finalmente en 143 establecimientos (ver el listado en el Anexo N° 3). Las estimaciones y los análisis que siguen se efectúan sobre esta cantidad de establecimientos. En el Anexo N° 4 se verifica que los establecimientos seleccionados para la evaluación también cumplen con los requisitos de elegibilidad del programa.

– Efectividad en los aprendizajes del Programa Singapur⁹

- o Descripción de la prueba aplicada: aspectos técnicos

El enfoque metodológico utilizado por el CFK para la construcción, análisis y obtención de los resultados de esta prueba se basó principalmente en Chevallard (1999)¹⁰,

⁷ Se define ρ , la correlación intra clase, como la razón entre la variabilidad entre clusters (τ) y la variabilidad total, compuesta por la variabilidad entre clusters más la variabilidad intra clusters (σ^2): $\rho = \tau / (\tau + \sigma^2)$.

⁸ Para la estimación se utilizó el programa *Optimal Design 3.0*.

⁹ Esta sección se basa en Centro Félix Klein (2012), documento que contiene detalles del servicio de diseño y aplicación del instrumento solicitado por Mineduc.

¹⁰ Chevallard, Y. (1999). *L'analyse des pratiques enseignantes en théorie anthropologique du didactique*. Recherches en Didactique des Mathématiques 19 (2), 221-266.

Brousseau (2007)¹¹, el marco teórico contenido en Espinoza (2008)¹² y en el marco teórico desarrollado por PISA para matemática (OCDE, 2006)¹³.

La construcción del instrumento utilizado consideró las características señaladas en la literatura antes citada, utilizando la noción de habilidad matemática para construir la tabla de especificaciones y los ítems que constituyen el instrumento, por medio de la siguiente estrategia metodológica: para cada tarea matemática nuclear considerada en la tabla de especificaciones de la prueba, se asoció una habilidad matemática que predominantemente utilizan los estudiantes al abordar el ítem de evaluación, así como el nivel de complejidad en que lo hacen. Además, se incluyó el nivel de logro al que responde dicha tarea. Los resultados de la evaluación y sus posteriores análisis se efectuaron considerando como criterio las tareas matemáticas, que a su vez respondían a los objetivos de aprendizaje o bien las habilidades matemáticas asociadas.

Uno de los requerimientos realizados por el Mineduc en la entrega de los resultados fue establecer la cantidad de estudiantes que se encontraban en niveles de logro inicial, intermedio y avanzado. Si bien el SIMCE describe estos niveles a partir de las competencias que un estudiante al finalizar el primer ciclo básico, no están definidos para 2° Básico, por lo cual se debieron construir los niveles de logro para este nivel, partiendo del análisis de las bases curriculares 2012. En cada eje temático, los expertos del Centro Félix Klein, establecieron las tareas matemáticas que los estudiantes debían resolver en cada nivel.

Posteriormente se diseñó la prueba teniendo en cuenta ésta primera definición. Una vez diseñada se retomó el trabajo de construcción de los niveles de logro. En esta etapa se estableció un orden de los ítems por grado de dificultad. Este ordenamiento se realizó, en primera instancia con los resultados obtenidos en el pilotaje y, posteriormente, se sometió a juicio de expertos para concordar el orden definitivo. Una vez que los expertos estuvieron de acuerdo en dicho orden, se establecieron los puntos de corte para cada nivel de logro. Es decir, se definió hasta qué ítem debía ser capaz de resolver cada estudiante para lograr las competencias de cada nivel.

La selección de los ítems de la prueba se realizó utilizando el banco de ítems de evaluación de aprendizajes y competencias matemáticas de 2° Básico con el cual cuenta el CFK, los cuales han sido sometidos a un proceso de validez y confiabilidad a través de juicio de expertos y análisis estadísticos. Los ítems de evaluación fueron calibrados utilizando la teoría de respuesta al ítem (IRT) con un modelo de tres parámetros: dificultad, discriminación y adivinanza. Una vez establecidos los niveles de logro, del banco de ítems se seleccionaron aquellos para segundo básico que evaluaban los objetivos de aprendizaje de las bases curriculares 2012.

¹¹ Brousseau, G (2007). *Iniciación al estudio de la teoría de las situaciones didácticas*. Primera edición, Libros del Zorzal, Buenos Aires.

¹² Espinoza, L. et al. (2009). *Análisis de las competencias matemáticas en NB1. Caracterización de los niveles de complejidad de las tareas matemáticas*. Universidad de Santiago de Chile, Proyecto FONIDE N°: DED0760.

¹³ OCDE (2006). *Marco de la evaluación. Conocimientos y habilidades en Ciencias, Matemáticas y Lectura*. Programa para la evaluación internacional de alumnos, OCDE.

Estos ítems fueron también revisados y sancionados por el Ministerio. Además, se crearon ítems para medir aquellos objetivos que se incorporaron a este nivel y que, según el currículo anterior, o no se estudiaban o se estudiaban en otro nivel. Finalmente, el instrumento constó de 30 ítems en cada uno de los ejes temáticos de las bases curriculares: Números y Operaciones, Geometría, Patrones y Álgebra, Medida y Datos y Probabilidad.

Posteriormente se elaboró la maqueta de la prueba para su posterior diseño y pilotaje, que se efectuó en 3 escuelas, una que utilizaba el método Singapur (40 estudiantes) y dos que no lo utilizaban (106 estudiantes). En la prueba piloteada se incluyeron los 30 ítems que era necesario validar. Una vez que se contó con las respuestas digitalizadas, se realizaron los análisis psicométricos que permitieran obtener información respecto de la prueba y de cada ítem. Se calculó el Alpha de Cronbach (0,7), logrando una prueba de una confiabilidad razonable. La prueba diseñada y finalmente aplicada por CFK se adjunta en el Anexo N° 5 de este informe.

- Aplicación de la prueba: aspectos operativos

De los establecimientos seleccionados, el equipo central del CFK solicitó a la contraparte el cambio de 10 escuelas por problemas de accesibilidad. El Ministerio procedió a reemplazarlos identificando establecimientos con similares características que fueran geográficamente accesibles. También se determinaron otros establecimientos para reemplazar a aquellos que no pudieron participar de la evaluación, ya sea por tener comprometidas evaluaciones internas o de instituciones externas en las fechas propuestas o porque no se logró establecer contacto telefónico o por correo electrónico de manera oportuna.

Para aplicar la prueba, se ideó un plan de acción diferenciado para regiones y para la Región Metropolitana. En el caso de regiones, las escuelas se distribuyeron por zonas geográficas tomando como centros zonales las capitales provinciales de las regiones pertenecientes al estudio y las escuelas pertenecientes a la zona de la capital provincial fueron aquellas escuelas para las que el tiempo de viaje no era de más de una hora y media desde dicha capital. En el caso de la Región Metropolitana, se realizó una distribución distinta. Se comenzó a aplicar de las comunas más lejanas hasta las más cercanas a Santiago. La cantidad de examinadores por región se calculó a partir de esta distribución, tomando en cuenta la cantidad de escuelas y cursos por cada zona.

Operativamente se definieron los siguientes actores: Supervisores Zonales, Supervisores de Escuela y Examinadores. Ellos forman parte de los equipos de trabajo con que el CFK ha abordado anteriormente proyectos de este tipo. Además, se hizo uso de contactos existentes en universidades regionales para completar los equipos regionales. Para informar y capacitar a los distintos actores involucrados sobre los roles y acciones que cada uno debía llevar a cabo durante el proceso de aplicación, se confeccionó un "Manual de Aplicación" (Centro Félix Klein, 2012), el cual se basó en los distintos protocolos de evaluación utilizados por el CFK, complementados con los manuales de aplicación de la prueba SIMCE.

En la capacitación, se estudiaron las acciones presentes en aquel manual, se respondieron las consultas y se dieron indicaciones para solucionar los distintos problemas que pudieran surgir al aplicar la prueba. Además, se confeccionó un

instructivo para aplicar la prueba a niños con necesidades educativas especiales. Estas capacitaciones se realizaron por regiones y a cada una de ellas asistió un integrante del CFK.

Una vez seleccionados los Supervisores Zonales, se comenzó la visita a los establecimientos antes de la aplicación de la prueba, con la finalidad de informar y coordinar, con directivos y docentes de cada establecimiento, el proceso de aplicación de la prueba, además de verificar en terreno las condiciones de las aulas.

La prueba fue finalmente aplicada entre los días 27 y 29 de Noviembre de 2012 en todos los establecimientos seleccionados, contando con un total final de 8.932 alumnos que rindieron la prueba, alcanzando al 94% del total de la matrícula de los 143 establecimientos evaluados.

Los estudiantes debían marcar sus respuestas en la prueba misma y para el proceso de análisis de resultados las respuestas fueron vaciadas a una hoja especialmente diseñada para ser leída por un lector óptico y almacenadas en un archivo digital. CFK contrató a personas capacitadas para realizar esta tarea. Se asignó a cada digitalizador un curso completo; después de realizado el vaciado a las hojas de respuestas, se revisaban tres pruebas al azar para confirmar que este vaciado se hubiera realizado de forma correcta. Luego, las hojas de respuestas por curso eran leídas por el lector óptico para generar el archivo digital con las respuestas. En este proceso, la persona encargada de escanear estas hojas revisaba que los datos estuvieran correctamente traspasados por el lector.

A partir del archivo digital de las respuestas, se confeccionó una base de datos que contenía el RBD del establecimiento, el Curso del alumno, el Número de Lista del alumno, el Sexo y las Respuestas a cada una de las Preguntas.

Para poder complementar los análisis estadísticos de la prueba aplicada, el CFK elaboró, diseñó y aplicó dos encuestas: una para docentes de las escuelas que utilizaban el set de libros del programa y otra para docentes de las escuelas que no lo hicieron (detalles de la encuesta en Centro Félix Klein, 2012).

En el Anexo N° 6 hemos resumido los principales hallazgos y conclusiones obtenidos por CFK en el proceso de aplicación de la prueba, que permiten contar con información estadística más detallada de los resultados de los grupos de tratamiento y control, analizando la información por eje temático, habilidades, niveles de logro, dependencia administrativa y nivel socio económico de los establecimientos, además de analizar los resultados tomando en cuenta los resultados de las encuestas aplicadas a los docentes.

– **Evaluación de impacto**

- Análisis del cumplimiento de condiciones para la evaluación de impacto

Para efectuar la medición del impacto inmediato o de corto plazo del programa en los aprendizajes de los niños y niñas de 2° básico, utilizamos las bases de datos de los resultados entregadas al Ministerio por el Centro Félix Klein, generados a través de la prueba elaborada y aplicada por dicha institución.

Antes de efectuar la medición misma del impacto es importante determinar la calidad de la aleatorización, la que puede ser corroborada al aplicar una regresión que relaciona la condición de tratado y control, T_i , con las variables de selección, para los años 2007 a 2009, mediante una regresión del tipo:

$$T_i = \beta_0 + \beta_1 X_i + u_i \quad (1)$$

donde $T_i = 1$, si el establecimiento i participa en el programa y $T_i = 0$ de lo contrario y donde X_i corresponde a las características observables del establecimiento que explican la variable de resultado y preceden al tratamiento y donde u_i es un error con media 0 y varianza constante. Si el tratamiento está asignado de manera aleatoria, no debería estar correlacionado con las características individuales, es decir, cuando X_i no está correlacionada con el tratamiento T_i los coeficientes que acompañan a X_i deben ser cero. La siguiente tabla muestra el resultado de esta regresión

Tabla 2: Verificación de la calidad de la aleatorización

Variables	Coefficiente	Error estándar	T	p-value
SIMCE Mat. 07	,0046	,0026	1,77	0,078
SIMCE Mat. 08	,0038	,0028	1,35	0,180
SIMCE Mat. 09	-,0021	,0028	-0,75	0,453
Dependencia ¹	-,1026	,1024	-1,00	0,318
Zona ²	-,0800	,1228	-0,65	0,516
Región ³	,0552	,0827	0,67	0,506
Constante	-,9700	,7241	-1,34	0,183

1: Dependencia toma el valor 1 si el establecimiento es particular subvencionado y 0 si es municipal.

2: Zona toma el valor 1 si el establecimiento es urbano y 0 si es rural.

3: Región toma el valor 1 si el establecimiento es de la Región Metropolitana y 0 si es de cualquier otra región.

Los *p-value* indican que todas las variables relevantes disponibles para explicar la participación en el tratamiento no son significativas (excepto en el caso de la prueba SIMCE de 2007, que lo es al 10%), lo que permite concluir que la aleatorización fue, en general, bien realizada o que los grupos son comparables.

La selección aleatoria se efectúa para evitar el sesgo de selección¹⁴ y, por ende, asegurar que éste sea 0, o lo que es lo mismo, que $E[Y_i(0)|T_i = 1] = E[Y_i(0)|T_i = 0]$, donde $Y_i(T_i)$ se define como la variable de resultado para cada individuo i , dado su estado T_i . La condición anterior es equivalente a asegurar que $E[u_i|T_i] = 0$ en la ecuación (2), donde u_i es el término de error que contiene el efecto de otras variables que pudieran afectar la variable de resultado. Es decir, cuando la aleatorización es exitosa no hay sesgo de selección, por lo cual el efecto del tratamiento se puede estimar como la diferencia simple entre el promedio de la variable de resultado de tratados y controles $[\bar{Y}|T = 1] - [\bar{Y}|T = 0]$ o, lo que es equivalente, mediante una regresión por Mínimos Cuadrados Ordinarios (MCO) del tipo

$$Y_i = \beta_0 + \beta_1 T_i + u_i \quad (2)$$

¹⁴ El sesgo de selección ocurre cuando los individuos se autoseleccionan en un programa, debido a la existencia de características observadas y no observadas en ellos que condicionan su decisión de participar, por lo cual los grupos de tratamiento y control podrían ser sistemáticamente diferentes entre sí.

donde β_1 es llamado el *estimador de diferencias*, que es insesgado y consistente aunque no necesariamente eficiente, porque otras variables, distintas al tratamiento, pueden también determinar la variable de resultado aun cuando la aleatorización sea exitosa. Por ello, la eficiencia del estimador se puede mejorar dependiendo de la disponibilidad de datos, estimando una regresión del tipo

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 X_i + u_i \quad (3)$$

donde β_1 es llamado el *estimador de diferencias con regresores adicionales* y X_i corresponde a las características del individuo que explican también la variable de resultado y preceden al tratamiento.

En lo que sigue efectuamos el análisis de impacto a nivel de estudiantes y a nivel de establecimientos. El análisis a nivel de estudiantes lo realizamos utilizando MCO con datos agrupados en *clusters* y modelos multinivel, mientras que el análisis a nivel de establecimientos lo efectuamos mediante MCO. El análisis a nivel de estudiantes se realiza de ese modo debido a que, por la naturaleza de los datos, los resultados a nivel de alumnos normalmente están correlacionados al interior de los establecimientos (*clusters*). Por ello, una correcta estimación debe dar cuenta de esta correlación, pues de lo contrario se subestimarían los errores estándares, generando inferencia inválida. Esto ocurre porque una regresión en MCO convencional a nivel de estudiantes asume que sus resultados son independientes entre sí, lo que genera errores estándares usualmente subestimados. Este tipo de correlación entre las observaciones organizadas en *clusters*, se llama correlación intraclase¹⁵. Así, antes de efectuar la estimación del impacto del programa entre los alumnos, verificamos la correlación intraclase, utilizando el siguiente modelo de medias incondicionales expresado en dos niveles, donde en el primer nivel, ecuación (4), se expresa la variable de resultado de los alumnos como la suma de un *intercepto* para el estudiante del establecimiento, β_{0j} , y de un término de error aleatorio, $\varepsilon_{ij} \sim N(0, \sigma^2)$, asociado con el i -ésimo estudiante en el j -ésimo establecimiento. En el segundo nivel, ecuación (5), se expresan los interceptos a nivel de establecimiento como la suma de una media general, γ_{00} , y una serie de desviaciones aleatorias en relación a esa media, con $v_{0j} \sim N(0, \tau_{00})$. El modelo es el siguiente

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad (4)$$

$$\beta_{0j} = \gamma_{00} + v_{0j} \quad (5)$$

Al sustituir (5) en (4) se genera el modelo multinivel

$$Y_{ij} = \gamma_{00} + v_{0j} + \varepsilon_{ij} \quad (6)$$

donde $v_{0j} \sim N(0, \tau_{00})$ y $\varepsilon_{ij} \sim N(0, \sigma^2)$. Este primer modelo consta de dos partes, una parte fija que contiene un intercepto general, γ_{00} , y una parte aleatoria que contiene dos efectos aleatorios, uno para el intercepto, v_{0j} , y otro para el residuo dentro de los establecimientos, ε_{ij} . Se ajusta este modelo para estimar tanto el efecto fijo γ_{00} , que

¹⁵ Ver la nota 7 anterior.

nos habla del promedio de preguntas correctas en la población, y de los dos efectos aleatorios: τ_{00} , que se refiere a la variabilidad en las medias de los establecimientos y σ^2 , que indica la variabilidad de los resultados de la prueba al interior de los establecimientos. Los resultados de la estimación de (6) son los siguientes

Tabla 3: Resultados del modelo de medias incondicionales

VARIABLES	ppc
Constante	59,12*** (0,682)
Observaciones	8.839
Número de grupos	143
Errores estándares en paréntesis *** p<0.01, ** p<0.05, * p<0.1	

Los resultados encontrados en la estimación de los componentes de la varianza del modelo de medias incondicionales a partir de (6) son $\tau_{00} = 61,04$ y $\sigma^2 = 243,40$, los que sugieren que los establecimientos difieren en el promedio de respuestas correctas en la prueba y que hay aún más variación entre los estudiantes, pues la varianza dentro de los establecimientos es casi 4 mayor que el tamaño de la varianza entre establecimientos (ver Tabla 4).

Tabla 4: Estimación de los parámetros de covarianza del modelo de medias incondicionales

Efectos aleatorios	Estimación	Errores estándares
Var. de la constante (τ_{00})	61,04	7,912
Var. del residuo (σ^2)	243,40	3,691

Otra forma de ver esto es mediante la correlación intraclase ya mencionada ($\hat{\rho} = \frac{\hat{\tau}_{00}}{\hat{\tau}_{00} + \hat{\sigma}^2}$), que nos indica qué porción de la varianza total corresponde a la que se observa entre los establecimientos. En nuestro caso $\hat{\rho} = 0,20$ lo que sugiere que considerar los datos en *clusters* es correcto y que utilizar un modelo MCO produciría resultados incorrectos¹⁶.

- Estimación del impacto a nivel de alumnos
 - Estimación por MCO con datos agrupados (*clusters*)

La utilización correcta de MCO asume el cumplimiento de los siguientes supuestos:

- i) $E(u_i|X_i) = 0$ (exogeneidad de los regresores),
- ii) $E(u_i^2|X_i) = \sigma^2$ (homocedasticidad condicional) y
- iii) $E(u_i u_j | X_i, X_j) = 0$ con $i \neq j$ (observaciones condicionalmente no correlacionadas).

¹⁶ Para una interpretación de ρ véase Singer (1998), p. 329. La correlación intraclase es diferente de la correlación de Pearson que se observa entre dos variables; Kreft & Leeuw (2002, p. 10) muestran que una correlación intraclase de 0,2 (que como correlación de Pearson puede ser considerada moderada o baja) es en realidad muy alta, pues en vez de corresponderse con un nivel de confianza de 0,05 (que es lo más usual) lo hace con un nivel de confianza de 0,28 cuando se usan diez observaciones y llega a 0,7 al utilizar 100 observaciones.

El supuesto i) es esencial para la estimación consistente de los parámetros β de interés y los supuestos ii) y iii) determinan la forma de la matriz de varianzas-covarianzas de $\hat{\beta}$. Cuando ii) no se cumple tenemos errores heterocedásticos y para una adecuada estimación se requiere un estimador robusto a heterocedasticidad. Cuando no se cumple el supuesto iii), los errores para diferentes observaciones están correlacionados de modo que las estimaciones robustas antes indicadas tampoco son válidas, pues cuando los datos están organizados en *clusters* no siempre es posible tratarlos de manera independiente.

De este modo es posible estimar, mediante MCO, regresiones a nivel de estudiantes agrupados en *clusters* (establecimientos). Para ello estimamos un modelo donde la variable dependiente es el *porcentaje de respuestas correctas* a nivel de alumnos contra la condición de tratamiento, T , que toma el valor 1 si el alumno pertenece a un establecimiento que participó en el programa Textos Singapur y 0 en otro caso. Se utilizan, además, un conjunto de variables de control para mejorar la eficiencia de la estimación¹⁷. Las estimaciones se efectúan la muestra completa de alumnos y para varias sub muestras de alumnos; entre las sub muestras se consideran la dependencia del establecimiento (municipal y particular subvencionada), la región de los establecimientos de los alumnos (si son de la región metropolitana o de las demás regiones), la zona de los establecimientos de los alumnos (urbana o rural), el sexo de los alumnos (hombres y mujeres) y dos sub muestras relacionadas con los niveles de logro (intermedio o avanzado) de los resultados promedios en las pruebas SIMCE de 2007 a 2009, considerando, en el primer caso, tanto el promedio de las pruebas de matemática y lenguaje de 4° básico y, en el segundo, sólo el promedio de las pruebas de matemáticas de dicho nivel. Los resultados se muestran en la Tabla 5

¹⁷ En todas las estimaciones se utilizan, como variables de control para mejorar la eficiencia de la estimación, el nivel socio económico del establecimiento, mediante variables *dummy* para cada nivel socio económico (usado *GSE C* como grupo de comparación), la variable *zona*, que toma el valor 1 si el establecimiento es urbano y 0 si es rural, la variable *Género*, que toma el valor 1 si el alumno es mujer y 0 si es hombre, *Dependencia*, que toma el valor 1 si el establecimiento es particular subvencionado y 0 si es municipal; se usan las variables *RBD femenino* y *RBD masculino*, que indican si el establecimiento es exclusivamente de mujeres u hombres respectivamente (usando como grupo de comparación a los establecimientos mixtos) y *RM*, que toma el valor 1 si el establecimiento es de la Región Metropolitana y 0 si es de otras regiones.

Tabla 5: Estimaciones para la muestra completa y para distintas sub muestras

Dimensión	Muestra	Impacto
Muestra completa		2,029* (1,09)
Dependencia	Municipal	1,679 (1,959)
	Part. Subvencionado	2,468* (1,305)
Región	Reg. Metropolitana	1,959 (1,821)
	Otras regiones	0,538 (1,461)
Zona	Rural	0,871 (3,204)
	Urbano	1,778 (1,148)
Sexo del alumno	Hombre	1,0628 (1,114)
	Mujer	3,000** (1,327)
SIMCE promedio lenguaje y matemáticas 2007-2009	Logro Intermedio	1,781 (1,341)
	Logro Avanzado	0,421 (1,836)
SIMCE promedio matemáticas 2007-2009	Logro Intermedio	1,941 (1,335)
	Logro Avanzado	0,632 (1,918)

Errores estándares en paréntesis
*** p<0.01, ** p<0.05, * p<0.1

Considerando la muestra completa de alumnos se constata que el impacto es positivo para los niños acogidos al programa y alcanza al 2,02%, siendo estadísticamente significativo, pero de magnitud más bien modesta.

Al mirar los efectos del programa sobre las sub muestras de alumnos, encontramos que el impacto es estadísticamente significativo entre los alumnos de establecimientos particulares subvencionados, pero no para el caso de alumnos de establecimientos municipales. En el primer caso, el impacto alcanza a 2,46% a favor de los alumnos acogidos al programa, mientras que en el segundo caso no hay diferencia estadísticamente significativa. Tampoco se encontraron diferencias estadísticamente significativas cuando los alumnos se agrupan por la región de sus establecimientos o por la zona geográfica donde éstos se encuentran. Sí se observa un impacto del programa que favorece a las niñas que participan en él, alcanzando un 3%, pero no en el caso de los niños. Tampoco se encontraron diferencias estadísticamente significativas al agrupar a los alumnos según el logro de los niños de los establecimientos en las pruebas SIMCE de los años anteriores.

- Análisis Multinivel

Nuestra segunda estrategia de evaluación a nivel de estudiantes consiste en la aplicación de un modelo multinivel. Para determinar el impacto puro del programa en los resultados de los alumnos, se debería ajustar un modelo en dos niveles con efectos a nivel de escuela (nivel 2), incluyendo como control la variable de tratamiento (participación o no del establecimiento en el Programa Singapur). De ese modo tendríamos el siguiente modelo

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad (7)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}T + v_{0j} \quad (8)$$

donde suponemos que $v_{0j} \sim N(0, \tau_{00})$ y $\varepsilon_{ij} \sim N(0, \sigma^2)$. Al sustituir (8) en (7) se genera el modelo multinivel

$$Y_{ij} = [\gamma_{00} + \gamma_{01}T] + [v_{0j} + \varepsilon_{ij}] \quad (9)$$

Para mejorar la eficiencia de la estimación del efecto del tratamiento en el rendimiento de los alumnos en la aplicación de la prueba, se agregan variables de control, tanto a nivel de estudiantes como de establecimientos. Nos encontramos con dos opciones: por un lado, podemos especificar un modelo multinivel con interceptos aleatorios y, por otro, un modelo multinivel con interceptos y pendientes aleatorias. La decisión de cuál modelo ajusta mejor se puede resolver usando los criterios de selección de modelos usuales (AIC y BIC). En la primera opción, sólo con interceptos aleatorios, tenemos un modelo del tipo

$$Y_{ij} = \beta_{0j} + \beta_{1j}genero + \varepsilon_{ij} \quad (10)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}T + \gamma_{02k}GSE_k + \gamma_{03}dependencia + \gamma_{04}RM + v_{0j} \quad (11)$$

donde las variables tienen el mismo significado que el utilizado en la sección anterior y donde GSE_k representa el nivel socio económico del establecimiento, con $k = \{A, B, C, D, E\}$. Reemplazando (11) en (10) y ordenando generamos el siguiente modelo con interceptos aleatorios

$$Y_{ij} = [\gamma_{00} + \gamma_{01}T + \gamma_{02k}GSE_k + \gamma_{03}dependencia + \gamma_{04}RM + \beta_{1j}genero] + [\varepsilon_{ij} + v_{0j}] \quad (12)$$

donde los términos contenidos en el primer paréntesis cuadrado representan la parte fija del modelo y los términos en el segundo paréntesis cuadrado la parte aleatoria.

En el segundo caso, con interceptos y pendientes aleatorios, tenemos un modelo del tipo

$$Y_{ij} = \beta_{0j} + \beta_{1j}genero + \varepsilon_{ij} \quad (13)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}T + \gamma_{02k}GSE_k + \gamma_{03}dependencia + \gamma_{04}RM + v_{0j} \quad (14)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}T + \gamma_{12k}GSE_k + \gamma_{13}dependencia + \gamma_{14}RM + v_{1j} \quad (15)$$

Reemplazando (14) y (15) en (13), simplificando la notación y ordenando, generamos el siguiente modelo

$$Y_{ij} = [\gamma_{00} + \gamma_{01}T + \gamma_{02k}GSE_k + \gamma_{03}dep + \gamma_{04}RM + \gamma_{10}gen + \gamma_{11}T \cdot gen + \gamma_{12k}GSE_k \cdot gen + \gamma_{13}dep \cdot gen + \gamma_{14}RM \cdot gen] + [v_{0j} + v_{1j} \cdot gen + \varepsilon_{ij}] \quad (16)$$

Al efectuar las estimaciones correspondientes y comparar los criterios de información, encontramos que AIC y BIC son levemente menores en el primer caso, por lo cual el modelo seleccionado es sólo aquel con interceptos aleatorios.

Como antes, se presentan los resultados para la muestra completa de alumnos y para varias sub muestras de alumnos, considerando también la dependencia del establecimiento, la región de los establecimientos de los alumnos, la zona de los establecimientos de los alumnos, el sexo de los alumnos y dos sub muestras relacionadas con los niveles de logro (intermedio o avanzado) de los resultados

promedios en las pruebas SIMCE de 2007 a 2009. Los resultados se muestran en la Tabla 6

Tabla 6: Estimaciones para la muestra completa y para distintas sub muestras (multinivel)

Dimensión	Muestra	Impacto
Muestra completa		2,845** (1,14)
Dependencia	Municipal	2,187 (1,931)
	Part. Subvencionado	3,616** (1,416)
Región	Reg. Metropolitana	3,270* (1,805)
	Otras regiones	2,604 (1,601)
Zona	Rural	3,585 (4,426)
	Urbano	2,804** (1,235)
Sexo del alumno	Hombre	2,537** (1,242)
	Mujer	2,739** (1,312)
SIMCE promedio lenguaje y matemáticas 2007-2009	Logro Intermedio	2,091 (1,315)
	Logro Avanzado	2,310 (2,262)
SIMCE promedio matemáticas 2007-2009	Logro Intermedio	2,262* (1,310)
	Logro Avanzado	2,372 (2,295)
Errores estándares en paréntesis *** p<0.01, ** p<0.05, * p<0.1		

En este caso, tomando la muestra completa de alumnos se observa que el impacto del programa es positivo y alcanza al 2,84% de rendimiento de los niños acogidos al programa en relación a los demás; es estadísticamente significativo al 5% de confianza y también la magnitud es más bien modesta.

Al mirar los efectos sobre las sub muestras, encontramos que el impacto también es estadísticamente significativo entre los alumnos de establecimientos particulares subvencionados (3,61%), pero no para el caso de alumnos de establecimientos municipales. Sí se constata, en este caso, un impacto positivo y estadísticamente significativo al 10%, entre los alumnos del programa de establecimientos de la región metropolitana, pero no entre alumnos de las demás regiones. También se observa un impacto positivo y estadísticamente significativo del programa al comparar los establecimientos urbanos, favoreciendo a aquellos que participaron en el programa, con un 2,84% de mayor rendimiento en relación a los establecimientos del grupo de control. Usando el modelo multinivel sí observamos un impacto del programa que favorece el rendimiento tanto de las niñas como de los niños que participan en él, siendo levemente mayor en el caso de las niñas (2,73% contra un 2,53% de los niños); este resultado resultó estadísticamente significativo al 5%. Finalmente, al agrupar a los niños por niveles de logro en las pruebas SIMCE de 2007 a 2009, se encontró un impacto estadísticamente significativo del programa en favor de los niños de establecimientos que obtuvieron un nivel de logro intermedio en el SIMCE de matemáticas entre esos años (de 2,26% a favor de los tratados), pero no en los demás casos.

- o Estimación del impacto a nivel de establecimientos

En la estimación a nivel de establecimientos la variable dependiente corresponde al porcentaje promedio de respuestas correctas a nivel de establecimientos y se utilizan, como variables de control, la variable de tratamiento, T , que toma el valor 1 si el establecimiento participa del programa Singapur y 0 de lo contrario. Como antes, también se utilizan otras variables de control¹⁸, además de la variable de tratamiento, con la finalidad de mejorar la eficiencia de la estimación del efecto puro del tratamiento.

Los resultados también se presentan para la muestra completa de alumnos y para las mismas sub muestras anteriores, excepto para la dimensión sexo del alumno, y son los siguientes

Tabla 7: Estimaciones del impacto del programa a nivel de establecimientos

Dimensión	Muestra	Impacto
Muestra completa		2,995** (1,206)
Dependencia	Municipal	2,441 (2,028)
	Part. Subvencionado	3,681** (1,509)
Región	Reg. Metropolitana	3,468* (1,937)
	Otras regiones	2,339 (1,618)
Zona	Rural	3,408 (5,038)
	Urbano	2,987** (1,295)
SIMCE promedio lenguaje y matemáticas 2007-2009	Logro Intermedio	2,184 (1,384)
	Logro Avanzado	2,619 (2,500)
SIMCE promedio matemáticas 2007-2009	Logro Intermedio	2,352* (1,374)
	Logro Avanzado	2,699 (2,523)

Errores estándares en paréntesis
 *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Considerando la muestra completa de establecimientos se observa que el impacto del programa es positivo, estadísticamente significativo al 5% de confianza, y alcanza a un rendimiento 2,99% superior por parte de los establecimientos acogidos al programa en relación a los no participantes.

Al analizar el efecto del programa sobre las distintas sub muestras, encontramos que el impacto es estadísticamente significativo para los establecimientos particulares subvencionados (3,61%), pero no para el caso de alumnos de establecimientos municipales, y también se observa un impacto positivo, aunque modesto y estadísticamente significativo al 10%, entre los establecimientos del programa de la región metropolitana, pero no de las demás regiones. Sí se constata, en este caso, un

¹⁸ Se utilizan como variables de control la dependencia, el nivel socio económico, la región y la zona del establecimiento.

impacto positivo y estadísticamente significativo al 10%, entre los alumnos del programa de establecimientos de la región metropolitana, pero no entre alumnos de las demás regiones. También, en este caso, se observa un impacto positivo y estadísticamente significativo entre los establecimientos urbanos, que favorece a los del programa en 2,98% de mayor rendimiento que los rurales. Finalmente, al analizar el efecto del programa agrupando a los establecimientos por niveles de logro en las pruebas SIMCE de 2007 a 2009, se encontró un impacto estadísticamente significativo del programa en favor de los establecimientos que obtuvieron un nivel de logro intermedio en el promedio de la prueba SIMCE de matemáticas entre esos años (de 2,35% a favor de los tratados), pero no en los demás casos.

- Estimaciones por grado de dificultad y eje temático de las preguntas

En las estimaciones anteriores la variable de resultado correspondía al porcentaje de preguntas correctamente contestadas sobre el total de preguntas. En lo que sigue desagregamos esta variable de dos maneras distintas: por un lado, generamos variables que miden el resultado de agrupar las preguntas según los grados de dificultad a los que correspondía cada una (preguntas de nivel inicial, intermedio y avanzado) y, por otro, según los ejes temáticos a los cuales pertenecía cada una de ellas (números, operatoria, patrones y álgebra, medición, geometría y datos y probabilidad). Nos interesa conocer, en este caso, los efectos del programa en el porcentaje de respuestas correctas en cada una de estas dimensiones por separado. Además, como en las estimaciones anteriores encontramos que, sistemáticamente, los resultados fueron estadísticamente significativos cuando se consideraron las submuestras de género y dependencia administrativa de los establecimientos, utilizaremos este mismo desglose con estas nuevas variables de resultado.

La Tabla 8 muestra el impacto del programa según el grado de dificultad de las preguntas, diferenciando el género de los alumnos. Se agrupan los datos en *cluster* y se usan modelos multinivel. Sistemáticamente, el mayor efecto del programa en favor de los niños participantes en el programa se observa en las preguntas de complejidad intermedia, mientras que en las preguntas cuyo grado de dificultad era inicial o avanzado no hubo diferencias estadísticamente significativas entre ambos grupos bajo ambas metodologías. Por otro lado, al descomponer la variable de resultados de esta manera, no es posible concluir de manera evidente si los niños tuvieron o no un mejor desempeño que las niñas, pues ambas metodologías arrojan resultados en sentido contrario, de modo que el mayor logro de la niñas observado en las Tablas 5 y 6, podría deberse a que, en dicho caso, la variable de resultado comprendía todas las preguntas sin distinguir por grado de dificultad.

Tabla 8: Impacto del programa Singapur por nivel de dificultad y género

Grado de dificultad	Estudiantes-Clúster		Estudiantes-Multinivel	
	Hombres	Mujeres	Hombres	Mujeres
Inicial	0,238 (0,837)	1,209 (1,153)	0,739 (0,887)	1,287 (1,058)
Intermedio	2,608** (1,167)	4,052*** (1,532)	3,727*** (1,303)	3,640** (1,481)
Avanzado	0,904 (1,439)	1,691 (1,437)	1,996 (1,493)	1,937 (1,551)

Errores estándares en paréntesis
 *** p<0.01, ** p<0.05, * p<0.1

Cuando se estima el impacto del programa a nivel de establecimientos, aplicando MCO y considerando su dependencia administrativa (Tabla 9), se encuentra que los

establecimientos particulares subvencionados acogidos al programa presentan un mejor desempeño en las preguntas cuyo grado de dificultad es intermedio o avanzado, en relación a los particulares subvencionados que no participaron en él y no se detectan diferencias estadísticamente significativas entre los establecimientos municipales participantes o no en el programa. Además, en ambos casos la magnitud del efecto es mayor entre los establecimientos particulares subvencionados en relación a los municipales.

Tabla 9: Impacto del programa Singapur por nivel de dificultad y dependencia (establecimientos)

Grado de dificultad	Establecimientos-MCO	
	Municipales	Part. Subvencionados
Inicial	2,308 (1,652)	1,447 (1,090)
Intermedio	3,230 (2,272)	5,431*** (1,702)
Avanzado	0,985 (2,284)	3,488* (2,060)

Errores estándares en paréntesis
 *** p<0.01, ** p<0.05, * p<0.1

La Tabla 10 cuantifica el impacto del programa, a nivel de estudiantes, en el porcentaje de respuestas correctas agrupadas en cada uno de los ejes temáticos a los cuales corresponden. También para realizar las estimaciones se agrupan los datos en *cluster* y se usan modelos multinivel. En este caso, se encuentra un efecto sistemáticamente favorable para los estudiantes acogidos al programa en las preguntas relacionadas con los ejes de operatoria, patrones y álgebra y datos y probabilidad y en dos de ellos, operatoria y datos y probabilidad, el efecto es mayor entre las mujeres en relación a los hombres, mientras que en el eje de patrones y álgebra el efecto es sistemáticamente mayor entre los hombres. En los ejes de medición y geometría no se encuentran diferencias estadísticamente significativas entre niños y niñas participantes y no participantes en el programa, mientras que en el eje números el efecto del programa, medido mediante ambas metodologías, muestra que los niños que sí participan obtienen un resultado menor que los demás niños, no siendo estadísticamente significativa esta diferencia entre las niñas. Por otro lado, si bien las diferencias no son estadísticamente significativas entre ambos grupos, se observa que en el eje Geometría los resultados parecen desfavorecer a los alumnos del programa, mientras que en el eje medición ocurriría lo contrario.

Tabla 10: Impacto del programa Singapur por eje temático y género de alumnos (estudiantes)

Eje temático	Estudiantes-Clúster		Estudiantes-Multinivel	
	Hombres	Mujeres	Hombres	Mujeres
Números	-2,823*** (1,052)	-1,345 (1,360)	-2,514** (1,087)	-1,668 (1,247)
Operatoria	2,465** (1,145)	5,424*** (1,543)	3,872*** (1,289)	4,803*** (1,449)
Patrones y Álgebra	4,230*** (1,422)	3,256** (1,415)	4,644*** (1,382)	3,607** (1,402)
Medición	0,443 (1,580)	-0,011 (1,929)	1,942 (1,752)	1,331 (1,977)
Geometría	-2,474 (1,630)	-1,783 (1,882)	-1,799 (0,602)	-2,273 (1,832)
Datos y Probabilidad	5,292** (2,171)	8,820*** (2,421)	7,838*** (2,213)	8,237*** (2,456)

Errores estándares en paréntesis
 *** p<0.01, ** p<0.05, * p<0.1

De todos modos, se debe ser cauto al interpretar estos resultados, porque podrían estar fuertemente condicionados por el orden en que fueron desarrollados los ejes en el programa Textos Singapur en relación al orden seguido en el desarrollo de los ejes

en los demás establecimientos. Es decir, es posible que la aparente ventaja de un grupo sobre el otro en un determinado eje no se deba necesariamente a las características de cada programa en sí, sino a la oportunidad en que fueron desarrollados los ejes, de modo que el tiempo dedicado al aprendizaje de ciertos contenidos o el tiempo disponible para su apropiación por parte de los alumnos generara logros distintos.

Finalmente, la Tabla 11 muestra los efectos del programa por eje temático y dependencia de los establecimientos utilizando MCO. Se encuentra, también aquí, que los mayores efectos positivos para quienes participaron del programa se observan entre los establecimientos particulares subvencionados en los ejes datos y probabilidad, operatoria y patrones y álgebra, mientras que en el eje geometría el mayor efecto del programa se observa entre los establecimientos particulares subvencionados que no participaron en él, lo mismo que el eje números, aun cuando en este caso el impacto no resultó estadísticamente significativo.

Tabla 11: Impacto del programa Singapur por eje temático y dependencia (establecimientos)

Eje temático	Establecimientos-MCO	
	Municipal	Part. Subvencionados
Números	-1,305 (1,696)	-1,127 (1,375)
Operatoria	3,494 (2,097)	6,155*** (1,675)
Patrones y Álgebra	4,375* (2,321)	4,469** (1,690)
Medición	2,309 (2,838)	2,649 (2,426)
Geometría	0,993 (2,698)	-3,971* (2,023)
Datos y Probabilidad	5,223 (3,376)	10,444*** (2,802)
Errores estándares en paréntesis *** p<0.01, ** p<0.05, * p<0.1		

Al efectuar estas nuevas estimaciones del efecto del programa en el porcentaje de respuestas correctas, agrupando las preguntas según grado de dificultad y eje temático al cual pertenecen, es posible encontrar resultados que permiten matizar los hallazgos de las secciones anteriores. Por un lado, no es tan evidente que las niñas obtengan un resultado sistemáticamente superior al de los niños cuando se desagregan las variables de resultado por grado de dificultad de las preguntas o por el eje temático al cual pertenecen. Si bien las magnitudes de los efectos son claramente mayores en valor absoluto, en unos casos parecen favorecer a los niños y en otros a las niñas. Sí es evidente que, agrupando las preguntas por grado de dificultad o eje temático, en los establecimientos particulares subvencionados el efecto del programa en el porcentaje de respuestas correctas parece ser sistemáticamente mayor en ellos en relación al efecto observado en los establecimientos municipales.

– Conclusiones

Se aplicaron tres estrategias de medición del efecto del programa en el aprendizaje de las matemáticas por parte de los estudiantes evaluados. Dos a nivel de estudiantes (mediante datos agrupados en *clusters* y mediante un análisis multinivel) y una a nivel de establecimientos. En todas las estimaciones se consideró, por un lado, la muestra completa de alumnos y/o establecimientos y, por otro, distintas sub muestras para determinar los efectos diferenciados del programa. Se efectuaron, además, evaluaciones utilizando como variables de resultados el porcentaje de respuestas

correctas agrupando las preguntas según grado de dificultad y eje temático al cual pertenecen.

Considerando la muestra completa, en los tres casos se pudo determinar que el efecto del programa en el porcentaje de logro es entre 2,02% y 2,99% mayor entre los alumnos y establecimientos que participaron en el Programa en relación a los que no lo hicieron, siendo la diferencia estadísticamente significativa.

Al considerar las distintas muestras de alumnos y establecimientos, se puede concluir que el programa ha tenido un mayor efecto entre los alumnos de establecimientos particulares subvencionados. En efecto, bajo las tres estrategias de medición se observan porcentajes de logro entre 2,46% y 3,68% mayores en el caso de los establecimientos particulares subvencionados participantes en el programa, mientras que en el caso de los establecimientos municipales no se observa un efecto estadísticamente distinto de cero. Este resultado es consistente con lo encontrado por el Centro Félix Klein (2012) y consistente también con el desempeño de los establecimientos en las pruebas SIMCE, en general, al considerar la dependencia administrativa de los establecimientos.

Del mismo modo, se pudo observar un efecto mayor en los establecimientos participantes en el programa que son de la región metropolitana, en relación a los establecimientos de las demás regiones, de entre 3,27% y 3,4%, mientras que en el rendimiento de alumnos y establecimientos de otras regiones no se observan diferencias de rendimiento estadísticamente significativas entre quienes participan en el programa y quienes no lo hacen.

El efecto del programa es también mayor entre los establecimientos urbanos en relación a los rurales, con un rendimiento de entre 2,80% y 2,98% mayor entre los establecimientos del programa; en las zonas rurales no se encontraron efectos estadísticamente distintos de cero entre tratados y controles.

Por sexo de los alumnos se encontró que el efecto del programa entre las mujeres sujetas al programa ha sido entre 2,73% y 3% mayor que entre las mujeres de los demás establecimientos, mientras que entre los hombres se encontró un efecto positivo y estadísticamente significativo del programa de 2,53% de mayor rendimiento, pero sólo cuando se aplica el modelo multinivel. Sí se debe observar que cuando se desagregan las variables de resultado por grado de dificultad de las preguntas o por el eje temático al cual pertenecen, no es tan evidente que las niñas obtengan un resultado sistemáticamente superior al de los niños: si bien las magnitudes de los efectos son claramente mayores en valor absoluto, en unos casos parecen favorecer a los niños y en otros a las niñas. Esto podría estar condicionado por la oportunidad en que fueron desarrollados los ejes temáticos bajo un programa en relación al otro.

Al estimar el efecto del programa utilizando las sub muestras relacionadas con los resultados promedio en las pruebas SIMCE de matemáticas y los promedios de las pruebas SIMCE de lenguaje y matemáticas de los años 2007 a 2009, encontramos que el único caso en que se observa un efecto estadísticamente significativo del programa a favor de los niños sujetos al programa es entre los niños y establecimientos con un nivel de logro intermedio en dichas pruebas. En los demás casos no se detectaron efectos estadísticamente distintos de cero.

Se debe tener presente que esta medición del efecto del programa Singapur en el aprendizaje de matemáticas de los niños de 2º básico de la muestra de establecimientos estudiados, constituye una evaluación de corto plazo, dado el escaso tiempo transcurrido entre la aplicación del programa y el período de evaluación, considerando, además, que los procesos educativos suelen tomar algún tiempo en madurar y mostrar sus efectos. En particular, el Centro Félix Klein (2012) observa que los textos del método Singapur de 1º y 2º básico no resultarían disruptivos respecto de la cultura escolar instalada en este sub-ciclo de la enseñanza, pero que los textos de 3º y 4º básico sí incorporarían claras diferencias respecto de la práctica habitual de la enseñanza en estos cursos, resultando más complejos y distintos de lo que los profesores de estos cursos están acostumbrados a enseñar, planteando la hipótesis de que el resultado de la presente evaluación, levemente favorable para las escuelas que utilizaron los textos del método Singapur en 2º básico, podría ser más nítido de observar en cursos posteriores.

Es por ello que la aplicación de la prueba SIMCE de matemáticas de 4º básico en el año 2014, aplicada a la misma cohorte que rindió la presente prueba, permitirá verificar o no esta hipótesis planteada por el Centro Félix Klein.

- Anexos

Anexo N° 1. Principales resultados de la evaluación de implementación

Entre noviembre de 2011 y abril de 2012 se efectuó una evaluación de implementación del programa, ejecutada por Adimark GfK, mediante la aplicación de 283 entrevistas a docentes que impartían matemáticas en 1° o 2° básico en 176 Establecimientos y mediante la aplicación de 295 entrevistas telefónicas a Jefes de UTP en 295 establecimientos. Además se ejecutaron 3 grupos focales con docentes de establecimientos que aplicaban el Programa "Singapur" en la Región Metropolitana. Los objetivos del estudio fueron: i) levantar información acerca de la distribución y uso de los textos escolares de matemática "Textos Singapur" en los establecimientos educacionales que participaron del programa, ii) levantar información sobre la pertinencia de la capacitación que los docentes realizan para trabajar con estos textos y iii) levantar información sobre la valoración que tienen los docentes del material del Programa Singapur entregado a sus establecimientos educacionales.

Entre los principales hallazgos sobre la implementación del programa, es posible destacar lo siguiente:

- Los resultados obtenidos tanto de jefes de UTP como de los docentes entrevistados, dieron cuenta de algunos problemas en el período de recepción del material: el 16,6% de los Jefes de UTP y 31,8% de los Docentes entrevistados, señalaron haber recibido el material relacionado al Programa una vez empezado el año escolar. Esta situación podría haber afectado la implementación del programa, y podría dar cuenta tanto de problemas reales en los plazos de distribución a los establecimientos, como de problemas de distribución interna del material al interior de estos.
- En relación a la cantidad de material recibido por los establecimientos, la mayor parte de ellos afirmó haber recibido la cantidad de material requerido. De los materiales que componían el programa, los Jefes de UTP señalaron a los Libros para el alumno como aquel recibido en menor proporción en la cantidad requerida por el establecimiento (16,6%). Mientras que los Docentes destacaron al Kit didáctico como el componente con mayores deficiencias en términos de cantidad de material recibido (20,5%). Dado que el material fue otorgado por el MINEDUC sobre la base de la matrícula de estudiantes del año 2010, es posible pensar que modificaciones en términos de aumento de matrícula repercutieron negativamente al respecto, lo que podría sugerir la entrega de material extra a cada establecimiento.
- Sólo en el 50,5% de los establecimientos habría un encargado que sirve de apoyo y guía para los docentes. No obstante, tanto jefes de UTP (95,6%) como docentes (88,3%) reconocieron que existía coordinación entre los profesores para aplicar el programa. Estos resultados indicarían que el soporte en términos de apoyo en la implementación del programa se dio principalmente entre colegas, por sobre el apoyo institucional que brindaría el establecimiento.
- El elemento inicial con el cual se dio inicio a la implementación del programa consistió en un proceso de capacitación para docentes de establecimientos seleccionados, gestionado inicialmente por el MINEDUC. En términos de la cobertura alcanzada por este proceso, de los docentes entrevistados el 83,0% señaló haber sido capacitado acerca del programa Singapur, siendo las capacitaciones oficiales organizadas por el MINEDUC aquellas con mayor cobertura. Las capacitaciones recibidas lograron positivas evaluaciones, no obstante la gran mayoría de los

docentes señaló como necesario o muy necesario el recibir una capacitación para lograr una correcta y efectiva aplicación del programa.

- A pesar de que el proceso de capacitación logró una buena cobertura y valoración, los docentes demandaron más y mejores capacitaciones, considerando que estas deberían ser permanentes en el tiempo. Esta situación podría indicar cierta inseguridad entre los docentes frente al cambio de metodología de enseñanza del programa, así como la necesidad del establecimiento de ser reforzado durante el año escolar en el proceso de aplicación en aula del programa.
- En ese sentido, en la fase cualitativa se advirtieron algunas diferencias según el tipo de capacitación recibida por los profesores: los que recibieron capacitación directa manifiestan sentirse mucho más seguros a la hora de aplicar el método, ya que tenían conocimiento de los textos que utilizarían, los materiales y en general los contenidos, lo que les permitió manejar los tiempos de manera adecuada y diseñar estrategias para concretar la entrega de contenidos; mientras que los profesores que recibieron capacitación indirecta, tuvieron dificultades para utilizar los materiales y aplicar los contenidos, por lo que tuvieron que crear estrategias para abarcar la cantidad de contenidos, lo que muchos reconocen no haber logrado.

En relación al uso y valoración del programa se encontró que:

- Hubo una alta frecuencia de utilización de los materiales del programa, los que fueron utilizados por los docentes en todas o las mayorías de sus clases.
- Junto a la utilización del material del programa dentro de sus clases, los docentes declararon utilizar metodologías adicionales a las incluidas en el texto Singapur (68,6%), material complementario a los textos (70,9%), así como material didáctico complementario al entregado por el programa (63,3%). Esta amplia utilización de componentes adicionales al programa, podría dar cuenta de deficiencias en términos de su contenido o material, así como de inseguridad de parte de los docentes respecto a su implementación (que les llevaría a utilizar lo ya conocido), o bien de desconocimiento respecto a la correcta aplicación del mismo.
- Los resultados de la evaluación del componente valoración del programa Singapur, dieron cuenta de una alta aceptación y valoración del programa tanto por parte de los docentes, como por el establecimiento en general. En efecto, al indagar en la aceptación del programa por parte del cuerpo docente los resultados dieron cuenta de un alto nivel de aceptación. Los profesores indicaron estar motivados con la implementación del programa y consideraron que el programa facilitó su labor de enseñanza. No obstante, se reconoció que la implementación requirió esfuerzos adicionales de parte del profesor en el aula, así como que las clases eran más cansadoras para el profesor. En esta línea, alrededor del 20% los Jefes de UTP declararon percibir a un segmento de profesores que mostrarían resistencia al cambio de metodología de enseñanza del programa, y/o dificultades para trabajar con el material dentro de sus clases. La existencia de este segmento dentro de un contexto de alta valoración y aceptación del programa, podría respaldar la necesidad de mayor cobertura de las capacitaciones, monitoreo y apoyo externo en la implementación del programa.
- En concordancia con lo anterior, todos los profesores participantes de la fase cualitativa manifestaron una postura crítica en relación a las capacitaciones en las que participaron, ya que consideraron que las capacitaciones fueron insuficientes y además de mala calidad, en tanto no recibieron materiales, modelamiento de clases ni ejemplos concretos, por lo que creen que su experiencia con el Método Singapur podría haber sido más satisfactoria de contar con más herramientas.

- La evaluación del programa y de los materiales que lo componen en términos de notas, arrojaron resultados de excelencia. De los materiales evaluados, aquellos que lograron las mejores evaluaciones correspondían al material de uso personal del alumno, por sobre la evaluación del Kit de material didáctico. Respaldo la buena evaluación de los textos; la amplia mayoría de los entrevistados consideró a los textos Singapur como mejores a los entregados anteriormente por el MINEDUC.
- En general, todos los profesores participantes de la fase cualitativa destacaron las guías didácticas como una gran ayuda para llevar a cabo el Método Singapur, siendo en algunos casos la única fuente de conocimiento del método.
- En relación a los resultados percibidos en los estudiantes tras la aplicación del programa, los resultados fueron nuevamente alentadores. Cerca del 90% de los docentes señaló haber percibido mayor interés y motivación con la asignatura de matemáticas, así como un mejor ambiente de aprendizaje de los alumnos. Este mejor ambiente se plasmaría en la percepción de una mejor comprensión de los contenidos de la asignatura, mejorías en la seguridad que sienten los alumnos a la hora de resolver problemas, así como mejores resultados académicos de los alumnos.
- De manera similar, en la fase cualitativa todos los profesores destacaron espontáneamente los efectos positivos que el Método Singapur tuvo sobre los niños y niñas, identificando mejoras tanto en el aprendizaje de las matemáticas como a nivel de la personalidad de los niños, lo que se atribuyó a un efecto del método en sí mismo, incentivando el razonamiento matemático, y además de la metodología de trabajo grupal. Por otra parte, los efectos de Singapur a nivel de la personalidad de los niños, repercutieron de manera positiva en el comportamiento de los niños en los demás subsectores, pues se mantenía la dinámica de participación, respeto y capacidad argumentativa que iban desarrollando.

Finalmente, las sugerencias más destacadas generadas a partir del estudio fueron las siguientes:

- Realizar capacitaciones a los profesores, ampliando cobertura y frecuencia durante el año escolar, así como mejorando la calidad de éstas. Esta demanda se sustentó en la percepción que tienen los profesores de que muchas de las falencias y dificultades que tuvieron que enfrentar en su experiencia con Singapur se debió a su falta de preparación y de capacidad de respuesta frente a los inconvenientes surgidos.
- Apoyo, asesoría, monitoreo y supervisión desde el MINEDUC a la implementación del programa, lo que en términos concretos implica contar con una persona que tenga dedicación específica para Singapur, a quién ellos se puedan dirigir en caso de tener dudas o para entregar sugerencias frente a los problemas cotidianos a los que se enfrentan. En este contexto, también esperan tener acceso a evaluaciones estandarizadas, para evitar mezclar contenidos y además comparar resultados entre colegios/escuelas y cursos, lo que permitiría también ir optimizando la aplicación del método.
- Aumento y entrega oportuna de materiales, en especial del material didáctico permitiendo reducir el número de niños que trabajan grupalmente con este material, lo que en paralelo repercutiría en favorecer un mejor aprendizaje para los niños ya que todos tendrían contacto directo con el material concreto.
- Corrección en términos de contenido y extensión del programa.

Anexo N° 2. Verificación del cumplimiento de requisitos de los establecimientos para participar en el programa

Los 482 establecimientos elegibles para participar en el programa presentan, en la prueba SIMCE de matemáticas, un promedio de 274,7 puntos entre 2007 y 2009, siendo el puntaje mínimo de 250 puntos, tal como lo exigían las condiciones para participar en el programa (Tabla 1.1).

Tabla 2.1: Puntajes SIMCE de matemáticas 2007-2009 en establecimientos elegibles

Variable	N° establecimientos	Promedio SIMCE 2007-2009	Desviación estándar	Mínimo	Máximo
SIMCE promedio 2007-2009	482	274,5	17,7	250	326,3

La matrícula, entre 2007 y 2009, era de 20 ó más alumnos en 474 de los 482 establecimientos elegibles y sólo 8 de ellos tenían una matrícula inferior a 20 alumnos (1,66% de los establecimientos). El resto cumplía la condición exigida por el programa, tal como muestra la Tabla 1.2:

Tabla 2.2: Matrícula 2007-2009 en establecimientos elegibles

Alumnos	N° establecimientos	% sobre el total
Menos de 20	8	1,66
20 ó más	474	98,34
Totales	482	100

La tercera exigencia para ser elegible para el programa era que la matrícula de 1° básico del 2010 debía ser superior a 20 alumnos, aspecto que, revisados los datos administrativos, cumplían 480 establecimientos (99,59%).

Tabla 2.3: Matrícula 2010 de 1° básico en establecimientos elegibles

Variable	N° establecimientos	Promedio Matrícula 1° básico 2010	Desviación estándar	Mínimo	Máximo
Matrícula 1° básico 2010	482	71,9	43,83	18	317

La cuarta exigencia era que el 40% de los establecimientos fueran de la región metropolitana y el resto de las demás regiones, cuestión que el programa también respetó, como indica la Tabla 1.4:

Tabla 2.4: Localización geográfica de los establecimientos elegibles

Región	N° establecimientos	Porcentaje del total
Otras regiones	279	58
Región Metropolitana	203	42
Totales	482	100

Se exigía al programa representar las proporciones nacionales de establecimientos según dependencia, que en 2009 correspondía a una relación de 1,07 establecimientos municipales en relación a los particulares subvencionados. En el programa, sin embargo, la proporción alcanzó a 0,81 establecimientos municipales por cada particular subvencionado.

Finalmente, dado que por los criterios de selección del programa no fue posible replicar la distribución nacional por GSE, sí se debía cumplir que el programa tuviera, al menos, representación de establecimientos de todos los GSE. Entre los 482 establecimientos elegibles se constata el cumplimiento de esta exigencia, según muestra la Tabla 1.5.

Tabla 2.5: Nivel socio económico de los establecimientos elegibles

Grupo socio económico	N° establecimientos	Porcentaje de establecimientos
A (bajo)	15	3,11
B (medio bajo)	153	31,74
C (medio)	206	42,74
D (medio alto)	96	19,92
E (alto)	12	2,49
Totales	482	100

Anexo N° 3. Establecimientos en los que se aplicó la prueba

N°	RBD	Nombre establecimiento	Región	Tratamiento
1	522	Colegio Javiera Carrera	4	No Singapur
2	579	Colegio Seminario Conciliar	4	No Singapur
3	13551	Colegio Manantial	4	No Singapur
4	13606	Colegio Leonardo Da Vinci - Vicuña	4	No Singapur
5	1371	Escuela Básica Arauco	5	No Singapur
6	1734	Colegio Rubén Castro-Diurno	5	No Singapur
7	2112	Escuela Municipal Moisés Mussa	6	No Singapur
8	2264	Escuela Nuestra Sra. de La Merced	6	No Singapur
9	2330	Esc. Carmen Gallegos de Roble	6	No Singapur
10	2400	Escuela Particular Amada Sofía García	6	No Singapur
11	15596	Instituto San Vicente de Tagua Tagua	6	No Singapur
12	15758	Colegio Christian College	6	No Singapur
13	2744	Escuela España	7	No Singapur
14	2782	Colegio San Ramón Nonato	7	No Singapur
15	2803	Escuela Las Arboledas	7	No Singapur
16	2948	Escuela Carlos Salinas Lagos	7	No Singapur
17	2991	Colegio Monseñor Manuel Larraín	7	No Singapur
18	3298	Liceo Nuestra Señora del Rosario	7	No Singapur
19	16502	Colegio Cordillera	7	No Singapur
20	4636	Escuela Marcela Paz	8	No Singapur
21	4725	Escuela Villa Independencia	8	No Singapur
22	4831	Escuela Básica Bellavista	8	No Singapur
23	18014	Escuela Básica Los Conquistadores	8	No Singapur
24	6500	Escuela Darío Salas Díaz	9	No Singapur
25	7338	Escuela Leonila Folch Lopez	10	No Singapur
26	7406	Colegio San Mateo de Osorno	10	No Singapur
27	7434	Escuela Particular Santa Cruz	10	No Singapur
28	7773	Escuela Eulogio Goycolea Garay	10	No Singapur
29	7775	Escuela Rural Pargua	10	No Singapur
30	22343	Colegio San Sebastián	10	No Singapur
31	8634	Liceo Ruiz Tagle	13	No Singapur
32	8821	Liceo Municipalizado Amanda Labarca	13	No Singapur
33	9205	Escuela Particular The New School	13	No Singapur
34	9967	Escuela Particular Rubén Darío	13	No Singapur
35	10085	Escuela Estado de Florida	13	No Singapur
36	10187	Escuela Particular Madre A. Eugenia	13	No Singapur
37	10259	Liceo Agustín Edwards	13	No Singapur
38	10338	Escuela Básica Santo Domingo	13	No Singapur
39	10497	Esc. Particular Josefina Gana de J.	13	No Singapur
40	10726	Centro de Educación General Básica	13	No Singapur
41	10753	Escuela Republica Checa	13	No Singapur
42	11906	Escuela Bas. Los Carrera	13	No Singapur
43	25282	Colegio Poliv. Cardenal Raúl Silva Henríquez	13	No Singapur
44	25508	Colegio Part. Gabriela Mistral Sociedad Educa	13	No Singapur
45	25580	Colegio Particular Mirador	13	No Singapur
46	25722	Colegio Part. Patrona Señora de Lourdes	13	No Singapur
47	25724	Colegio Particular Antilhue	13	No Singapur
48	25781	Colegio Poliv. Arzobispo Crescente Errázuriz	13	No Singapur
49	25785	Colegio Part. Royal American School	13	No Singapur
50	6899	Liceo Republica del Brasil	14	No Singapur
51	12534	Liceo Colegio Metodista Robert Johnson	1	Singapur
52	12567	Colegio San Antonio de Matilla	1	Singapur

N°	RBD	Nombre establecimiento	Región	Tratamiento
53	12669	Colegio Hispano Británico Iquique	1	Singapur
54	637	Escuela Lucila Godoy Alacayaga	4	Singapur
55	712	Escuela Básica Arturo Alessandri Palma	4	Singapur
56	13439	Colegio Pucara	4	Singapur
57	13542	Colegio Irma Salas Silva	4	Singapur
58	1569	Escuela Piloto 1° Luis Pardo Villalón	5	Singapur
59	1609	Esc. Part Fray Luis Beltrán	5	Singapur
60	14265	Escuela Saint Benedict College	5	Singapur
61	14363	Colegio Hispano	5	Singapur
62	2386	Escuela Part. Luis García de La Huerta	6	Singapur
63	2459	Colegio Giuseppe Bertoluzzi de Felip	6	Singapur
64	2795	Escuela Teno	7	Singapur
65	2796	Escuela Comalle	7	Singapur
66	2817	Escuela San Cristóbal	7	Singapur
67	2871	Escuela Lo Valdivia	7	Singapur
68	3002	Escuela Particular María Mazzarello	7	Singapur
69	3111	Escuela Villa Culenar	7	Singapur
70	3256	Escuela Ramón Belmar Saldías	7	Singapur
71	3266	Escuela San Antonio	7	Singapur
72	3297	Escuela Particular Prospero Villar Bravo	7	Singapur
73	3314	Escuela Maitencillo	7	Singapur
74	3433	Colegio Concepción de Parral	7	Singapur
75	3531	Instituto Regional del Maule	7	Singapur
76	16448	Instituto Santa Marta	7	Singapur
77	16477	Colegio Aquelarre	7	Singapur
78	3663	Escuela Republica de México	8	Singapur
79	3733	Escuela Particular Gabriela Mistral	8	Singapur
80	3797	Colegio Sagrado Corazón de Jesús	8	Singapur
81	4308	Colegio Abel Inostroza Gutiérrez	8	Singapur
82	4662	Colegio San Ignacio	8	Singapur
83	4672	Liceo Tecn. Profesional Jorge Sánchez U.	8	Singapur
84	4723	Escuela México Estado de Guerrero	8	Singapur
85	4841	Escuela Básica Cerro Estanque	8	Singapur
86	5029	Domingo Faustino Sarmiento	8	Singapur
87	5085	Escuela Ramiro Roa González	8	Singapur
88	11712	Escuela Particular Gaspar Cabrales	8	Singapur
89	5442	Escuela Básica Emilia Romagna	9	Singapur
90	5539	Escuela de Los Sauces	9	Singapur
91	5813	Colegio Centenario	9	Singapur
92	20180	Esc. Cristiana Desarrollo Integral	9	Singapur
93	20205	Escuela Particular Habit-Art	9	Singapur
94	7332	Escuela Juan Ricardo Sánchez	10	Singapur
95	7346	Escuela N_ 46 Italia	10	Singapur
96	7544	Escuela Básica Riachuelo	10	Singapur
97	7630	Escuela España	10	Singapur
98	22519	Colegio Mirador del Lago	10	Singapur
99	22608	Colegio Labrador	10	Singapur
100	8521	Escuela Carlos Condell de La Haza	13	Singapur
101	8536	Escuela Republica de Israel	13	Singapur
102	8825	Colegio Antártica Chilena	13	Singapur
103	9084	Esc. Presidente Eduardo Frei Montalva	13	Singapur
104	9104	Escuela Villa Macul	13	Singapur
105	9317	Instituto La Salle	13	Singapur
106	9426	Escuela Llano Subercaseaux	13	Singapur

N°	RBD	Nombre establecimiento	Región	Tratamiento
107	9642	Escuela Particular Tte. Dagoberto Godoy	13	Singapur
108	9655	Escuela Particular Claudio Matte	13	Singapur
109	9800	Escuela Bas. Rosa Elvira Matte de Prieto	13	Singapur
110	9891	República de Guatemala	13	Singapur
111	9892	Escuela Las Américas	13	Singapur
112	9941	Colegio Particular de Asís	13	Singapur
113	10006	Escuela Republica de La India	13	Singapur
114	10030	Liceo Poliv. San José de La Preciosa Sang	13	Singapur
115	10151	Colegio Cristian And Caren School	13	Singapur
116	10351	Escuela Particular Parque Asunción	13	Singapur
117	10363	Escuela Particular Jaime Eyzaguirre	13	Singapur
118	10407	Escuela Básica Algarrobal	13	Singapur
119	10415	Escuela Particular Desiree	13	Singapur
120	10662	Escuela Particular Saint Mary	13	Singapur
121	10671	Escuela Carmen Reveco Núñez	13	Singapur
122	10676	Escuela Francisco Letelier Valdés	13	Singapur
123	10689	Escuela El Roto Chileno	13	Singapur
124	10748	Escuela Emilia Lascar	13	Singapur
125	11942	Colegio Abersan	13	Singapur
126	24424	Colegio Particular La Unión	13	Singapur
127	24427	Colegio San Viator	13	Singapur
128	24486	Colegio Polivalente San Rafael	13	Singapur
129	24648	Colegio Los Nogales	13	Singapur
130	24956	Colegio Técnico Profesional Necedal	13	Singapur
131	25016	Esc. Básica San Marcelo Anexo	13	Singapur
132	25223	Colegio Monte de Asís	13	Singapur
133	25558	Colegio Particular Altazor	13	Singapur
134	25613	Colegio Pedro Apóstol	13	Singapur
135	25675	Colegio Part. Corazón de María de San Miguel	13	Singapur
136	25725	Escuela Part. Nahuel	13	Singapur
137	26017	Escuela Bas. Kings School	13	Singapur
138	26033	Colegio Poliv. San Carlos de Quilicura	13	Singapur
139	26035	Escuela Part. San José de Lampa	13	Singapur
140	6770	Colegio de Música Juan Sebastián Bach	14	Singapur
141	6774	Escuela El Laurel	14	Singapur
142	52	Liceo Domingo Santa María	15	Singapur
143	12570	Colegio Alta Cordillera	15	Singapur

Anexo N° 4. Verificación del cumplimiento de requisitos de los establecimientos considerados en la evaluación

El promedio SIMCE de matemáticas de 4° básico de los años 2007 a 2009 de los 143 establecimientos a evaluar era de 274,7 puntos, siendo el puntaje mínimo de 250 puntos, tal como lo exigían las condiciones para participar en el programa:

Tabla 4.1: Puntajes SIMCE de matemáticas 2007-2009 en establecimientos a evaluar

Variable	N° establecimientos	Promedio SIMCE 2007-2009	Desviación estándar	Mínimo	Máximo
SIMCE Promedio 2007-2009	143	274,7	18,21	250,0	313,3

Al revisar la matrícula de los 143 establecimientos entre 2007 y 2009, se constató que sólo 3 de ellos tenían una matrícula inferior a 20 alumnos (1 establecimiento con 15 alumnos y 2 con 18), lo que representa al 2% de la muestra evaluada. El resto cumplió la condición exigida por el programa, tal como indica la Tabla 3.2:

Tabla 4.2: Matrícula 2007-2009 en establecimientos a evaluar

Alumnos	N° establecimientos	% sobre el total
Menos de 20	3	2,09
20 ó más	140	97,91
Totales	143	100

La tercera exigencia era que la matrícula de 1° básico del 2010 debía ser superior a 20 alumnos, cuestión que, revisados los datos administrativos, todos los establecimientos cumplían, como se indica en la Tabla 3.3:

Tabla 4.3: Matrícula 2010 de 1° básico en establecimientos a evaluar

Variable	N° establecimientos	Promedio Matrícula 1° básico 2010	Desviación estándar	Mínimo	Máximo
Matrícula 1° básico 2010	143	70	43,52	20	317

La cuarta exigencia era que el 40% de los establecimientos fuera de la región metropolitana y el resto de las demás regiones, cuestión que la muestra de establecimientos a evaluar también cumplía, como se indica a continuación:

Tabla 4.4: Localización geográfica de los establecimientos a evaluar

Región	N° establecimientos	Porcentaje del total
Otras regiones	84	58,74
Región Metropolitana	59	41,26
Totales	143	100

Se exigía al programa representar las proporciones de establecimientos según dependencia, que en 2009 correspondía a una relación de 1,072 establecimientos municipales en relación a los particulares subvencionados. En la muestra de esta evaluación, sin embargo, la proporción, seleccionada aleatoriamente, arrojó una proporción de 0,77 establecimientos municipales por cada particular subvencionado.

Finalmente, dado que no fue posible replicar la distribución nacional por GSE, como se indicó, se debía cumplir que el programa sí tuviera representación de establecimientos

de todos los GSE, cuestión los 143 establecimientos a evaluar también cumplieron según indica la Tabla 3.5:

Tabla 4.5: Nivel socio económico de los establecimientos a evaluar

Grupo socio económico	N° establecimientos	Porcentaje de establecimientos
A (bajo)	5	3,50
B (medio bajo)	45	31,47
C (medio)	64	44,76
D (medio alto)	24	16,78
E (alto)	5	3,50
Totales	143	100

Anexo N° 5. Prueba aplicada

[Insertar aquí la prueba]

Anexo N° 6. Resumen de los principales hallazgos del Instituto Félix Klein

○ Principales hallazgos del Centro Félix Klein

El documento preparado por CFK contiene el detalle de la información de esta sección, que compara los resultados en la prueba entre los alumnos de establecimientos acogidos al programa Singapur y alumnos de establecimientos que no lo hicieron.

CFK logró determinar que el porcentaje de logro promedio de los estudiantes que rindieron la prueba fue del 60% (59% en el grupo no Singapur, 60% para el grupo de establecimientos del programa), promedio que se puede categorizar como bueno.

Si bien no se encontraron diferencias significativas entre los porcentajes de logro promedios obtenidos por el grupo Singapur y el grupo no Singapur, en los ejes de Operatoria y Datos y Probabilidad, sí observaron diferencias significativas. En ambos casos se trata de ejes que incorporan contenidos nuevos para la cultura escolar en este nivel de la enseñanza, pues los problemas multiplicativos del eje Operatoria eran tratados, en el currículo anterior, a partir de 3º básico, mientras que el eje de Datos y Probabilidad no tiene precedentes. Al analizar los resultados de este eje, encontraron que el grupo Singapur está por sobre el grupo no Singapur en 5 puntos porcentuales, aunque con un porcentaje de logro regular, cercano al 50%. En el caso del eje Operatoria, la diferencia entre ambos grupos fue de 3 puntos porcentuales a favor del grupo Singapur. De este modo, ahí donde aparecen temas nuevos para la cultura escolar, el Método Singapur hizo una diferencia positiva en términos de porcentajes de logro promedio, aunque todavía incipiente. Si bien el libro de texto distribuido por el Ministerio de Educación a todas las escuelas del país está ajustado a estos cambios curriculares, los textos del Método Singapur de 2º básico abordan el estudio de los problemas multiplicativos a lo largo de todo el texto, de manera integrada a los problemas aditivos, siguiendo la estrategia de enseñanza iniciada en 1º básico.

Al analizar los resultados por eje temático CFK observó que los dos ejes que presentan menor porcentaje de logro a nivel poblacional son Medición (38%) y Datos y Probabilidad (43%). Estos ejes son también los de menor porcentaje de logro promedio en ambos grupos. El eje Datos y Probabilidad presentó, además, la mayor desviación estándar.

Sin embargo, el mejor porcentaje de logro promedio por eje temático, no fue el mismo en ambos grupos. Mientras que en el grupo no Singapur el mayor porcentaje se alcanzó en el eje de Números (67%), en el grupo Singapur lo hizo en el eje de Patrones y Álgebra (68%).

En cuanto a las cuatro habilidades matemáticas consideradas en el estudio, en general se obtuvieron porcentajes de logro levemente superiores en el grupo Singapur, salvo en la habilidad de Resolver Problemas en que el grupo no Singapur resultó levemente superior.

La habilidad de manipulación de expresiones matemáticas es la que presentó un mayor porcentaje de logro promedio en ambos grupos. Las escuelas Singapur obtuvieron un 69% de logro promedio, mientras que las no Singapur un 65%, siendo esta diferencia estadísticamente significativa. Asimismo, la habilidad de Resolver Problemas presentó un menor porcentaje de logro promedio, también en ambos grupos; las escuelas

Singapur obtuvieron un 47% de logro promedio, mientras que las de control un 49%, aunque esta diferencia no resultó estadísticamente significativa.

La mayor cantidad de estudiantes que rindió la prueba se concentró en el Nivel de Logro Intermedio (73,6% de estudiantes). Esta mayor concentración en el nivel intermedio se mantuvo en cada grupo; en las escuelas de Singapur fue de un 73,2%, y en las de control de un 74,5%. El porcentaje de estudiantes en los niveles de logro inicial y avanzado resultó claramente inferior en ambos grupos. En el caso del nivel inicial, el porcentaje de estudiantes en el grupo no Singapur fue de 11,8% (342 alumnos de 2.897), mientras que en el grupo Singapur de 10,7% (637 estudiantes de un total de 5.951). Esta diferencia en la distribución, sin embargo, no resultó significativa estadísticamente. En el caso del nivel avanzado, la diferencia entre ambos grupos fue mayor a favor del grupo Singapur. Mientras que la cantidad de estudiantes en nivel avanzado en el grupo no Singapur fue del 13,7%, en el grupo Singapur fue de 16,1%. Esta diferencia entre cantidad de estudiantes por niveles de logro según grupo de escuela, resultó estadísticamente significativa únicamente en el nivel avanzado. Este resultado es particularmente interesante puesto que señala que, si bien el método Singapur no muestra diferencias significativas entre los grupos en el nivel inicial ni en el intermedio, sí las muestra en el nivel avanzado. Así, los resultados parecen indicar que los estudiantes de mejores desempeños, son los que se favorecen mayoritariamente con la aplicación del método Singapur.

Haciendo un análisis por escuela, se observó que los mejores resultados en términos de porcentajes de logro promedio se concentraron mayoritariamente en escuelas que utilizaron los libros del método Singapur. En efecto, de las 143 escuelas que participaron en el estudio (93 del programa Singapur y 50 no Singapur), 17 de ellas obtuvieron sobre un 70% de logro promedio. De estas 17 escuelas, 14 resultaron ser del grupo Singapur, mientras que sólo 3 del grupo no Singapur. Si bien la cantidad de escuelas de la muestra que pertenecían al grupo Singapur (64% eran escuelas Singapur, mientras que el 36% eran del grupo no Singapur), el 82% de las escuelas que obtuvieron mejores resultados eran del primer grupo y sólo el 18% del segundo. Este cambio en la distribución de porcentajes a favor de las escuelas Singapur (pasaron del 64% de la muestra al 82% de las escuelas con mejores resultados), da cuenta de que el método Singapur, a nivel de escuela, favoreció que determinadas escuelas del grupo Singapur alcanzaran un nivel de logro destacado.

Por otro lado, al observar las 16 escuelas que obtuvieron un nivel de logro promedio insuficiente (bajo un 49%), CFK indica que la mitad de ellas pertenecía a un grupo y la otra mitad al otro, lo que señalaría qué tipo de escuelas podrían aprovechar mejor al método Singapur. Sería conveniente, señala CFK, acompañar este resultado con estudios que permitieran conocer, de un modo más amplio y profundo, distintos aspectos de la "realidad de la escuela", de los docentes, de los estudiantes y de los padres entre otras, que podrían explicar estos resultados.

Al analizar los porcentajes de logro obtenidos por los estudiantes, tomando en cuenta el Nivel Socio-Económico de las escuelas, se determinó que, a mayor nivel socio-económico, se lograron mejores resultados en la prueba. Este es un resultado concordante con el comportamiento nacional de los resultados en otras pruebas estandarizadas, como el SIMCE. En el grupo medio y medio alto se generaron diferencias significativas a favor del grupo Singapur, mientras los demás grupos socioeconómicos no presentaron diferencias significativas.

Tomando en cuenta los resultados del análisis por dependencia administrativa de las escuelas, se obtuvo una diferencia significativa entre establecimientos municipales y particulares subvencionadas, con un 56% y 62% de logro promedio respectivamente, pero no hubo diferencias significativas entre los porcentajes de logro promedio al comparar establecimientos municipales y particulares subvencionados entre ambos grupos. Este es un resultado consistente con el desempeño de las escuelas en las pruebas SIMCE, considerando el tipo de dependencia a nivel nacional.

A partir de los resultados de la encuesta aplicada a los docentes (ver detalles en anexo N° 4), específicamente a los docentes que implementaron el método Singapur, y estableciendo una relación con los resultados de los estudiantes en la prueba, se observó que quienes tuvieron docentes capacitados en el método obtuvieron mejor porcentaje de logro promedio respecto de los estudiantes cuyos docentes no recibieron capacitación en dicho método. Esta diferencia resultó significativa al 5%. Es decir, la capacitación hizo una diferencia significativa. En este mismo sentido, los estudiantes de aquellos docentes que declararon haber utilizado solamente el texto "Pensar Sin Límites", obtuvieron un porcentaje de logro promedio mayor respecto de aquellos estudiantes cuyos docentes utilizaron material adicional (por ejemplo, otros textos). Es decir, la utilización del texto también hizo una diferencia significativa.

Finalmente, el CFK indica que si bien para los profesores chilenos de 1° a 4° básico capacitados en el uso de los libros de textos del programa, los textos de 1° y 2° básico resultan abordables y manejables, los textos de 3° y 4° resultarían más complejos y difíciles de implementar, aun para profesores que hayan tenido capacitación sistemática. Ello en parte porque los textos del método Singapur de 1° y 2° básico no resultarían disruptivos respecto de la cultura escolar instalada en este sub-ciclo de la enseñanza. Por lo general incorporan contenidos y modos de enseñanza relativamente cercanos para los profesores que hacen clases en estos niveles. Sin embargo, los textos de 3° y 4° básico incorporan claras diferencias respecto de la práctica habitual de la enseñanza en estos cursos. Resultan más complejos y distintos de lo que los profesores de estos cursos están acostumbrados a enseñar, tanto desde la perspectiva matemática como didáctica. Es claro que estas mayores dificultades se relacionan también con las mayores complejidades provenientes de la propia disciplina, y que el currículo va incorporando a medida que avanza por los distintos cursos de la enseñanza básica. Sin embargo, de modo general se puede establecer que los textos Singapur "avanzan más rápido" que las bases curriculares de básica de nuestro país. Esto es, hay contenidos en los textos Singapur para un determinado curso que las bases curriculares incorporan en cursos posteriores. En consecuencia, el resultado que este informe ha documentado, levemente favorable para las escuelas que utilizaron los textos del método Singapur en 2° básico, podría ser más nítido de observar en cursos posteriores.

Finalmente, concluye CFK, sería necesario añadir que, según la investigación actualizada, a dos años de implementación de una innovación en la enseñanza, es prematuro esperar un impacto significativo en resultados de aprendizaje. Se trata de un plazo muy inmediato. En ese corto período de tiempo, las investigaciones más bien suelen documentar sobre las grandes dificultades y resistencias que presentan los profesores a la hora de adoptar una nueva metodología de enseñanza, que conlleva en muchos casos incluso a un leve descenso en los resultados de aprendizaje, aunque los resultados de este estudio, en particular, muestran que no ha existido tal rechazo por

parte de docentes y estudiantes, puesto que los porcentajes de logro de niños y niñas que usaron los textos Singapur, resultaron mejores, aunque todavía con una diferencia leve, que los de sus pares que no lo hicieron.

– Bibliografía

Cameron, A. C. & Trivedi, P. (2010). "Microeconometrics using stata", *Stata Press, Texas, Revised Edition*.

Centro Félix Klein (2012). *Servicio de elaboración y aplicación de prueba para medir la efectividad en los aprendizajes del Proyecto Singapur. Informe Final*. Universidad de Santiago de Chile.

Kreft, I. & de Leeuw. J (2002). "Introducing multilevel modeling", *SAGE Publications, London*.

Mullis, Ina V.S. et al. (2000). *TIMSS 1999. International Mathematics Report, Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*, p. 32. http://timss.bc.edu/timss1999i/pdf/T99i_Math_All.pdf

_____ (2004). *TIMSS 2003. International Mathematics Report, Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*, p. 34. http://timss.bc.edu/PDF/t03_download/T03INTLMATRPT.pdf

Singer, J. (1998). "Using SAS PROC MIXED to fit multilevel models, hierarchical models and individual growth models", *Journal of Education and Behavioral Statistics, Winter 1998, vol. 24, N° 4, pp. 323-355*.