



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
ESCUELA DE INGENIERIA

**PESO RELATIVO DE LOS EJES Y  
HABILIDADES CURRICULARES EN LA  
PREDICCIÓN DE ÉXITO ACADÉMICO SIMCE:  
UN ESTUDIO LONGITUDINAL A ESCALA  
NACIONAL**

**MACARENA PAZ SANTANA SEPÚLVEDA**

Tesis para optar al grado de  
Magister en Ciencias de la Ingeniería

Profesor Supervisor:  
**MIGUEL NUSSBAUM**

Santiago de Chile, Enero, 2015

© 2014, Macarena Santana Sepúlveda



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
ESCUELA DE INGENIERIA

**PESO RELATIVO DE LOS EJES Y  
HABILIDADES CURRICULARES EN LA  
PREDICCIÓN DE ÉXITO ACADÉMICO SIMCE:  
UN ESTUDIO LONGITUDINAL A ESCALA  
NACIONAL**

**MACARENA PAZ SANTANA SEPÚLVEDA**

Tesis presentada a la Comisión integrada por los profesores:

**MIGUEL NUSSBAUM VOEHL**

**NICOLÁS MAJLUF SAPAG**

**RICARDO DANIEL PAREDES MOLINA**

**MARIO URIBE BRICEÑO**

Para completar las exigencias del grado de  
Magister en Ciencias de la Ingeniería

Santiago de Chile, Enero, 2015

*Dedicatoria*

*A quienes no dejan de creer en que lo pueden lograr y por todos aquellos que creerán.*

## **AGRADECIMIENTOS**

Un gran agradecimiento a todos los que formaron parte de este proceso, en especial a mi familia y amigos que siempre han estado conmigo, a los incondicionales. Quiero destacar la participación de Miguel Nussbaum quien me ha brindado infinitas oportunidades y en especial la de trabajar con Magdalena Claro y Tania Cabello, a quienes admiro profundamente y sin cuya colaboración este trabajo no hubiera sido posible. Muchas gracias al profesor Luis Maldonado por su energía y pasión para enseñar y guiar los análisis estadísticos, y al profesor Guillermo Zamora cuya clase sin duda marcó un antes y un después en mi concepción de la educación.

## INDICE GENERAL

<i>Dedicatoria</i> .....	ii
AGRADECIMIENTOS .....	iii
INDICE GENERAL .....	iv
INDICE DE FIGURAS.....	vii
INDICE DE TABLAS .....	viii
RESUMEN.....	ix
ABSTRACT .....	x
1 INTRODUCCIÓN .....	1
1.1 Motivación.....	1
1.1.1 Bajos desempeños a través del tiempo.....	1
1.1.2 Motivaciones personales .....	3
1.2 Características prueba Simce .....	4
1.3 Objetivos.....	7
1.4 Trabajo realizado.....	7
2 METODOLOGÍA .....	9
2.1 Descripción general.....	9
2.1.1 Metodología análisis descriptivo .....	9
2.1.2 Metodología análisis predictivo.....	10
2.2 Universo y diseño muestral .....	11
2.2.1 Muestra .....	14

2.2.2	Representatividad de la muestra .....	16
3	Conformación modelo predictivo y variables.....	17
3.1	Variables independientes.....	20
3.1.1	Descripción .....	20
3.1.2	Análisis multicolinealidad.....	23
3.1.3	Justificación variables de control.....	26
4	RESULTADOS.....	29
4.1	Resultado análisis descriptivo.....	29
4.1.1	Aprobación 60% .....	29
4.1.2	Aprobación según forma.....	30
4.1.3	Puntaje Simce en cuarto y octavo básico .....	31
4.1.4	Criterio según puntaje Simce cuarto y clasificación 8vo según Nussbaum et al.....	32
4.1.5	Persistencia de las trayectorias. Síntesis de resultados.....	33
4.2	Resultado análisis predictivo .....	33
5	CONCLUSIONES .....	40
	BIBLIOGRAFÍA .....	43
	ANEXOS.....	46
	ANEXO A: Resultados por forma análisis descriptivo .....	47
	ANEXO B: Casos perdidos por limpieza de bases .....	48
	ANEXO C: Resultados diagnóstico de colinearidad para formas K, L y M. ....	50
	ANEXO D: Resultados por forma análisis descriptivo.....	52
	ANEXO E: diagnóstico de diferencias entre coeficientes. pruebas f.....	58
	ANEXO F: Alpha de cronbach y cantidad de preguntas por prueba y eje.....	64

ANEXO G: Representatividad.....	66
ANEXO H: Carta de recepción de paper.....	69
ANEXO I: Paper por enviar.....	70
ANEXO J: Cuarto criterio de clasificación (paper) .....	92
ANEXO K: Descripción Subscores (variables independientes modelo) .....	101
ANEXO L: Resultados modelo sin pesos y error estandar tradicional .....	105

## INDICE DE FIGURAS

Figura 4-1 Consistencia entre los coeficientes obtenidos en las regresiones (OLS) (ICs 90% y 95%).....	37
Figura 4-2 Coeficientes de regresión para cada sub-sample (95% y 99% de confianza)	38

## INDICE DE TABLAS

Tabla 3-3-1 Estadísticos descriptivos de las covariables estandarizadas por submuestra	21
Tabla 3-3-2 Diagnóstico de colinealidad entre ejes.....	24
Tabla 3-3-3 Diagnóstico de colinealidad modelo final.....	25
Tabla 3-3-4 Proporción de varianza, <i>condition index</i> y <i>eigenvalues</i> para forma J.....	25
Tabla 3-3-5 Proporción de varianza, <i>condition index</i> y <i>eigenvalues</i> para forma J sin constante.....	26
Tabla 4-1 Resumen de etiquetas de clasificación.....	29
Tabla 4-2 Clasificación de alumnos de acuerdo al primer criterio (Aprobación simple 60%).....	30
Tabla 4-3 Clasificación de alumnos de acuerdo al primer criterio (Aprobación por forma).....	30
Tabla 4-4 Clasificación de alumnos de acuerdo al tercer criterio (puntaje cuarto y octavo según estándares ministeriales).....	31
Tabla 4-5 Clasificación de alumnos de acuerdo al tercer criterio (puntaje cuarto y octavo según estándares ministeriales).....	32
Tabla 4-6 Resultados modelos de regresión (OLS) por submuestra.....	34
Tabla 4-7 Coeficientes regresiones. Intervalos de confianza por submuestra.....	35

## RESUMEN

Actualmente en Chile existe una baja movilidad en las trayectorias académicas de los estudiantes y, en los últimos años, el país ha tenido un bajo rendimiento en la prueba estandarizada Simcede cuarto básico; nace así la pregunta ¿qué porcentaje de nuestros niños se mantiene en un bajo desempeño entre cuarto y octavo básico? Por otra parte, en el contexto de baja implementación curricular, existe poca evidencia acerca del peso que tienen los distintos ejes de contenido que evalúa cada prueba en cuanto a la predicción del desempeño académico; surge la pregunta ¿qué conocimientos previos en los ejes de la asignatura de matemáticas son más importantes para explicar el desempeño futuro en dicha asignatura? Además, existe evidencia acerca del efecto del grupo socioeconómico, lenguaje y género en el desempeño académico, por lo que es relevante preguntar ¿qué factores sociodemográficos y de desempeño académico son más importantes para determinar el desempeño anteriormente mencionado? Se buscó dichas respuestas, mediante un análisis descriptivo y luego a través regresiones lineales (OLS) aplicadas a las cohortes que rinden las pruebas SIMCE 2007 y 2011 de matemática. Se desarrollaron sub-escalas por eje de contenido (Números, Álgebra, Datos y Azar y, Formas y Espacio) a través del método Barlett. Como resultado de este estudio se concluye que al menos un 77% de los alumnos que presentan bajo desempeño en cuarto permanece en esta situación hasta octavo en ambas asignaturas y que el eje curricular Números es el más importante para la predicción de desempeño en la misma asignatura, seguido por Álgebra y lenguaje, y mejorar dos desviaciones estándar en su desempeño podría compensar el efecto de pertenecer a los grupos socioeconómicos más desaventajados. Esta tesis contó con el apoyo del Centro de Investigación en Políticas y Prácticas en Educación, CIE01- CONICYT.

Palabras claves: predicción desempeño académico, trayectorias académicas, ejes curriculares, Simce.

## ABSTRACT

Currently in Chile there is a low performance of schools measured by Simce (4<sup>th</sup> grade), this implies that many of those students are not able to demonstrate minimum levels of achievement of the learning objectives stipulated in the existing curriculum; that is how arises question, what is the reason for those bad results? Besides, there is some evidence of low mobility in the academic trajectories of students, therefore, what percentage of students remain in low performance between 4<sup>th</sup> and 8<sup>th</sup> grade? There is evidence that among the factors predicting students' results there are socioeconomic status (SES), language and gender. However, there is little evidence about which specific content areas measured by Simce best predict students' general performance in the test. This is important because it identifies, among the content areas measured by Simce, those that should be worked more intensively in order to increase students' chances of future success in the test. The study seeks to examine this relationship applying a linear regression (OLS) analysis to students' scores in Simce test of mathematics, administered on 2007 and 2011. Sub-scales for curricular axes (Numbers, Algebra, Geometry and Statistics and Probabilities) are obtained through the Barlett method. As a result of this study, at least 77% of students with low performance remain in this situation and it is suggested that an improvement in the Numbers of two standard deviation, the most important followed by Algebra and language, might offset the effect of belonging to lower socioeconomic groups.

This thesis had the support of the Center for Research on Educational Policy and Practice, Grant CIE01- CONICYT.

Keywords: predicting academic performance, achievement trajectories, curricular axes, Simce

# 1 INTRODUCCIÓN

## 1.1 Motivación

### 1.1.1 Bajos desempeños a través del tiempo

Uno de los últimos informes de resultados Simce mostró que durante el 2013 sólo un 24% de los alumnos lograron el estándar de aprendizaje adecuado en la prueba de matemáticas de cuarto básico (Agencia de la Calidad de la Educación, 2013a), situación que cada vez mejora menos (Chávez, 2006; Mizala et al., 1998). A raíz de lo anterior, inevitablemente surge la pregunta: ¿A qué se deben estos malos desempeños?

Por otra parte, en Chile existe una baja movilidad en las trayectorias académicas de los estudiantes. Los estudios longitudinales muestran que existe una alta correlación entre los puntajes SIMCE obtenidos por un mismo alumno en dos años distintos de rendición (Valenzuela et al., 2012). En consideración de la baja movilidad de las trayectorias académicas escolares, nace la pregunta ¿cuál es el porcentaje de niños que, habiendo fallado durante cuarto básico en lograr las competencias esperadas según estándares ministeriales, alcanzan el nivel avanzado durante octavo básico o, por el contrario, permanecen bajo los estándares? De confirmarse la persistencia de la baja movilidad académica, nos preguntamos ¿cuánto inciden los conocimientos previos en el desempeño futuro de un niño?, y ¿qué conocimientos y características sociodemográficas son los más esenciales y podrían determinar su trayectoria académica?

Es importante destacar que hoy existe amplia evidencia de factores predictivos de resultados SIMCE entre los cuales destacan el capital cultural, rendimiento previo y género, entre otros (Caprara et al., 2011; Jiménez et al., 2000; Valenzuela et al., 2012). Sin embargo, existe poca evidencia acerca del peso que tienen los distintos ejes de contenido que evalúa cada prueba en cuanto a la predicción del desempeño académico de nuestros alumnos. Conocer esto es importante ya que permite identificar los ejes de contenidos más fundamentales de trabajar en el sistema para aumentar las probabilidades de logro académico futuro de los estudiantes.

Lo anterior se vuelve más importante si se considera que en muchos sistemas escolares no se logran impartir los contenidos que se propone el marco curricular formal para cada año académico. Por ejemplo, un estudio realizado en Estados Unidos mostró que no se estaban trabajando los contenidos y habilidades correspondientes a cada nivel (Schmidt, et.al. 2011). En Chile los resultados son similares, en el sentido de que ninguno de los niveles de enseñanza analizados alcanza un 100% de la cobertura curricular (Ministerio de Educación, 2004, 2013), variando de un 41% a un 73% de cobertura en matemáticas dependiendo de la fuente de información consultada (Ministerio de Educación, 2013).

Por otra parte, desde el ámbito de la investigación cognitiva, existen estudios que encuentran asociaciones entre la alfabetización numérica temprana y el rendimiento posterior en matemáticas (Passolunghi & Lanfranchi, 2012; Geary, 2011; Gersten, Jordan, & Flojo, 2005; Jordan, Glutting, & Ramineni, 2010; Träff, 2013). Específicamente, se encuentra que las habilidades tempranas de números

son una precondition para el pensamiento de segundo orden matemático (Chard et al., 2005; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Krajewski & Schneider, 2009; Passolunghi & Lanfranchi, 2012). Sin embargo, la gran mayoría de estudios encontrados en este ámbito se realizan en base a aplicaciones de pequeña y mediana escala, con un número de casos entre 70 y 411 estudiantes en un número bajo de escuelas. Además, todos los estudios encontrados abordan el tema desde una mirada cognitiva, que no permite orientar En este contexto de baja implementación curricular y la necesidad de dejar espacio para desarrollar en los estudiantes las habilidades básicas para pensar de forma autónoma, en este contexto se plantearon otras dos preguntas de investigación en el contexto de la asignatura de matemáticas. Primero, ¿qué factores son más importantes para determinar el desempeño curricular futuro de los estudiantes en matemáticas? Y segundo, ¿qué conocimientos o ejes curriculares previos son más importantes para explicar el desempeño futuro en matemáticas de los estudiantes?

### **1.1.2 Motivaciones personales**

Que el fracaso de muchos de nuestros niños es algo anunciado tempranamente me parece una idea terrible. En muchos casos los estudiantes de clase baja:

“se consideran un producto de lo que son y el presentimiento de su destino oscuro no hace más que reforzar las posibilidades de fracaso, según la lógica de la profecía que contribuye a su propio cumplimiento” (Bourdieu & Passeron, 2009, p. 108).

Estoy fuertemente convencida y muy de acuerdo la idea de que el capital cultural es un piso y no un techo con respecto a los logros que pueden alcanzar nuestros

estudiantes. Sin duda, que el contexto socioeconómico influye, pero estoy segura de que no nos determina. De hecho según Eyzaguirre:

“En el ambiente educativo circula la noción de que la falta de apoyo de los padres, la ausencia de capital cultural de las familias, la baja capacidad intelectual, las carencias de lenguaje, los problemas de aprendizaje, que se perciben como prevalentes en los sectores pobres, impedirían alcanzar altos rendimientos. Si bien la evidencia indica que estos factores son condicionantes, no muestra que sean determinantes” (2004, p. 252).

La escuela de Ingeniería Civil brinda a sus estudiantes el manejo de herramientas computacionales y conocimiento matemático, que permiten trabajar con bases de datos y llevar a cabo un buen análisis estadístico. En consideración de lo anterior, me gustaría estudiar el tema de la movilidad académica y los factores que permiten escaparse del paradigma de fracaso anunciado, de manera de contribuir con evidencia que favorezca las políticas públicas en torno a desarrollo curricular. Además, quisiera aportar al Departamento de Ciencias de la Computación dando luces acerca de los temas a priorizar en términos de contenido al momento de desarrollar soluciones educativas.

## **1.2 Características prueba Simce**

Para realizar esta investigación fue necesario contar con los patrones de respuesta y descriptores de las preguntas de las pruebas Simce correspondientes a las pruebas de cuarto básico en el año 2007 y octavo básico durante el 2011, tanto para la asignatura de lenguaje como matemáticas. Esta evaluación es de tipo

nacional poblacional, es decir, se registra para su rendición a todos los estudiantes de Chile que cursan los niveles evaluados.

Cada asignatura fue evaluada mediante cuatro formas o cuadernillos diferentes, constituidas tanto por preguntas comunes como exclusivas de la forma. Por lo que en lo que compete a este trabajo se considera que cada asignatura es evaluada, en los años y asignaturas de interés, mediante la aplicación de cuatro pruebas diferentes.

Simce surge en 1988 con el propósito de mejorar la calidad y equidad de la educación, mediante la provisión de información relevante para el quehacer de los distintos actores del sistema educativo. En particular, se comunican los logros de aprendizaje de los estudiantes en diferentes áreas de aprendizaje del currículo nacional, y se relacionan con el contexto escolar y social en el que estos aprenden. Con este fin se recoge también información sobre docentes, estudiantes, padres y apoderados a través de cuestionarios. Desde el 2012 el Simce es el sistema de evaluación que utiliza la Agencia de la Calidad de la Educación para evaluar los resultados de aprendizaje de los establecimientos, a través de la medición del logro de los contenidos y habilidades del programa de estudios vigente en diferentes áreas de aprendizaje o asignaturas (Agencia de la Calidad de la Educación, 2013).

Es importante destacar que la prueba nacional SIMCE no ha sido construida con la finalidad de evaluar ni reportar estos cuatro ejes de forma separada, sino que busca entregar información integrada de los Estándares de Aprendizaje logrados por los estudiantes en los distintos niveles de enseñanza (Agencia de la Calidad de la

Educación, 2013) y representar a partir de un solo puntaje por alumno una habilidad matemática general. Conforme a la Teoría de Respuesta al Ítem (TRI), cada prueba se construye para ser unidimensional, esto para que, de acuerdo con la psicometría, se mida un solo constructo o rasgo latente. Además, para que se pueda elaborar un puntaje de acuerdo a la probabilidad de que el alumno acierte correctamente al ítem en función de la habilidad que se pretende medir, se exige que las preguntas se puedan contestar de manera independiente (Ministerio de Educación, 2008, Agencia de la Calidad de la Educación, 2013).

En consideración de que los descriptores de los objetivos de evaluación por pregunta de la prueba de lenguaje se identifican más con habilidades que contenidos (Hidden Curriculum, 2014), se decide conservar el enfoque psicométrico y no estratificar la prueba en más habilidades que el constructo latente, no así en matemáticas. Luego, dado que la segunda parte de este estudio tiene un enfoque curricular que busca entender, a partir de los contenidos evaluados en los diferentes ítems de la prueba, la importancia relativa de cada uno de los ejes curriculares, se decide realizar el análisis predictivo en la disciplina cuyos objetivos de evaluación se identifiquen más con ejes de contenidos que con habilidades. Mientras que los análisis descriptivos se realizarán en ambas asignaturas.

Los ejes definidos por el Ministerio de Educación para ser evaluados en la prueba Simce de cuarto 2007 en la asignatura de matemáticas fueron: Números, Operaciones aritméticas, Álgebra, y Formas y Espacio (2007) (para más

información ver **Anexo A**). Sin embargo, posteriormente, se incluyeron los ejes de Geometría y Datos y Azar (2010). Por otra parte, los ejes evaluados en octavo 2011 fueron: Números, Algebra, Formas y Espacio y, Datos y Azar (2011). Por lo que los ítems evaluados durante el año 2007 y correspondientes a este eje se re clasificaron según fuera pertinente de acuerdo al lenguaje más reciente relativo al currículo. Por otra parte, en la prueba de lenguaje cuarto 2007 las habilidades evaluadas fueron Localizar, Interpretar y Relacionar, y Reflexionar (2007), mientras que en octavo 2011 se cambia Localizar por Extraer información (2011).

### **1.3 Objetivos**

El principal objetivo de esta tesis es la aplicación de un adecuado modelo predictivo que permita la comparación entre los distintos ejes y/o habilidades medidas en una misma prueba en cuarto básico en relación a su efecto en la predicción del desempeño Simce en octavo básico. En segundo lugar, se busca realizar un análisis descriptivo acerca de la movilidad de las trayectorias académicas entre los mismos años.

### **1.4 Trabajo realizado**

Se realiza la limpieza y procesamiento de las bases de datos. Se desarrollan y ejecutan cuatro métodos de clasificación por tipo de trayectoria para el análisis descriptivo de las mismas tanto en lenguaje como matemática, en particular, el cuarto método se desarrolla utilizando los criterios de clasificación de pertenencia a octavo básico desarrollados por Nussbaum et al. (2015) y dichos resultados

constituyen el aporte realizado al mismo estudio. Luego, se escoge un método adecuado(ya existente) para la generación de subcores por eje curricular en matemática, se realiza un diagnóstico de colinealidad de las variables y ejecutan cuatro regresiones lineales (OLS) correspondientes a las cuatro submuestras disponibles. Finalmente, se interpretan y reportan los resultados.

## 2 METODOLOGÍA

### 2.1 Descripción general

#### 2.1.1 Metodología análisis descriptivo

- 1) Para la realización del análisis descriptivo, se clasificará a los alumnos tanto en cuarto como en octavo de acuerdo a si cumple o no con el nivel avanzado según dos criterios. En ambos años: verificar si aprueba con un 60% la prueba de acuerdo a una suma simple de sus respuestas,
- 2) En ambos años: verificar si aprueba con un porcentaje de éxito según forma que utiliza como referencia los estándares de aprendizaje ministeriales<sup>1</sup> y corresponde con la proporción desarrollada en Frolich(2014) expresada en la siguiente fórmula:

$$P_{\text{éxito}} = \frac{NA_{\text{adecuado}} - pje_{\text{min}}}{pje_{\text{max}} - pje_{\text{min}}}$$

Donde  $P_{\text{éxito}}$  corresponde al porcentaje de éxito,  $NA_{\text{adecuado}}$  al puntaje Simce adecuado según estándares de aprendizaje ministeriales<sup>2</sup>,  $pje_{\text{max}}$  y  $pje_{\text{min}}$  que corresponden a los puntajes máximos y mínimos obtenidos por cada cohorte.

- 3) En ambos años: verificar si el puntaje Simce obtenido por el niño (que es equivalente entre formas) alcanza el nivel adecuado según los estándares de aprendizaje ya mencionados.

---

<sup>1</sup> En este estudio se utilizan también los Estándares de Aprendizaje Matemática 4° básico - Decreto 129\_2013 además de los Estándares de Aprendizaje Matemática 8° básico - Decreto 129\_2013.

<sup>2</sup> Para ver un resumen de los puntajes correspondientes a los estándares, ver **Anexo D**.

4) En cuarto básico: verificar si cumple con el nivel avanzado según estándares de aprendizaje del nivel. En octavo básico: verificar si domina los contenidos de a octavo básico de acuerdo con la clasificación realizada en Nussbaum et al (2015), (ver **Anexo J**). Posteriormente, a partir de las clasificaciones anteriores, se calcula el porcentaje de alumnos que: 1) se mantienen en un nivel adecuado en ambos años (A1), 2) transitan de un nivel avanzado a uno intermedio o insuficiente (A2), 3) los que logran alcanzar el nivel avanzado aun cuando previamente pertenecían al nivel bajo (A3) y, finalmente, aquellos que permanecen en ambos años bajo el mínimo dominio esperado para un alumno de su nivel (A4). Las etiquetas de clasificación de trayectorias académicas recién definidas y expuestas varían según asignatura donde M1, M2, M3 y M4 corresponden a matemáticas y L1, L2,L3,L4 a lenguaje.

, para ver en detalle el resultado por par de formas **Da partir** .

### **2.1.2 Metodología análisis predictivo**

Por otra parte, se propone encontrar la relación entre los distintos ejes curriculares y el posterior desempeño académico Simce mediante la aplicación de regresiones lineales (Franzese et al., 2009; Gelman et al., 2006) aplicadas sobre la cohorte de alumnos que rinden las pruebas SIMCE de matemáticas cuando cursan 4to básico, el año 2007 y también en 8vo básico el año 2011. En cuanto a las variables independientes, se utiliza el método Barlett (DiStefano et al., 2009) que permite el cálculo de subscores por ejes de contenido (Números, Algebra, FyA y DyA) lo que

se realiza independientemente para cada forma de la prueba aplicada durante el 2007 sobre un total de 212.018 alumnos. Para información acerca de la cantidad de preguntas utilizadas por eje ver **Anexo F**. Además, se establecen como variables de control el puntaje Simce obtenido el 2007 en lenguaje, género y grupo socioeconómico. Luego de aplicado el modelo se revisa su ajuste y se comparan los coeficientes obtenidos. Es importante destacar que para cada eje se construye un subpuntaje a partir de la submuestra que rindió cada forma, por lo que se generó un total de dieciséis variables a partir de cuatro ejes de contenido y cuatro formas. Para que estos puntajes fueran comparables y se pudiera interpretar de mejor manera el efecto del grupo socioeconómico y género, se dividieron todas las variables independientes por dos desviaciones estándar exceptuando las variables categóricas (Galmer, 2008).

## **2.2 Universo y diseño muestral**

La primera fase de la investigación estuvo constituida por la obtención de las bases de datos a través de la institución responsable, lo cual fue un proceso largo y que requirió mucha paciencia. Luego de más de un año de espera fue posible obtener los archivos solicitados. Sin embargo, la información recibida poseía una serie de inconsistencias, las que solo lograron ser resueltas tras nuevas solicitudes y consultas. Finalmente, fue posible disponer de la información para los años 2005, 2007, 2009 y 2011, en formato .xlsx, correspondiente a un conjunto de datos con la información tanto de la prueba de lenguaje como de matemáticas para cada año.

Las bases de interés para esta investigación son las cuarto 2007 y octavo 2011 con un (n) total de 492.231 y 442.685 alumnos, respectivamente. Estos archivos constaban de las variables: mrun, patrones de respuestas (alternativas correspondientes a las respuestas de los alumnos) y asignatura. Además se recibe un archivo con las descripciones de los ejes y/o objetivos evaluados en cada pregunta según forma, año y asignatura.

La variable mrun corresponde a la máscara de identificación de alumno que permite cuidar la identidad del mismo, en otras palabras, cumple la función de un R.U.T. virtual. Por otra parte los patrones de respuestas podían poseer letras válidas de alternativa, estar en blanco en caso de estar omitidas, o un guión (-) cuando no aplicaban para ser respondidas con una letra. Finalmente, una misma base estaba constituida tanto por casos de matemáticas como de lenguaje, por lo que para cada año se incluye la variable asignatura. Se dispone además con anterioridad de la información del alumno, de su establecimiento, id del alumno (número de uso interno, diferente al mrun) y forma, almacenadas en cuatro conjuntos de datos diferentes. Se solicita además los datos del cuestionario de padres correspondiente a cuarto 2007.

Con disponibilidad de la información ya descrita, se procede a la revisión, limpieza e integración de las bases de datos. Con este fin se utilizó el software IBM SPSS Statistics 20 y, posteriormente, StataSE 12.

Se revisan los archivos y son abiertos con software Excel para realizarse una inspección visual. En primera instancia se importa se importa al software SPSS y se realiza la limpieza de datos a través del mismo, se exige que contenga una forma válida y se realizan las correcciones sin mayor prolijidad. Sin embargo, este método resultó en obtener casos extraños, existiendo alumnos que poseían patrones de respuesta que no coincidían con su forma y respuestas no válidas (asteriscos en vez de letra) consideradas como legítimas. Por otra parte, el código utilizado por el programa es poco limpio y difícil de almacenar, por lo que no se consiguió un buen registro del trabajo realizado. Como resultado de esta labor, se reinicia el proceso de limpieza y procesamiento de las bases. Para esto se importan nuevamente los archivos a SPSS, sin embargo esto fue solo con la finalidad de cambiar su formato a .dta y poder trabajar los datos en Stata.

Se realiza la división de los archivos según asignatura y se chequea que los patrones de respuesta sean coherentes con las descripciones recibidas, es decir, que aquellas respuestas que tuvieran alternativa efectivamente coincidieran con aquellas preguntas que debían tener. De esta manera se detecta el primer error en los archivos recibidos. El conjunto de datos correspondiente a la prueba de cuarto 2005 no era correcto y se debe realizar nuevamente la solicitud del mismo.

Los conjuntos de datos para los años de interés por asignatura se constituyen de la siguiente manera: 246.259 casos para matemáticas 2007, 245.966 para lenguaje

en el mismo año, 221.198 casos para matemáticas 2011 y 221.375 para lenguaje en el mismo año.

### 2.2.1 Muestra

Para la conformación de la base de datos final fue necesario eliminar aquellos mrun no válidos o cuyos valores estuvieran perdidos. Frente a la existencia de duplicidad de mrun se procedió a preservar el primero de los casos. Luego se eliminan todos aquellos casos que poseyeran valores perdidos en forma de asterisco (\*) en sus patrones. No obstante lo anterior, se mantuvieron los alumnos que poseían respuestas omitidas (casillas en blanco o con etiqueta NULL), en consideración de que estas podrían haber dejado en blanco de manera intencional. Luego se fusiona cada conjunto de datos actual con aquel que contiene idforma e idalumno. Es importante mencionar que, por lo general, cuando se fusionan conjuntos de datos existirán casos descartados. Se eliminan también aquellos casos que presentaban guiones en preguntas que, de acuerdo con su forma, esperaban como respuesta una alternativa. Para ver el detalle de valores perdidos por asignatura y año ver **Anexo B**. Con la finalidad de maximizar la cantidad de casos considerados en cada análisis, estas bases ( $n = 225.638$  y  $n = 225.493$ , en matemáticas y lenguaje respectivamente) es utilizada para el análisis descriptivo de acuerdo con el requerimiento del 60% de respuestas correctas, dificultad por forma y para el cálculo de subpuntajes por eje de contenido en matemáticas.

Posteriormente, se conforma una nueva base de datos al unir la información tanto de alumno (puntaje lenguaje cuarto, puntaje matemática 8vo y género) como de establecimiento (grupo socioeconómico). Se realiza una fusión a través de idalumno y RBD (utilizado para identificar el establecimiento), para formar una base con la información de alumno (puntaje y género, entre otros) e información de establecimiento (grupo socioeconómico y tipo de dependencia, entre otros). Para confirmar la validez de la información socioeconómica y de género, se exige que el grupo no varíe en más de un nivel y que el género registrado se mantenga entre las dos mediciones. Se eliminan todos aquellos alumnos que no contaban con puntaje de lenguaje en cuarto ni puntaje en matemáticas en octavo.

De esta manera, se obtiene una muestra efectiva de 162.676 casos en matemáticas, constituida por cuatro submuestras ( $n_1$ : 42.789,  $n_2$ : 41.342,  $n_3$ : 40.099,  $n_4$ : 38.446) pertenecientes a 5.827 colegios sobre un universo de 226.785 alumnos registrados para rendir la prueba (Agencia de la Calidad de la Educación, 2014). Estos son los casos con información válida (tienen puntajes asociados a las pruebas utilizadas en este estudio en ambos años de rendición). Dada esta pérdida de casos y para asegurar la representatividad de la muestra en cuanto a región y tipo de dependencia administrativa se calcularon pesos muestrales según la metodología utilizada para una prueba nacional de carácter muestral llamada Simce TIC (Enlaces, xx) que se detallará en la siguiente sección.

Para el caso de lenguaje se obtiene una muestra efectiva de 220.200 alumnos. El análisis de representatividad no se realiza en términos de esta asignatura, ya que, como se explicó en la sección 1.2., sólo se lleva a cabo un análisis descriptivo.

### **2.2.2 Representatividad de la muestra**

En consideración de la pérdida de casos, para revisar la representatividad de la muestra con el universo en términos de puntajes SIMCE, tipo de dependencia, región y género, se realizan pruebas de diferencias de medias y de diferencias de proporciones (Enlaces, 2013) mediante el comando `ttest` mediante software Stata 12. Estas pruebas indican que la muestra no es representativa a nivel nacional en términos de ninguna de las variables presentadas, para más detalle ver **Anexo G**. En consecuencia, se requiere el uso de ponderadores o pesos muestrales para que la muestra se pueda considerar representativa del universo. Para asegurar representatividad nacional de la muestra en cuanto a cobertura país y type of school (public, subsidized or private), se calcularon pesos muestrales utilizando la misma metodología que la usada en una prueba nacional de carácter muestral para evaluar habilidades ICT, SIMCE TIC (Enlaces, 2011). Este cálculo se realiza en el software estadístico SAS mediante un código desarrollado por el profesor Ernesto San Martín (2011) que fue adaptado a la muestra y universo de este estudio.

### 3 CONFORMACIÓN MODELO PREDICTIVO Y VARIABLES

Para responder al objetivo se ajustan modelos de regresiones lineales (OLS) con *cluster-robust standard errors using sampling weights*, donde la variable dependiente corresponde al puntaje SIMCE Matemáticas obtenido en 8vo básico (año 2011). Este es un puntaje con rango entre 135,35 y 395,66 y una media 264,78 calculados el total de la muestra. Esta prueba evalúa Números, Algebra, Geometría y, Datos y azar.

La prueba SIMCE de matemáticas para cuarto básico durante el año 2007 se aplicó en cuatro formas diferentes (J, K, L y M) con un conjunto de ítems comunes a partir de los cuales se realiza un procedimiento de *equating* para obtener un sólo puntaje por estudiante comparable entre formas a nivel de escuela. Sin embargo, para este estudio se generaron scores por subconjuntos de ítems de la prueba (para cada eje curricular), por lo que se debe ajustar un modelo de O'brien regresión lineal para cada submuestra que rindió cada una de las cuatro formas del test.

Los modelos de regresión consideran los pesos muestrales calculados y el cálculo del error estándar se realiza por colegio (clúster). Este cálculo de los errores estándar es más conservador, ya que su cálculo convencional se vuelve más engañoso en la medida que aumenta el n de la muestra, y considera la agrupación de estudiantes por colegio, en términos de que relaja la suposición de que los errores estándar son independientes entre sí y permite la correlación arbitraria

entre los errores dentro de una escuela (Angrist & Pischke, 2008). Todos los análisis fueron realizados utilizando Stata 12.

Finalmente, mediante el software Stata/SE 12.0 se realizan cuatro regresiones lineales, una por cada muestra de estudiantes según la forma de prueba rendida en cuarto básico, mediante comando regress y pesos pweight (Hamilton, 2012) correspondientes al inverso de la probabilidad de una observación.

Los modelos de regresión consideran los pesos muestrales calculados y el cálculo del error estándar se realiza por colegio (clúster). Este cálculo de los errores estándar es más conservador, ya que su cálculo convencional se vuelve más engañoso en la medida que aumenta el n de la muestra, y considera la agrupación de estudiantes por colegio, en términos de que relaja la suposición de que los errores estándar son independientes entre sí y permite la correlación arbitraria entre los errores dentro de una escuela (Angrist & Pischke, 2008). Todos los análisis fueron realizados utilizando Stata 12.

El clásico modelo de regresión lineal puede escribirse matemáticamente como:

Ecuación 1: Modelo regresión lineal clásico

$$y_i = \beta_1 X_{i1} + \dots + \beta_k X_{ik} + \epsilon_i, \text{ for } i = 1, \dots, n,$$

## Ecuación 2: Modelo regresión lineal actual

$$y_i = \beta_{Num}X_{Num} + \beta_{Alg}X_{Alg} + \beta_{FyE}X_{FyE} + \beta_{DyA}X_{DyA} + \vec{\beta}_c\vec{X}_c + \epsilon_i$$

$$\vec{\beta}_c = \begin{bmatrix} \beta_{GSE A} \\ \beta_{GSE B} \\ \beta_{GSE C} \\ \beta_{GSE D} \\ \beta_{GSE E} \\ \beta_{Female} \\ \beta_{Language} \end{bmatrix} \quad \vec{X}_c = \begin{bmatrix} X_{GSE A} \\ X_{GSE B} \\ X_{GSE C} \\ X_{GSE D} \\ X_{GSE E} \\ X_{Female} \\ X_{Language} \end{bmatrix}$$

Donde  $\beta_{Num}$  corresponde al coeficiente de la regresión relacionado con el puntaje de Números ( $X_{Num}$ ),  $\beta_{Alg}$  al coeficiente relacionado con el puntaje en Algebra ( $X_{Alg}$ ),  $\beta_{FyE}$  a Formas y Espacio ( $X_{FyE}$ ),  $\beta_{DyA}$  a Datos y Azar ( $X_{DyA}$ ) y  $\vec{\beta}_c$  es el vector de coeficientes correspondientes a las variables de control  $\vec{X}_c$  (grupo socioeconómico, género y puntaje lenguaje),  $\epsilon_i$  tiene distribución normal independiente con media 0 y desviación estandar  $\sigma$ .

Posteriormente, para descartar problemas de multicolinealidad perjudicial o excesiva se verifica que el indicador VIF sea menor que 10 (Mandel, 1985; Marquardt, 1970 citado en Mason & Perreault 1991), incluso se ha sugerido que sea menor que 4 (O'brien, 2007). Por otra parte, se revisa que el *condition number* no sea mayor de 30, ya que de acuerdo con Belsley, Kuh, y Welsh (1980, citado en Mason & Perreault 1991), en tales casos se evidenciarían fuertes dependencias entre variables. Estos procedimientos se realizan primero entre las cuatro variables correspondientes a los ejes de contenido y, finalmente, en el modelo completo para cada uno de los cuatro cuadernillos.

Finalmente, se realiza de una serie de pruebas Wald que permite comparar si el efecto de dos variables es el mismo dentro de la regresión (Williams, 2014) a través del comando post regresión test mediante software Stata.

### **3.1 Variables independientes**

#### **3.1.1 Descripción**

Cabe destacar que los puntajes de los cuatro ejes curriculares de Matemáticas evaluados el año 2007 se componen de diferentes cantidades de ítems, que además varían por forma de la prueba. Las variables independientes incluidas en el modelo se presentan a continuación y se detallan en la **Tabla 3-1** presentada al final de esta sección. Para más información descriptiva acerca de las variables no estandarizadas y el output desde el programa ver **Anexo H**.

a) **Números:** Corresponde al puntaje obtenido en las preguntas sobre el eje curricular de Números en el SIMCE Matemáticas en 4to básico aplicado el año 2007 (detalle en **Anexo A**). Para obtener un puntaje sobre el desempeño de los estudiantes en este eje, se realizó un análisis factorial exploratorio a partir de correlaciones tetracóricas de todos los ítems binarios (entre 15 y 17 por forma) donde se estimó un puntaje a través del método Barlett (DiStefano et al., 2009) del primer componente (eigenvalue del primer factor [5,84 a 6,81]; Cronbach's alpha= [0,81 a 0,84]). Finalmente, para luego poder comparar los resultados, los puntajes se estandarizaron dividiendo la variable generada en cada forma por dos de sus desviaciones estándar ([0,324 a 0,338]). Este método se llevó a cabo

independientemente para cada una de las cuatro sub muestras (según forma rendida) de la prueba aplicada durante el 2007.

Tabla 3-3-1 Estadísticos descriptivos de las covariables estandarizadas por submuestra

Variable	Forma J (n=42.789)				Forma K (n=41.342)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Números	1,49	0,50	0,00	2,14	1,63	0,50	0,00	2,26
Álgebra	1,26	0,50	0,00	1,86	1,26	0,50	0,00	1,83
FyE	1,68	0,50	0,00	2,29	1,84	0,50	0,00	2,43
DyE	1,56	0,50	0,00	1,96	1,75	0,50	0,00	2,12
Lenguaje	2,72	0,50	1,14	3,78	2,64	0,50	1,19	3,76
GSE	3,10	1,09	1,00	5,00	3,09	1,07	1,00	5,00
Género	0,52	0,50	0,00	1,00	0,52	0,50	0,00	1,00

Variable	Forma L (n=40.099)				Forma M (n=38.446)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Números	1,47	0,50	0,00	2,11	1,55	0,50	0,00	2,14
Álgebra	1,19	0,50	0,00	1,73	1,33	0,50	0,00	1,83
FyE	1,84	0,50	0,00	2,42	1,90	0,50	0,00	2,48
DyE	1,43	0,50	0,00	1,93	1,38	0,50	0,00	1,89
Lenguaje	2,59	0,50	1,17	3,72	2,61	0,50	1,21	3,73
GSE	3,07	1,06	1,00	5,00	3,06	1,05	1,00	5,00
Género	0,52	0,50	0,00	1,00	0,52	0,50	0,00	1,00

b) **Álgebra:** Corresponde al puntaje obtenido en las preguntas sobre el eje curricular de Álgebra en el SIMCE Matemáticas en 4to básico aplicado el año 2007 (detalle en **Anexo A**). Se llevó a cabo la misma metodología que en a) para generar un puntaje a partir de los ítems (entre 4 y 5 por forma) (eigenvalue primer factor = [1,5 a 1,90]; Cronbach's alpha= [0,53 a 0,61]). Finalmente, se estandariza dividiendo por dos desviaciones estándar ([0,301 a 0,324]) en cada forma.

- c) Formas y espacio: Corresponde al puntaje obtenido en las preguntas sobre el eje curricular de Geometría en el SIMCE Matemáticas en 4to básico aplicado el año 2007 (detalle en **Anexo A**). Se llevó a cabo la misma metodología que en a) para generar un puntaje a partir de los ítems (entre 9 y 11 por forma), (eigenvalue primer factor = [2,84 a 3,59]; Cronbach's alpha= [0,65 a 0,71]). Finalmente, se estandariza dividiendo por dos desviaciones estándar ([0,287 a 0,301]) en cada forma.
- d) Datos y azar: Corresponde al puntaje obtenido en las preguntas sobre el eje Datos y Azar en el SIMCE Matemáticas en 4to básico aplicado el año 2007 (detalle en Anexo 1). Se llevó a cabo la misma metodología que en a) para generar un puntaje a partir de los ítems (entre 4 y 6 por forma), (eigenvalue primer factor = [2,20 a 2,56]; Cronbach's alpha= [0,45 a 0,67]). Finalmente, se estandariza dividiendo por dos desviaciones estándar ([0,301, 0,324]) en cada forma.
- e) Puntaje SIMCE Lenguaje 4to básico (2007): Corresponde al puntaje obtenido en la prueba SIMCE Lenguaje en 4to básico aplicado el año 2007. Esta prueba evalúa comprensión lectora a través de tareas de localización, interpretación relación y reflexión de información en textos literarios y no literarios (Ministerio de Educación, 2007). El rango de puntajes está entre 112,32 y 379,35 con media 265,23 (calculados sobre el total de la muestra). Este puntaje se estandariza a partir de la división por dos desviaciones estándar de la variable (50,23).

f) Género: autodeclaración obtenida a partir de los cuestionarios de contexto SIMCE. Corresponde a una variable binaria donde 1 representa femenino y 0 a masculino (agregar % de hombres y mujeres en la muestra efectiva).

g) Grupo socioeconómico: corresponde a la clasificación por colegio realizada por el Ministerio de Educación y sus valores son nominales (A, B, C, D y E) donde A corresponde al grupo más bajo y E corresponde al más alto. Esta variable se construye en base al nivel educacional de la madre, nivel educacional del padre, ingreso mensual del hogar e índice de vulnerabilidad de la escuela (Ministerio de Educación, 2012).

Vale la pena mencionar que, en general, los resultados estimados tienen aceptable a buena consistencia interna con alfas que varían de 0,62 a 0,84. Sin embargo, en cuatro casos no se alcanzan estos niveles de consistencia interna, lo que significa que los resultados relativos Álgebra en submuestras J (alfa = 0,53), L (alfa = 0,55) y M (alfa = 0,55) y en probabilidad y estadística en J submuestra (alfa = 0,45) debe ser interpretado con cautela.

### **3.1.2 Análisis multicolinealidad**

En consideración de que la unidimensionalidad<sup>3</sup> de la prueba (presencia de un solo constructo latente) constituye un diagnóstico alternativo de colinealidad

---

<sup>3</sup> Las características de la prueba Simce fueron descritas en la sección 1.2.

(Kleinbaum et al., 2013), entre las preguntasse realiza un diagnóstico de multicolinealidad entre las cuatro variables generadas a partir de las mismas correspondientes a los subpuntajes por ejes de contenidos evaluados. En la **Tabla 3-2** se puede observar que los indicadores del VIF son menores a 3 tanto en cada una de las variables como en el promedio. Por otra parte, en las cuatro formas se presentan *condition numbers* menores a 15. En consideración de lo anterior se descartan problemas graves de colinealidad entre ellas.

Análogamente se descartan problemas graves de multicolinealidad en el modelo general. En la **Tabla 3-3** se observa también que en todas las formas el indicador VIF es menor que 3 y que los *condition number* son menores a 30.

Tabla 3-3-2 Diagnóstico de colinealidad entre ejes

	VIF			
	Forma J	Forma K	Forma L	Forma M
Numeros*	2.62	2.62	2.79	2.67
Algebra*	1.80	1.78	1.81	1.70
FyE*	1.93	2.05	2.06	2.06
DyA*	1.90	1.76	2.08	2.11
Mean VIF	2.06	2.05	2.19	2.13
Condition number	13.18	14.56	14.13	14.58
Det (correlation matrix)	0.167	0.171	0.143	0.152

\*Las variables se construyen a partir de distintas submuestras

Tabla 3-3-3 Diagnóstico de colinealidad modelo final

	VIF			
	Forma J	Forma K	Forma L	Forma M
Numeros*	2.93	2.92	3.21	2.90
Algebra*	1.88	1.87	1.88	1.77
FyE*	2.02	2.15	2.17	2.15
DyA*	1.98	1.84	2.17	2.26
Lenguaje	2.29	2.23	2.29	2.29
GSE	1.04	1.05	1.05	1.05
Género	1.16	1.18	1.19	1.17
Mean VIF	1.90	1.89	2.00	1.94
Condition number	27.55	26.14	26.75	26.74
Det (correlation matrix)	0.0626	0.0643	0.0510	0.0560

\*Las variables se construyen a partir de distintas submuestras

Para comprender mejor el aumento en el *condition number* se realiza un análisis mediante los indicadores de proporción de varianza y *condition index* (Kleinbaum et al., 2013). En la **Tabla 3-4** y **Tabla 3-5** se presentan los resultados sólo para la forma J, ya que se observó consistencia general a través de las formas. En la primera tabla se puede identificar que para el *condition index* 8 (27.55) una alta proporción de la varianza es explicada por el intercepto (.74). Luego, en consideración de dicha observación, se realiza el análisis ajustado sin constante, correspondiente a la segunda tabla, a partir del cual se ve que en el *condition index* 7 (18.34) es altamente atribuible a lenguaje (.98). Para ver el detalle de las demás formas ver **Anexo C**.

Tabla 3-3-4 Proporción de varianza, *condition index* y *eigenvalues* para forma J

Condition	eigenvalue	Index	_cons	Num.	Alg.	FyE	DyA	Leng.	Fem.	GSE
1	7,1590	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00
2	0,4815	3,86	0,00	0,00	0,01	0,00	0,00	0,00	0,84	0,00
3	0,1970	6,03	0,01	0,01	0,03	0,00	0,01	0,00	0,08	0,33
4	0,0598	10,94	0,00	0,00	0,82	0,07	0,13	0,00	0,02	0,04
5	0,0375	13,81	0,02	0,00	0,01	0,36	0,73	0,01	0,00	0,04
6	0,0287	15,79	0,03	0,88	0,10	0,14	0,10	0,00	0,00	0,07
7	0,0271	16,25	0,21	0,00	0,02	0,43	0,02	0,14	0,05	0,35
8	0,0094	27,55	0,74	0,10	0,00	0,00	0,01	0,85	0,00	0,16

Tabla 3-3-5 Proporción de varianza, *condition index* y *eigenvalues* para forma J sin constante

Condition	Index	Num.	Alg.	FyE	DyA	Leng.	Fem.	GSE
1	1	0	0	0	0	0	0,01	0
2	3,59	0	0	0	0	0	0,84	0
3	5,78	0,01	0	0	0	0	0,07	0,7
4	10,19	0,01	0,8	0,1	0,1	0	0,02	0
5	12,96	0,01	0	0,5	0,7	0,01	0	0
6	14,73	0,79	0,1	0,3	0,1	0,01	0	0
7	18,34	0,18	0	0,1	0	0,98	0,06	0,2

### 3.1.3 Justificación variables de control

De acuerdo con la revisión bibliográfica realizada, existe evidencia que muestra que el nivel socioeconómico, el dominio del lenguaje y el género son variables importantes para explicar el desempeño académico en general y matemáticas en particular. Específicamente, en relación al nivel socioeconómico, se ha encontrado

amplia evidencia de su importancia en el rendimiento en general (Bradley & Corwyn, 2002; Buchmann, 2002; Dahl & Lochner, 2005; OECD, 2013a, 2013b; Sirin, 2005; Sullivan, 2007) y específicamente en matemáticas (Jordan, Kaplan, Olah, & Locuniak, 2006; Sirin, 2005). Además, hay evidencia de la asociación entre el nivel socioeconómico inicial en el rendimiento posterior en matemáticas (Krajewski & Schneider, 2009).

En relación a la alfabetización lectora y escrita, se ha encontrado que su conocimiento es fundamental en otros desempeños académicos como las matemáticas y las ciencias (Mullis, Martin, & Foy, 2013; Norris & Phillips, 2003; Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008; Vilenius-Tuohimaa, Aunola, & Nurmi, 2008). Específicamente con respecto al rendimiento en matemáticas, la evidencia muestra que los estudiantes con peor desempeño en lectura demuestran consistentemente un peor desempeño en matemáticas (Mullis et al., 2013) y que la habilidad de comprensión lectora está fuertemente asociada al desempeño para resolver problemas matemáticas escritos (Vilenius-Tuohimaa et al., 2008). Además, con respecto al progreso, se ha encontrado una asociación entre las dificultades en lectura y un progreso lento en superar dificultades en matemáticas (Gersten et al., 2005).

Finalmente, también existe amplia evidencia hay mucha investigación sobre las diferencias de género en el rendimiento en matemáticas, aunque estas varían entre países. . En la prueba PISA 2012, en los países de la OECD los hombres obtienen

en promedio 11 puntos más que las mujeres, sin embargo, solo en 37 de los 65 países los hombres tienen un mejor desempeño que las mujeres (OECD, 2013b). En la misma línea, (Else-Quest, Hyde, & Linn, 2010) analizan las diferencias de género en las pruebas TIMSS y PISA del año 2003 y encuentran que el tamaño promedio del efecto de género es muy bajo, aunque este varía bastante entre países. Por otro lado, se ha encontrado que desde edades tempranas, los hombres tienden a tener mejor desempeño que las mujeres y que estas diferencias pueden ser incluso mayores en los grupos de padres con alta educación (Penner & Paret, 2008). En Chile, los resultados de TIMSS 2003 y 2011 indican que, en promedio, los hombres tienen significativamente un mejor resultado que las mujeres (Mullis, Martin, Foy, & Arora, 2012).

## 4 RESULTADOS

### 4.1 Resultado análisis descriptivo

Tal como se mencionó en la introducción y metodología, se busca describir las trayectorias académicas de los estudiantes. Para esto se construyen y definen cuatro clases de trayectorias según asignatura las que se resumen en la siguiente tabla (4-1) para facilitar la lectura:

Tabla 4-1 Resumen de etiquetas de clasificación

M1/L1	4th APRUEBA-8th APRUEBA
M2/L2	4th APRUEBA-8th FALLA
M3/L3	4th FALLA-8th APRUEBA
M4/L4	4th FALLA-8th FALLA

#### 4.1.1 Aprobación 60%

A continuación se presentan los resultados a partir del primer criterio de clasificación. A partir de la tabla 4-2 de lenguaje es posible decir que un aproximadamente 77% de los niños que fallan en cumplir los estándares en cuarto básico vuelve a fallar en octavo. Por otra parte, un 75% de los niños que aprueban en cuarto vuelve a aprobar en octavo. Por otra parte en matemáticas, un 97% de los alumnos que fallan en matemática, vuelve a fallar en octavo y un 44% de los que aprueba en cuarto vuelve a hacerlo en octavo

Tabla 4-2 Clasificación de alumnos de acuerdo al primer criterio  
(Aprobación simple 60%)

Matemáticas				
Resultados a nivel país* (%)				
Caso	M1	M2	M3	M4
% total**	29,84%	38,70%	1,09%	30,37%
Lenguaje				
Resultados a nivel país* (%)				
Caso	L1	L2	L3	L4
% total**	37,66%	12,31%	11,44%	38,59%

Para ver el detalle (resultados por combinación de formas) ver **Anexo D**.

#### 4.1.2 Aprobación según forma

A continuación se presentan los resultados a partir del segundo criterio de clasificación:

Tabla 4-3 Clasificación de alumnos de acuerdo al segundo criterio  
(Aprobación por forma)

Matemáticas				
Resultados a nivel país* (%)				
Caso	M1	M2	M3	M4
% total**	25,75%	21,36%	3,69%	49,20%
Lenguaje				
Resultados a nivel país* (%)				
Caso	L1	L2	L3	L4
% total**	38,78%	17,98%	6,99%	36,24%

A partir de la segunda tabla 4-3, del total de niños que falla en cuarto un 84% vuelve a fallar en octavo en la asignatura de lenguaje. Así también un 68% de los niños que aprueban en cuarto vuelve a aprobar en octavo. Por otra parte en matemáticas, un 93% de los alumnos que fallan en matemática, vuelve a fallar en octavo y un 55% de los que aprueba en cuarto vuelve a hacerlo en octavo. Para ver el detalle de este procedimiento ver **Anexo D**.

#### 4.1.3 Puntaje Simce en cuarto y octavo básico

A continuación se presentan los resultados a partir del tercer criterio de clasificación:

Tabla 4-4 Clasificación de alumnos de acuerdo al tercer criterio (puntaje cuarto y octavo según estándares ministeriales)

Matemáticas				
Resultados a nivel país* (%)				
Caso	M1	M2	M3	M4
% total**	16,93%	8,79%	9,04%	65,24%
Lenguaje				
Resultados a nivel país* (%)				
Caso	L1	L2	L3	L4
% total**	21,23%	16,49%	7,38%	54,90%

A partir de la tabla 4-4, del total de niños que falla en cuarto un 88% vuelve a fallar en octavo en la asignatura de lenguaje. Así también un 56% de los niños que

aprueban en cuarto vuelve a aprobar en octavo. Por otra parte en matemáticas, un 88% de los alumnos que fallan inicialmente en esta asignatura, vuelve a fallar en octavo y un 66% de los que aprueba en cuarto vuelve a hacerlo en octavo.

#### 4.1.4 Criterio según puntaje Simce cuarto y clasificación 8vo según Nussbaum et al.

A continuación se presentan los resultados a partir del cuarto criterio de clasificación:

Tabla 4-5 Clasificación de alumnos de acuerdo al tercer criterio (puntaje cuarto y octavo según estándares ministeriales)

Matemáticas				
Resultados a nivel país* (%)				
Caso	M1	M2	M3	M4
% total**	15,7%	9,4%	8,7%	66,3%
Lenguaje				
Resultados a nivel país* (%)				
Caso	L1	L2	L3	L4
% total**	27,7%	17,1%	8,7%	46,5%

A partir de la tabla 4-5, del total de niños que falla en cuarto un 84% vuelve a fallar en octavo en la asignatura de lenguaje. Así también un 62% de los niños que aprueban en cuarto vuelve a aprobar en octavo. Por otra parte en matemáticas, un 88% de los alumnos que fallan en esta asignatura, vuelve a fallar en octavo y un 62% de los que aprueba en cuarto vuelve a hacerlo en octavo.

#### 4.1.5 Persistencia de las trayectorias. Síntesis de resultados

En lenguaje para todas las metodologías de clasificación más de un 77% de los alumnos que falla en cuarto básico vuelve a fallar en octavo. La permanencia en un alto desempeño se da al menos en un 56% de los alumnos que aprueba cuarto según las cuatro clasificaciones. En la asignatura de matemática, en todas las clasificaciones al menos un 88% de los alumnos que falla en cuarto vuelve a fallar en octavo. Por otra parte, un 44% de los alumnos que aprueba vuelve a aprobar.

#### 4.2 Resultado análisis predictivo

En la **Tabla 4-6** se presentan los resultados de los modelos de regresión ajustados para cada submuestra. El ajuste de los modelos es adecuado con un  $R^2$  de 0.52, por lo que se puede decir que los modelos se ajustan bien a los datos.

Al trasladarse a un modelo representativo y aplicar pesos como ponderadores, se pierde en cierto grado la significancia de la variable género a través de las distintas formas y se presentan ciertas variaciones en la significancia de las variables de control. Sin embargo, se observó consistencia a nivel general al aplicar los pesos (ver **Anexo L**).

Los resultados indican que la gran mayoría de las variables incluidas en el modelo (números, álgebra, formas y espacio, puntaje en SIMCE Lenguaje, género y grupo socioeconómico) tienen un efecto estadísticamente significativo (y consistente entre formas) en el rendimiento de los estudiantes en la prueba SIMCE

Matemáticas rendida en 8vo básico (con al menos un  $p < 0.1$ ). Excepcionalmente, los resultados indican que el score obtenido en el eje curricular de Datos y Azar no tiene un efecto estadísticamente significativo en el rendimiento en la prueba SIMCE de Matemáticas 8vo básico en dos de las submuestras (forma J y K).

Tabla 4-6 Resultados modelos de regresión (OLS) por submuestra

VARIABLES	Forma 4to J Puntaje Matemáticas 8vo	Forma 4to K Puntaje Matemáticas 8vo	Forma 4to L Puntaje Matemáticas 8vo	Forma 4to M Puntaje Matemáticas 8vo
Números*	26.55*** (2.286)	26.92*** (1.165)	29.01*** (1.339)	25.09*** (0.708)
Algebra*	16.32*** (0.855)	17.10*** (1.152)	16.14*** (0.565)	15.56*** (0.971)
Formas y Espacio*	9.622*** (0.0769)	10.50*** (0.476)	10.19*** (0.871)	9.086** (1.776)
Datos y Azar*	1.784 (0.914)	1.418 (1.415)	5.410*** (0.235)	4.859*** (0.399)
Puntaje SIMCE Lenguaje 4to básico	15.95*** (1.421)	13.74*** (1.318)	12.00** (2.332)	15.85*** (0.900)
Género (mujer)	-5.100** (0.976)	-5.166** (0.634)	-3.242* (0.948)	-5.083** (0.815)
2.Grupo D (GSE)	-16.34*** (0.977)	-17.29*** (0.337)	-15.30** (1.937)	-18.05** (2.161)
3.Grupo C (GSE)	-27.17** (3.338)	-27.64** (3.715)	-22.60** (3.068)	-25.88** (4.226)
4.Grupo B (GSE)	-30.18*** (2.865)	-31.20*** (2.926)	-25.32** (2.741)	-29.88** (3.663)
5.Grupo A (GSE)	-28.33*** (1.679)	-28.66*** (0.788)	-23.91*** (0.405)	-28.65** (3.435)
Constante	166.5*** (8.176)	166.2*** (5.881)	164.7*** (8.506)	163.8*** (8.177)
Observaciones	42,789	41,342	40,099	38,446
R-cuadrado	0.520	0.518	0.523	0.521

\*Las variables indicadas se construyen a partir de distintas muestras correspondientes a las cuatro formas de la prueba.

Robust standard errors in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$  two tailed test

Tabla 4-6 Coeficientes regresiones. Intervalos de confianza por submuestra

Intervalos de confianza (95%)				
	Forma J	Forma K	Forma L	Forma M
Números*	[ 24,88 , 28,23 ]	[ 25,14 , 28,69 ]	[ 27,25 , 30,77 ]	[ 23,37 , 26,82 ]
Algebra*	[ 14,97 , 17,68 ]	[ 15,59 , 18,61 ]	[ 14,80 , 17,49 ]	[ 14,26 , 16,85 ]
FyE*	[ 8,20 , 11,05 ]	[ 9,08 , 11,91 ]	[ 8,60 , 11,79 ]	[ 7,56 , 10,61 ]
DyE*	[ 0,36 , 3,21 ]	[ 0,03 , 2,80 ]	[ 3,87 , 6,95 ]	[ 3,34 , 6,38 ]
Lenguaje*	[ 14,33 , 17,57 ]	[ 12,16 , 15,32 ]	[ 10,40 , 13,60 ]	[ 14,38 , 17,32 ]
Género*	[ -6,13 , -4,07 ]	[ -6,31 , -4,02 ]	[ -4,39 , -2,09 ]	[ -6,11 , -4,05 ]
GSE				
D	[ -18,98 , -13,69 ]	[ -20,55 , -14,02 ]	[ -19,27 , -11,32 ]	[ -21,46 , -14,64 ]
C	[ -29,73 , -24,61 ]	[ -30,79 , -24,48 ]	[ -26,38 , -18,82 ]	[ -29,14 , -22,61 ]
B	[ -32,85 , -27,51 ]	[ -34,44 , -27,95 ]	[ -29,19 , -21,44 ]	[ -33,17 , -26,58 ]
A	[ -31,30 , -25,37 ]	[ -32,19 , -25,14 ]	[ -28,11 , -19,71 ]	[ -32,20 , -25,10 ]
Constante	[ 161,74 , 171,18 ]	[ 161,66 , 170,83 ]	[ 158,96 , 170,46 ]	[ 159,10 , 168,41 ]

En cuanto al nivel de significancia por forma, tanto el puntaje en Números como en Algebra y el intercepto resultaron ser significativos para todas las formas con un 99% de confianza. FyE resultó ser significativo para tres de las cuatro formas (J, K, L) con un 99% de confianza y con un 95% en la forma M. El puntaje de lenguaje también resulta ser altamente significativo (99%) en tres formas y con un 95% de confianza (forma L). En cuanto a género, en tres de las formas aparece con un 95% de confianza y sólo en una (L) con un 90%. DyA es significativo con un 99% de confianza tanto para las formas L como M. En cuanto a las variables que representan grupo socioeconómico resultan ser todas significativas con al menos un 95% de confianza, alcanzando un 99% de confianza tanto el coeficiente del grupo B como el del grupo D de en dos formas (J y K), mientras que el grupo A lo

alcanza en tres de las cuatro (J, K y L). En síntesis, tal como se indicó en el párrafo anterior, se obtiene que todas las variables de las regresiones son significativas con al menos un 90%, exceptuando Datos y Azar (DyA) en las formas J y K que resultaron ser no significativas.

En cuanto a la significancia por forma, cada uno de los modelos aplicados a las cuatro submuestras está conformados por once coeficientes. Tanto en la forma J como en la K, ocho de estos presentan un 99% de significancia, otros dos un 95% y sólo una variable resulta no ser significativa. Con respecto a la forma L, sólo seis de las variables presentan un 99% de significancia, cuatro un 95% y una un 90%. En cuanto a la forma M, cinco de los coeficientes alcanzan el 99% y seis un 95%.

Aun cuando a diferencia de la regresión logística, en la regresión OLS no existen problemas con la varianza no observadas, los coeficientes estandarizados no se pueden comparar entre grupos en tanto varianza entre los grupos no sea la misma (William, 2009). No obstante lo anterior, en la **Figura 4-1** se puede observar consistencia general entre los cuatro modelos.

Al observar los IC 95%, se puede observar que el efecto de números es estadísticamente diferente al de todos los otros ejes para todas las formas, ver Tabla 3. Sin embargo, no es posible afirmar que exista una diferencia significativa entre el efecto absoluto de números con respecto a los grupos socioeconómicos A, B ni C. Tampoco se puede asegurar que exista una diferencia significativa entre ellos. Por otra parte, es posible ver que Lenguaje y Algebra se diferencian solo en

dos de las formas (K y L). En cuanto a las diferencias entre Lenguaje y, Formas y Espacio se observan diferencias significativas en todos los casos a excepción la forma L. Lo anterior queda más claramente representado en la **Figura 4-2**.

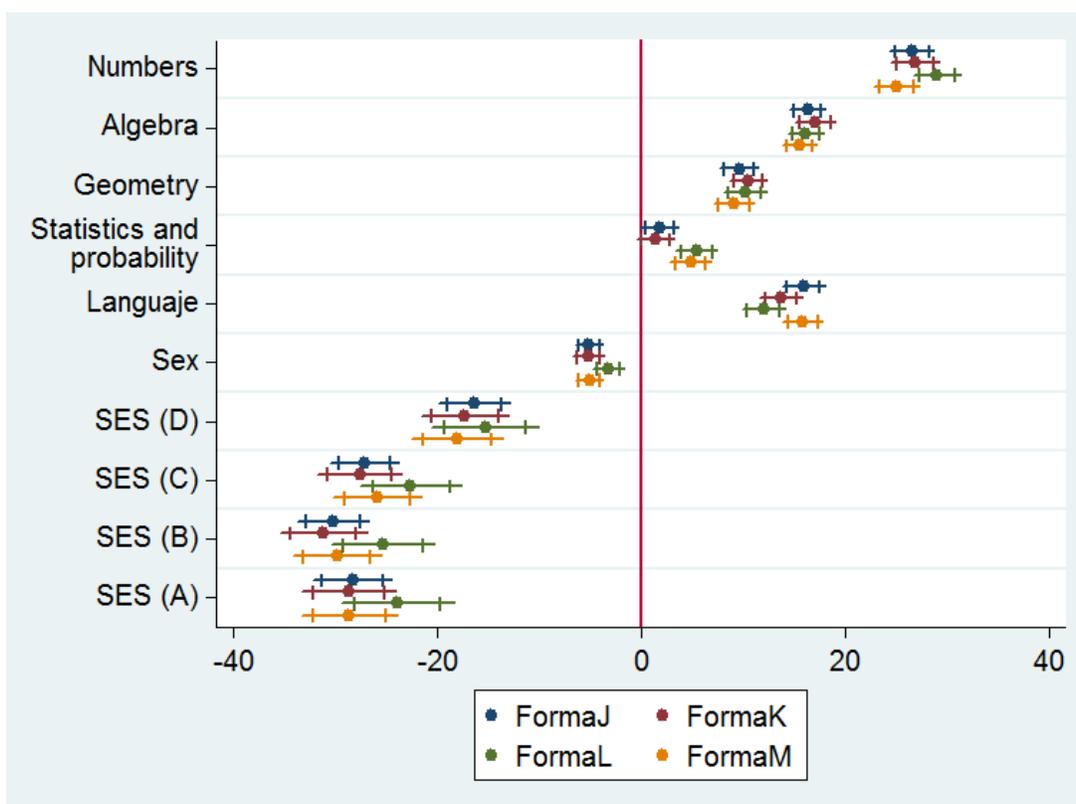


Figura 4-1 Consistencia entre los coeficientes obtenidos en las regresiones (OLS) (ICs 90% y 95%)

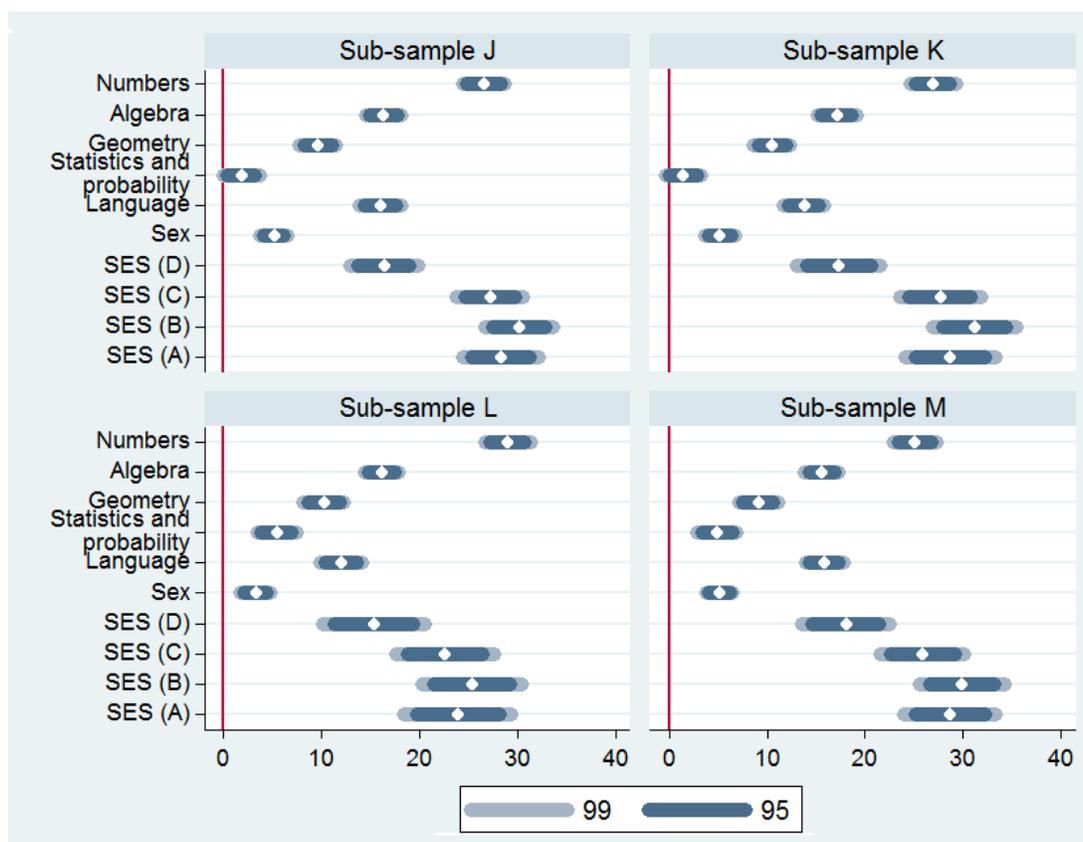


Figura 4-2 Coeficientes\* de regresión para cada sub-sample (95% y 99% de confianza)

\*Todos los coeficientes se presentan en valor absoluto para facilitar la interpretación del gráfico.

Para respaldar esto, al observar la forma J, Números, IC 95% [24.88, 28.23], difiere significativamente de todas las demás variables relacionadas con dominio curricular. Con respecto a Álgebra, IC 95% [14.97, 17.68], se puede asegurar que su efecto es significativamente diferente,  $F(1,6798)=[64.64]$ ,  $p<.001$ , también con respecto a FyE , IC 95% [8.20, 11.05],  $F(1, 6798)=[187.41]$ ,  $p<.001$ ; así también con respecto a DyA, IC 95% [0.36, 3.21],  $F(1, 6798)=[398.93]$ ,  $p<.001$ , igualmente con respecto a lenguaje, IC 95% [14.33, 17.57], se encuentra una

diferencia significativa,  $F(1, 6798)=[60.75]$ ,  $p<.001$ . Además, el efecto de números no difiere significativamente con respecto al efecto de los grupos A, B, C, ni tampoco difieren entre ellos, con ICs 95%  $[-31.30, -25,37]$ ,  $[-32.85, -27.51]$ ,  $[-29.73, -24.61]$ . Por otra parte, la diferencia entre lenguaje y álgebra tampoco resulta significativa,  $F(1, 6798)=[60.75]$ , ns. En cambio sí se encuentran diferencias significativas entre lenguaje y FyE,  $F(1, 6798)=[27.57]$ ,  $p>.001$ . Para revisar el resultado de las pruebas F para las demás formas ver **Anexo E**.

Resulta interesante que, controlando por las demás covariables, el efecto marginal (correspondiente al aumento de dos desviaciones estándar) del score en el eje de Números es similar en magnitud al efecto marginal de pertenecer a los grupos socioeconómicos A, B o C en comparación a pertenecer al grupo E. Esto podría significar que el aumento de dos desviaciones estándar en el eje de Números permitiría “compensar”, en términos de puntaje Simce, el hecho de pertenecer a los grupos socioeconómicos más bajos.

Adicionalmente, se puede ver que controlando por las demás covariables, las mujeres obtienen en todas las submuestras un puntaje significativamente peor que los hombres (entre 3.2 y 5.2 puntos menos). El grupo socioeconómico se presenta en comparación al grupo E (más alto) y se puede observar como al disminuir el grupo al que pertenece el estudiante, el puntaje de Matemáticas va disminuyendo significativamente.

## 5 CONCLUSIONES

A lo largo de este estudio se buscó, primero, describir las trayectorias académicas de los estudiantes entre cuarto y octavo básico a partir de cuatro distintas metodologías de clasificación. En particular, se busca encontrar el porcentaje de niños que se mantienen en el grupo de bajo desempeño. En segundo lugar, se busca definir qué conocimientos previos en los ejes de la asignatura de matemáticas son más importantes para determinar el desempeño general futuro en este ámbito. Y tercero, cuáles son los factores sociodemográficos y de desempeño más importantes para determinar el desempeño curricular futuro de los estudiantes en matemáticas. El abordar el tema en una muestra representativa nacional y desde una perspectiva curricular, permite que las conclusiones obtenidas orienten el diseño de las políticas públicas en dicha área. En particular, es importante recordar que en Chile, tal como muchos sistemas escolares (como es el caso de Estados Unidos presentados en la introducción), no se logran impartir los contenidos que se propone el marco curricular formal para cada año académico.

En primer lugar, se evidencia que los porcentajes que describen cada clase trayectoria varían importantemente según los criterios de clasificación empleados en cada año. Peso lo anterior, en todas las clasificaciones y para las dos asignaturas, el grupo que logra recuperarse de un bajo desempeño inicial corresponde a menos de un 12% del total de alumnos entre cuarto y octavo básico. Por otra parte, los resultados evidencian una alta persistencia de las trayectorias académicas, en especial de los desempeños bajos por sobre los altos,

en los cuales más de un 77% de los alumnos que falla en cuarto, vuelve a fallar en octavo. En particular, para matemáticas se observa más de un ... y para lenguaje más de un.

Sin embargo, y en segundo lugar, la importancia del eje curricular de números para el desempeño futuro en matemáticas, muestra que los resultados a nivel de desempeño curricular son consistentes con los resultados obtenidos desde la investigación cognitiva. En un contexto de baja implementación curricular, este resultado aporta evidencia a la política educativa sobre la importancia de trabajar las habilidades numéricas tempranas para un mejor desempeño general en matemática de los estudiantes.

En tercer lugar, la importancia de números es reforzada al considerar otro hallazgo importante de este estudio, en relación a que una mejora de dos desviaciones estándar en el puntaje del eje de números permite en alguna medida compensar las desventajas relacionadas con el nivel socioeconómico de los estudiantes. Este resultado es particularmente relevante en un país como Chile, donde las diferencias socioeconómicas y culturales en los resultados de aprendizaje de los estudiantes es una de las más grandes, tal como muestran los resultados de PISA (OECD, 2013a).

En cuarto lugar, los resultados muestran la importancia también de lenguaje en el desempeño general en matemáticas. Aún más, tener un buen desempeño en lenguaje resultó ser casi tan importante como manejar álgebra y más importante

que manejar los ejes de formas y espacio, y datos y azar. Este resultado es consistente con otras investigaciones que muestran la importancia del lenguaje para el desempeño en matemáticas (Mullis et al., 2013; Norris & Phillips, 2003; Savolainen et al., 2008; Vilenius-Tuohimaa et al., 2008). Por otra parte, entrega indicios sobre la relevancia de afianzar las alfabetizaciones básicas como números, lectura y escritura para un buen desempeño académico en matemáticas. Considerando este resultado, sería relevante estudiar si este es el caso sólo para matemáticas o también para otras asignaturas como ciencias.

Finalmente, los resultados son consistentes con investigaciones previas que muestran diferencias de género en el desempeño de los estudiantes chilenos en matemáticas (Mullis et al., 2012; OECD, 2013b). Este resultado en conjunto con los demás hallazgos, señala la importancia de reforzar el trabajo en el eje de Números focalizándose especialmente en las estudiantes mujeres.

En síntesis, los resultados de este estudio entregan evidencia a la política educativa sobre la importancia de reforzar los contenidos curriculares de números y de lenguaje para el desempeño futuro en matemáticas, y en especial, de hacerlo con estudiantes mujeres y de nivel socioeconómico más bajo. Finalmente, para saber si estos resultados trascienden a la realidad chilena, es relevante poder estudiar otros países en investigaciones futuras.

## BIBLIOGRAFÍA

Agencia de Calidad de la Educación (2014). Bases de datos nacionales. En *Archivos públicos*. Disponible en <http://www.agenciaeducacion.cl/simce/bases-de-datos-nacionales/>

Agencia de Calidad de la Educación (2013). Distribución nacional de estudiantes de 4° básico según Estándares de Aprendizaje. En *Síntesis de resultados Simce 4.° Básico 2013*. Disponible en <http://www.agenciaeducacion.cl/simce/resultados-simce/>

Agencia de Calidad de la Educación (2013). Documentos de trabajo. En *Documento N.° 3: Alineamiento de pruebas Simce y TIMSS de 4.º básico*. Disponible en <http://goo.gl/LH7cH9>

Agencia de Calidad de la Educación (2013). Simce. En *¿Qué es el Simce?* Disponible en <http://www.agenciaeducacion.cl/simce/que-es-el-simce/>

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Belsley, D. A., Kuh, E., & Welsch, R. E. Regression diagnostics: identifying influential data and sources of collinearity, 1980.

Bourdieu, P., & Passeron, J. C. (2009). Los herederos: los estudiantes y la cultura.

Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology*, 81(1), 78-96.

Cornejo Chávez, R. (2006). El experimento educativo chileno 20 años después: una mirada crítica a los logros y falencias del sistema escolar.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, 14(20), 1-11.

Eyzaguirre, B. (2004). Claves para la educación en pobreza. *Estudios Públicos*, 93(04).

Franzese, R., & Kam, C. (2009). *Modeling and interpreting interactive hypotheses in regression analysis*. University of Michigan Press.

Fröhlich López, Raimundo. (2014). *Análisis de la distribución del desempeño de los estudiantes en la prueba SIMCE en términos del dominio de los contenidos*

*curriculares* (Tesis de magíster). Pontificia Universidad Católica de Chile. Escuela de Ingeniería.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine*, 27(15), 2865-2873.

Hamilton, L. (2012). *Statistics with STATA: Version 12*. Cengage Learning.

Hidden curriculum (2014, August 26). In S. Abbott (Ed.), *The glossary of education reform*. Retrieved from <http://edglossary.org/hidden-curriculum>

Jiménez, R. M. G. (2003). Diferencias de género en el desempeño matemático de estudiantes de secundaria. *Educación Matemática*, 15(2), 129-161.

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: kindergarten number competence and later mathematics outcomes. *Developmental psychology*, 45(3), 850.

Krajewski, K., & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction*, 19(6), 513-526.

Kleinbaum, D., Kupper, L., Nizam, A., & Rosenberg, E. (2013). *Applied regression analysis and other multivariable methods*. Cengage Learning.

Long, J. S. (2009). Group comparisons in logit and probit using predicted probabilities. *Department of Sociology, University of Indiana. Unpublished manuscript*.

Long, J. S., & Freese, J. (2001). Listing and interpreting transformed coefficients from certain regression models. *Stata Technical Bulletin*, 10(57).

Marquardt, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, 12(3), 591-612.

Mizala, A., & Romaguera, P. (1998). *Desempeño escolar y elección de colegios: la experiencia chilena*. Centro de Economía Aplicada.

Mandel, J. (1985). The regression analysis of collinear data. *J. Res. Natl. Bur. Stand*, 90(6), 465-476.

Mason, C. H., & Perreault Jr, W. D. (1991). Collinearity, power, and interpretation of multiple regression analysis. *Journal of Marketing Research*, 268-280.

Ministerio de Educación (2011). Documentos de Orientaciones para Docentes y Directivos. En *Folleto de Orientaciones Simce 2011 / 8.º básico*. Disponible en <http://www.agenciaeducacion.cl/biblioteca-digital/folletos-de-orientaciones-2/>

Ministerio de Educación (2010). Documentos de Orientaciones para Docentes y Directivos. En *Folleto de Orientaciones Simce 2010 / 4.º básico*. Disponible en <http://www.agenciaeducacion.cl/biblioteca-digital/folletos-de-orientaciones-2/>

Ministerio de Educación (2008). Resumen Ejecutivo Metodología Puntajes de Corte Pruebas 4.º básico Simce. En *Documentos Técnicos Simce*. Disponible en <http://www.agenciaeducacion.cl/biblioteca-digital/documentos-tecnicos-simce/>

Ministerio de Educación (2007). Documentos de Orientaciones para Docentes y Directivos. En *Folleto de Orientaciones Simce 2007 / 4.º básico*. Disponible en <http://www.agenciaeducacion.cl/biblioteca-digital/folletos-de-orientaciones-2/>

Nussbaum, M., Claro, M., Frohlich, R., Claro, F., Cabello, T., Santana, M., San Martín, E., & Matus, C. (2015) Curricular achievement under a sequential learning system. Submitted for publication.

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673-690.

Passolunghi, M. C., & Lanfranchi, S. (2012). Domain-specific and domain-general precursors of mathematical achievement: A longitudinal study from kindergarten to first grade. *British Journal of Educational Psychology*, 82(1), 42-63.

Williams, R. (2014). Graduate Statistics II. En *Imposing and Testing Equality Constraints in Models* (En Part II). Recuperado de <https://www3.nd.edu/~rwilliam/stats2/>

Zamora, G. (2011). Movilidad escolar en Chile: Análisis de las implicancias para la calidad y equidad de la educación. *Estudios pedagógicos (Valdivia)*, 37(1), 53-69.

Valenzuela, Juan Pablo; Allende, Claudio; Sevilla, Alejandro; Egaña, Pablo (2013). *La (in)movilidad del desempeño educativo de los estudiantes chilenos: realidad, oportunidades y desafíos*". En Centro de Estudios MINEDUC. "Evidencias para políticas públicas en educación: Selección de Investigaciones Sexto Concurso FONIDE" (pp. 187-231). Recuperado de <http://centroestudios.mineduc.cl/index.php?t=96>

**ANEXOS**

## **ANEXO A: RESULTADOS POR FORMA ANÁLISIS DESCRIPTIVO**

### **1. Contenidos evaluados en SIMCE 4to Básico 2007**

Se evalúan tres ejes temáticos a través de los cuales se evalúa de manera integrada la resolución de problemas. Los ejes son:

*Números: En este eje se evaluará la capacidad de los estudiantes para **leer y escribir números** naturales y establecer relaciones entre estos, **ordenándolos**, buscando regularidades sencillas en secuencias o completando igualdades que tienen una incógnita en distintos lugares. También se contemplará la comprensión de **fracciones** a partir de su identificación en contextos gráficos o su relación con las partes de un todo. Finalmente, se evaluará el uso de los números para leer, interpretar y organizar información en tablas y gráficos. La resolución de problemas se evaluará en tareas relativas a la formación de números a partir de sus dígitos. En estos problemas, así como en la capacidad para descomponer y ordenar números se evaluará implícitamente la comprensión del **valor posicional de los dígitos que forman un número**.*

*Operaciones aritméticas: En este eje se evaluará la capacidad de los estudiantes para usar la suma, resta, multiplicación y división de números naturales en distintas situaciones, y el uso de algoritmos convencionales para realizar operaciones. Los estudiantes enfrentarán una gama amplia de problemas que requieren seleccionar datos y operaciones, realizar cálculos y verificar resultados. Al mismo tiempo enfrentarán problemas que demandan idear estrategias simples para su resolución.*

*Forma y espacio: En este eje se evaluará la orientación espacial, la capacidad para describir e interpretar trayectorias y ubicaciones en planos esquemáticos y la capacidad de relacionar el punto desde donde se observa un objeto con la representación gráfica de este. En este eje también se evaluará el reconocimiento, la comparación y la clasificación de figuras y cuerpos geométricos. La resolución de problemas se evaluará en tareas que requieren componer y descomponer formas geométricas de dos o tres dimensiones.*

(Ministerio de Educación, 2007, pp. 14-15).

**ANEXO B: CASOS PERDIDOS POR LIMPIEZA DE BASES**

Bases originales (lenguaje y matemáticas en un mismo archivo)

<b>Año</b>	<b>2007</b>	<b>2011</b>
<b>mrnun perdidos</b>	<b>6</b>	<b>112</b>

**A1.1. Primera fase de limpieza**

<b>Matemáticas</b>	<b>2007</b>	<b>2011</b>
<b>Por mrnun duplicados</b>	1.762	902
<b>Por valores perdidos</b>	14.936	8.031
<b>No match idalumno-idforma</b>	2717	47
<b>No match n/a (-) - forma</b>	206	200
<b>No match mrnun info. alumno</b>	0	0
<b>No match cuestionario padres</b>	54.753	No aplica

<b>Lenguaje</b>	<b>2007</b>	<b>2011</b>
<b>Por mrnun duplicados</b>	1.761	910
<b>Por valores perdidos</b>	14.727	-
<b>No match idalumno-idforma</b>	3.803	61
<b>No match n/a (-) - forma</b>	182	204
<b>No match mrnun info. alumno</b>	0	0
<b>No match cuestionario padres</b>	No aplica	No aplica

**A1.2. Segunda fase de limpieza:**

<b>Matemáticas</b>	<b>2007/2011</b>
<b>n parcial</b>	<b>168.673</b>

-

**Resultados merge: Casos perdidos**

<b>Matemáticas</b>	<b>2007/2011</b>
<b>Mantienen mismo género</b>	629
<b>GSE no cambia más de un nivel</b>	3.772
<b>Tiene ptje. Lenguaje cuarto</b>	1557
<b>Tiene ptje. Matemáticas octavo</b>	39

<b>Matemáticas</b>	<b>2007/2011</b>
<b>n parcial</b>	<b>162.676</b>

## ANEXO C: RESULTADOS DIAGNÓSTICO DE COLINEARIDAD PARA FORMAS K, L Y

### M.

- Diagnóstico con constante

Forma K										
Condition	eigenvalue	index	_cons	Num.	Alg.	FyE	DyA	Leng.	Fem.	GSE
1	7,1856	1	0	0	0	0	0	0	0,01	0
2	0,4772	3,88	0	0	0,01	0	0	0	0,85	0
3	0,1890	6,17	0	0,01	0,04	0	0	0	0,06	0,34
4	0,0588	11,06	0	0,01	0,82	0,04	0,11	0,01	0,03	0,07
5	0,0322	14,94	0,02	0,03	0,04	0,23	0,83	0,01	0	0,02
6	0,0239	17,34	0,14	0,8	0,06	0,04	0,01	0,03	0,01	0,23
7	0,0228	17,75	0,14	0,07	0,02	0,68	0,03	0,14	0,04	0,17
8	0,0105	26,14	0,69	0,08	0,01	0	0,01	0,8	0	0,17

Forma L										
Condition	eigenvalue	index	_cons	Num.	Alg.	FyE	DyA	Leng.	Fem.	GSE
1	7,1454	1	0	0	0	0	0	0	0,01	0
2	0,4871	3,83	0	0	0,01	0	0	0	0,81	0
3	0,2052	5,9	0,01	0,01	0,03	0	0,01	0	0,09	0,3
4	0,0635	10,61	0	0,01	0,86	0,02	0,15	0,01	0,04	0,03
5	0,0386	13,61	0,04	0,01	0,03	0,19	0,7	0,03	0,01	0,15
6	0,0279	16	0,06	0,82	0,06	0,08	0,12	0	0	0,12
7	0,0223	17,9	0,15	0	0	0,69	0	0,19	0,04	0,23
8	0,0100	26,75	0,75	0,16	0	0,01	0,01	0,78	0	0,18

Forma M										
Condition	eigenvalue	index	_cons	Num.	Alg.	FyE	DyA	Leng.	Fem.	GSE
1	7,1675	1	0	0	0	0	0	0	0,01	0
2	0,4823	3,86	0	0	0,01	0	0	0	0,84	0
3	0,1949	6,06	0,01	0,01	0,02	0	0,02	0	0,07	0,32
4	0,0576	11,15	0	0,01	0,86	0,01	0,2	0	0,03	0
5	0,0394	13,49	0,04	0,02	0,09	0,16	0,62	0,02	0,01	0,18
6	0,0264	16,47	0,09	0,83	0,01	0,02	0,1	0,03	0,02	0,17
7	0,0219	18,11	0,11	0,06	0,01	0,79	0	0,15	0,03	0,12
8	0,0100	26,74	0,75	0,06	0	0	0,05	0,79	0	0,2

- Diagnóstico sin constante

Forma K

condition	index	Num.	Alg.	FyE	DyA	Leng.	Fem.	GSE
1	1	0	0	0	0	0	0,01	0
2	3,61	0	0	0	0	0	0,85	0
3	5,89	0,01	0	0	0	0	0,06	0,8
4	10,33	0,01	0,8	0,1	0,1	0,01	0,03	0,1
5	14,03	0,06	0,1	0,3	0,8	0,01	0	0
6	16,31	0,72	0,1	0,5	0	0,01	0	0
7	19,05	0,2	0	0,1	0	0,97	0,04	0,2

Forma L

condition	index	Num.	Alg.	FyE	DyA	Leng.	Fem.	GSE
1	1	0	0	0	0	0	0,01	0
2	3,56	0	0	0	0	0	0,82	0
3	5,69	0,01	0	0	0	0	0,09	0,7
4	9,88	0,01	0,9	0	0,2	0,01	0,04	0
5	13,01	0,05	0,1	0,3	0,8	0,04	0,01	0,1
6	15,29	0,75	0	0,5	0	0	0	0
7	18,67	0,18	0	0,2	0	0,95	0,04	0,2

Forma M

condition	index	Num.	Alg.	FyE	DyA	Leng.	Fem.	GSE
1	1	0	0	0	0	0	0,01	0
2	3,59	0	0	0	0	0	0,85	0
3	5,82	0,01	0	0	0	0	0,06	0,8
4	10,37	0,01	0,9	0	0,2	0	0,03	0
5	12,97	0,08	0,1	0,3	0,7	0,03	0	0,1
6	15,87	0,83	0	0,4	0	0,01	0	0
7	18,65	0,07	0	0,3	0	0,96	0,05	0,1



*SIMCE 8vo en matemática*

\* Por aprobarse entiende cumplir con el 60% de aprobación

Porcentaje M3%				
4to\ 8vo	C	D	E	G
J	1,31%	1,51%	1,43%	1,58%
K	1,10%	0,81%	1,01%	1,00%
L	1,02%	0,86%	1,05%	1,10%
M	0,93%	0,77%	1,03%	0,82%

Desv. Est.	Min	Max
0,249%	0,77%	1,58%

M3 = Niños que fallan\*\* SIMCE 4to y aprueban\* SIMCE 8vo en matemática

\* Por aprobarse entiende cumplir con el 60% de aprobación

\*\* Por fallar se entiende obtener menor al 60% de aprobación

*SIMCE 8vo en matemática*

\* Por aprobarse entiende cumplir con el 60% de aprobación

\*\* Por fallar se entiende obtener menor al 60% de aprobación

Porcentaje M4%				
4to\ 8vo	C	D	E	G
J	33,30%	32,88%	33,08%	33,65%
K	29,30%	29,53%	28,94%	29,49%
L	31,47%	32,36%	30,37%	30,41%
M	27,60%	27,90%	27,33%	27,28%

Desv. Est.	Min	Max
2,241%	27,28%	33,65%

M4 = Niños que fallan\*\* SIMCE 4to y fallan\*\* SIMCE 8vo en matemática

\*\* Por fallar se entiende obtener menor al 60% de aprobación

Porcentaje L1%		Porcentaje L2%		
Niños por combinación de formas* (N)				
4to\ 8vo	C	D	E	G
J	14.674	11.879	10.981	10.713
K	11.900	12.647	11.564	10.645
L	10.990	11.449	11.611	10.986
M	10.608	10.688	10.901	11.109

\*Correspondientes a las pruebas de matemática 2007 y 2011

LENGUAJE
----------

4to\ 8vo	H	I	J	K
E	41,15%	40,50%	34,44%	43,68%
G	38,13%	37,09%	32,98%	40,54%
H	40,87%	39,30%	34,52%	42,56%
I	35,08%	33,54%	30,27%	36,89%

Desv. Est.	Min	Max
3,858%	30,27%	43,68%

4to\ 8vo	H	I	J	K
E	13,34%	15,24%	21,23%	10,84%
G	10,17%	11,40%	15,59%	8,24%
H	12,27%	13,84%	19,23%	10,55%
I	7,34%	8,64%	11,58%	6,20%

Desv. Est.	Min	Max
4,110%	6,20%	21,23%

L1 = Niños que aprueban\* SIMCE 4to y aprueban\* SIMCE 8vo en lenguaje

\* Por aprobarse entiende cumplir con el 69% de aprobación

L2 = Niños que aprueban\* SIMCE 4to y fallan\*\* SIMCE 8vo en lenguaje

\* Por aprobarse entiende cumplir con el 69% de aprobación

\*\* Por fallar se entiende obtener menor al 69% de aprobación

Porcentaje L3%				
4to\ 8vo	H	I	J	K
E	10,08%	8,91%	5,09%	12,22%
G	13,53%	11,36%	7,90%	15,34%
H	11,61%	9,22%	5,93%	13,17%
I	17,13%	14,65%	9,78%	18,57%

Porcentaje L4%				
4to\ 8vo	H	I	J	K
E	35,42%	35,34%	39,25%	33,26%
G	38,17%	40,15%	43,53%	35,88%
H	35,25%	37,64%	40,32%	33,71%
I	40,46%	43,17%	48,37%	38,35%

Desv. Est.	Min	Max
3,792%	5,09%	18,57%

L3 = Niños que fallan\*\* SIMCE 4to y aprueban\* SIMCE 8vo en lenguaje

\* Por aprobarse entiende cumplir con el 69% de aprobación

Desv. Est.	Min	Max
4,023%	33,26%	48,37%

L4 = Niños que fallan\*\* SIMCE 4to y fallan\*\* SIMCE 8vo en lenguaje

\*\* Por fallar se entiende obtener menor al 69% de aprobación

Niños por combinación de formas* (N)				
4to\ 8vo	C	D	E	G
J	14.674	11.879	10.981	10.713
K	11.900	12.547	11.564	10.645
L	10.990	11.449	11.611	10.986
M	10.608	10.686	10.901	11.109

\*Correspondientes a las pruebas de matemática 2007 y 2011

- Aprobación por forma
  - Criterio por forma

Asignatura
Matemática
Lenguaje

Criterios de Aprobación Finales			
Cuarto		Octavo	
Forma	Criterio	Forma	Criterio
J	0,73050454	C	0,607839994
K	0,73132713	D	0,614338207
L	0,72715295	E	0,616006899
M	0,75018944	G	0,627474575
E	0,66261306	H	0,714908787
G	0,64733033	I	0,709438618
H	0,64102564	J	0,6847953
I	0,63182485	K	0,728030539

Cuadro 1. Proporción de preguntas correctas mínimas necesarias para ubicarse en un desempeño adecuado. Cada forma se adecúa según los estándares nacionales.

Puntaje Simce Normativo Adecuado		
Asignatura	4to	8vo
Matemática	295 o más	297 o más
Lenguaje	284 o más	292 o más

### Estándares de aprendizaje

- Resultado por combinación de formas

Porcentaje M1%				
4to\ 8vo	C	D	E	G
J	25,92%	25,69%	24,18%	23,39%
K	26,96%	26,59%	26,02%	24,73%
L	27,34%	26,35%	25,98%	24,98%
M	27,15%	26,38%	24,95%	24,96%

Porcentaje M2%				
4to\ 8vo	C	D	E	G
J	18,17%	18,64%	19,77%	20,46%
K	21,37%	22,30%	22,67%	23,55%
L	21,58%	20,75%	22,88%	23,78%
M	20,38%	20,55%	22,25%	23,66%

Desv. Est.	Min	Max

Desv. Est.	Min	Max

1,110%	23,39%	27,34%
--------	--------	--------

M1 = Niños que aprueban\* SIMCE 4to y aprueban\* SIMCE 8vo en matemática

\* Por aprobarse entiende cumplir con el criterio de aprobación por forma

1,723%	18,17%	23,78%
--------	--------	--------

M2 = Niños que aprueban\* SIMCE 4to y fallan\*\* SIMCE 8vo en matemática

\* Por aprobarse entiende cumplir con el criterio de aprobación por forma  
\*\* Por fallar se entiende obtener menor al criterio de aprobación por forma

Porcentaje M3%				
4to\ 8vo	C	D	E	G
J	4,75%	4,99%	4,05%	4,29%
K	3,81%	3,57%	3,14%	2,65%
L	3,65%	3,50%	3,02%	2,83%
M	4,12%	3,59%	3,66%	3,02%

Desv. Est.	Min	Max
0,663%	2,65%	4,99%

M3 = Niños que fallan\*\* SIMCE 4to y aprueban\* SIMCE 8vo en matemática

\* Por aprobarse entiende cumplir con el criterio de aprobación por forma  
\*\* Por fallar se entiende obtener menor al criterio de aprobación por forma

Porcentaje M4%				
4to\ 8vo	C	D	E	G
J	51,17%	50,68%	52,00%	51,85%
K	47,87%	47,54%	48,17%	49,07%
L	47,42%	49,39%	48,11%	48,42%
M	48,35%	49,48%	49,14%	48,37%

Desv. Est.	Min	Max
1,484%	47,42%	52,00%

M4 = Niños que fallan\*\* SIMCE 4to y fallan\*\* SIMCE 8vo en matemática

\*\* Por fallar se entiende obtener menor al criterio de aprobación por forma

Niños por combinación de formas* (N)				
4to\ 8vo	C	D	E	G
J	14.674	11.879	10.981	10.713
K	11.900	12.547	11.564	10.645
L	10.990	11.449	11.611	10.986
M	10.608	10.686	10.901	11.109

N	183.243
---	---------

\*Correspondientes a las pruebas de matemática 2007 y 2011

## LENGUAJE

Porcentaje L1%				
4to\ 8vo	H	I	J	K
E	40,12%	39,29%	35,35%	43,26%
G	39,21%	38,16%	35,65%	42,38%
H	41,47%	39,03%	36,51%	43,67%

Porcentaje L2%				
4to\ 8vo	H	I	J	K
E	18,36%	20,89%	23,99%	15,39%
G	16,61%	18,62%	20,52%	13,82%
H	19,45%	21,95%	24,96%	17,30%

I	37,98%	35,48%	33,24%	39,60%
---	--------	--------	--------	--------

I	12,66%	14,44%	16,46%	10,86%
---	--------	--------	--------	--------

Desv. Est.	Min	Max
3,016%	33,24%	43,67%

L1 = Niños que aprueban\* SIMCE 4to y aprueban\* SIMCE 8vo en lenguaje

\* Por aprobarse entiende cumplir con el criterio de aprobación por forma

Desv. Est.	Min	Max
3,987%	10,86%	24,96%

L2 = Niños que aprueban\* SIMCE 4to y fallan\*\* SIMCE 8vo en lenguaje

\* Por aprobarse entiende cumplir con el criterio de aprobación por forma

\*\* Por fallar se entiende obtener menor al criterio de aprobación por forma

Porcentaje L3%				
4to\ 8vo	H	I	J	K
E	6,63%	5,52%	4,18%	8,13%
G	7,91%	6,14%	5,23%	9,12%
H	6,25%	5,08%	3,93%	7,42%
I	9,95%	8,59%	6,81%	11,77%

Porcentaje L4%				
4to\ 8vo	H	I	J	K
E	34,88%	34,29%	36,48%	33,22%
G	36,26%	37,07%	38,60%	34,69%
H	32,83%	33,93%	34,60%	31,61%
I	39,40%	41,49%	43,49%	37,77%

Desv. Est.	Min	Max
2,139%	3,93%	11,77%

L3 = Niños que fallan\*\* SIMCE 4to y aprueban\* SIMCE 8vo en lenguaje

\* Por aprobarse entiende cumplir con el criterio de aprobación por forma

\*\* Por fallar se entiende obtener menor al criterio de aprobación por forma

Desv. Est.	Min	Max
3,235%	31,61%	43,49%

L4 = Niños que fallan\*\* SIMCE 4to y fallan\*\* SIMCE 8vo en lenguaje

\*\* Por fallar se entiende obtener menor al criterio de aprobación por forma

Niños por combinación de formas* (N)				
4to\ 8vo	H	I	J	K
E	14.426	11.592	10.770	10.523
G	11.659	12.392	11.290	10.423
H	10.759	11.160	11.388	10.765
I	10.423	10.452	10.676	10.847

N
179.545

\*Correspondientes a las pruebas de lenguaje 2007 y 2011

**ANEXO E: DIAGNÓSTICO DE DIFERENCIAS ENTRE COEFICIENTES. PRUEBAS F.**

• **Forma J**

F-ratio (Sub-sample J)					
	Alg.	FyE	DyA	Leng.	Fem.
Num.	63,64 (0,0000)	187,41 (0,0000)	398,93 (0,0000)	60,75 (0,0000)	1063,94 (0,0000)
Alg.		41,27 (0,0000)	218,72 (0,0000)	0,11 (0,7349)	668,7 (0,0000)
FyE			48,22 (0,0000)	27,57 (0,0000)	282,06 (0,0000)
DyA				133,44 (0,0000)	59,84 (0,0000)
Leng.					423,86 (0,0000)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 6798

F-ratio (Sub-sample J)				
	GSE B	GSE C	GSE D	GSE E
	146,68 (0,0000)	434,19 (0,0000)	490,36 (0,0000)	350,59 (0,0000)
GSE A		168,52 (0,0000)	247,23 (0,0000)	120,25 (0,0000)
GSE B			17,81 (0,0000)	1,44 (0,2295)
GSE C				3,66 (0,0557)
GSE D				

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 6798

## F-ratio (Sub-sample J)

	Numeros
GSE A	965,29 (0,0000)
GSE B	717,96 (0,0000)
GSE C	1195,21 (0,0000)
GSE D	1254,15 (0,0000)
GSE E	1055,77 (0,0000)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 6798

- **Forma K**

## F-ratio (Sub-sample K)

	Alg.	FyE	DyA	Leng.	Fem.
Num.	53,7 (0,0000)	163,2 (0,0000)	395,82 (0,0000)	85,39 (0,0000)	894,24 (0,0000)
Alg.		32,53 (0,0000)	203,66 (0,0000)	8,23 (0,0041)	629,99 (0,0000)
FyE			70,45 (0,0000)	7,89 (0,0050)	299,26 (0,0000)
DyA				108,43 (0,0000)	51,71 (0,0000)
Leng.					298,24 (0,0000)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 6451

F-ratio (Sub-sample K)				
	GSE B	GSE C	GSE D	GSE E
	107,7	294,43	354,93	254,36
GSE A	(0,0000)	(0,0000)	(0,0000)	(0,0000)
GSE B		135,22	217,47	94,49
GSE C		(0,0000)	(0,0000)	(0,0000)
GSE D			22,02	1,02
GSE E			(0,0000)	(0,3129)
				354,93
				(0,0134)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 6451

F-ratio (Sub-sample K)	
	Numeros
GSE A	883,52 (0,0000)
GSE B	540,91 (0,0000)
GSE C	892,94 (0,0000)
GSE D	992,2 (0,0000)
GSE E	800,12 (0,0000)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 6451

- **Forma L**

F-ratio (Sub-sample L)					
	Alg.	FyE	DyA	Leng.	Fem.
Num.	100,12 (0,0000)	27,35 (0,0000)	97,4 (0,0000)	14,23 (0,0000)	956,35 (0,0000)
Alg.		27,35 (0,0000)	14,4 (0,0000)	2,1 (0,0002)	610 (0,0000)
FyE			14,4 (0,0001)	2,1 (0,1476)	171,42 (0,0000)
DyA				28,37 (0,0000)	70,02 (0,0000)
Leng.					232,97 (0,0000)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 6190

F-ratio (Sub-sample L)				
	GSE B	GSE C	GSE D	GSE E
GSE A	56,94 (0,0000)	137,47 (0,0000)	164,16 (0,0000)	124,58 (0,0000)
GSE B		52,7 (0,0000)	96,38 (0,0000)	44,82 (0,0000)
GSE C			16,67 (0,0000)	1,55 (0,2128)
GSE D				1,85 (0,1738)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 6190

F-ratio (Sub-sample L)	
	Numeros
GSE A	1044,05 (0,0000)
GSE B	446,62 (0,0000)
GSE C	682,94 (0,0000)
GSE D	720,1 (0,0000)
GSE E	591,56 (0,0000)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 6190

- **Forma M**

F-ratio (Sub-sample M)					
	Alg.	FyE	DyA	Leng.	Fem.
Num.	59,47 (0,0000)	139,93 (0,0000)	235,21 (0,0000)	50,78 (0,0000)	883,78 (0,0000)
Alg.		35,24 (0,0000)	96,62 (0,0000)	0,08 (0,7707)	650,1 (0,0000)
FyE			13,04 (0,0003)	34,11 (0,0000)	240,99 (0,0000)
Leng.					451,32 (0,0000)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 5844

F-ratio (Sub-sample M)				
	GSE B	GSE C	GSE D	GSE E
	107,73	241,41	315,36	250,63
GSE A	(0,0000)	(0,0000)	(0,0000)	(0,0000)

		81,59	174,4	90,1
GSE B		(0,0000)	(0,0000)	(0,0000)
GSE C			35,81	8,79
			(0,0000)	(0,0030)
GSE D				315,36
				(0,1925)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 5844

F-ratio (Sub-sample M)	
	Numeros
GSE A	816,22 (0,0000)
GSE B	430,2 (0,0000)
GSE C	644,2 (0,0000)
GSE D	748,16 (0,0000)
GSE E	650,17 (0,0000)

Prob > F in parenthesis

between-group degrees of freedom = 1

within-group degrees of freedom = 5844

## ANEXO F: Alpha de cronbach y cantidad de preguntas por prueba y eje

- **Alpha Cronbach y preguntas por eje en matemáticas**

Tabla D-1. Alpha de Cronbach y cantidad de preguntas en sub-puntajes por eje

	Forma J	Forma K	Forma L	Forma M
Numeros	0,8237 (16)	0,8113 (16)	0,84 (17)	0,8246 (15)
Algebra	0,5292 (5)	0,6072 (5)	0,5523 (4)	0,5466 (4)
DyA	0,6523 (9)	0,6859 (11)	0,7105 (10)	0,6736 (10)
FyA	0,4524 (4)	0,6181 (5)	0,6209 (5)	(1) (6)

Number of items in the scale in parenthesis

- **Descripción variables subscore estandarizadas**

- Código Stata

```
. summ z2_tnumerosJ z2_talgebraJ z2_tFyEJ z2_tdatosyazarJ z2_ptje_lect_4to GSE2_4to female if
idforma_4to=="J"
. summ z2_tnumerosK z2_talgebraK z2_tFyEK z2_tdatosyazarK z2_ptje_lect_4to GSE2_4to female
if idforma_4to=="K"
. summ z2_tnumerosL z2_talgebraL z2_tFyEL z2_tdatosyazarL z2_ptje_lect_4to GSE2_4to female if
idforma_4to=="L"
. summ z2_tnumerosM z2_talgebraM z2_tFyEM z2_tdatosyazarM z2_ptje_lect_4to GSE2_4to
female if idforma_4to=="M"
```

- Output reformateado

Variable	Forma J (n=42,789)				Forma K (n=41,342)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Números	1,491	0,500	0,000	2,144	1,634	0,500	0,000	2,256
Algebra	1,265	0,500	0,000	1,855	1,263	0,500	0,000	1,826
FyE	1,681	0,500	0,000	2,293	1,836	0,500	0,000	2,432
DyE	1,556	0,500	0,000	1,957	1,746	0,500	0,000	2,116

Lenguaje	2,715	0,500	1,140	3,784	2,640	0,500	1,188	3,756
GSE	3,102	1,088	1,000	5,000	3,085	1,075	1,000	5,000
Género	0,521	0,500	0,000	1,000	0,520	0,500	0,000	1,000

Variable	Forma L (n=40.099)				Forma M (n=38.446)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Números	1,472	0,500	0,000	2,112	1,552	0,500	0,000	2,138
Algebra	1,195	0,500	0,000	1,726	1,331	0,500	0,000	1,826
FyE	1,836	0,500	0,000	2,424	1,899	0,500	0,000	2,477
DyE	1,426	0,500	0,000	1,934	1,380	0,500	0,000	1,889
Lenguaje	2,594	0,500	1,165	3,723	2,614	0,500	1,211	3,733
GSE	3,071	1,062	1,000	5,000	3,057	1,054	1,000	5,000
Género	0,521	0,500	0,000	1,000	0,518	0,500	0,000	1,000

- **Descripción variables subscore sin estandarizar**

- Código Stata

```

summ tnumerosJ talgebraJ tFyEJ tdatosyazarJ ptje_lect_4to female if idforma_4to=="J"
summ tnumerosK talgebraK tFyEK tdatosyazarK ptje_lect_4to female if idforma_4to=="K"
summ tnumerosL talgebraL tFyEL tdatosyazarL ptje_lect_4to female if idforma_4to=="L"
summ tnumerosM talgebraM tFyEM tdatosyazarM ptje_lect_4to female if idforma_4to=="M"

```

- Output reformateado

Variable	Forma J (n=42.789)				Forma K (n=41.342)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Números	1,009	0,338	0,000	1,450	1,062	0,325	0,000	1,466
Algebra	0,808	0,320	0,000	1,185	0,820	0,325	0,000	1,186
FyE	1,015	0,302	0,000	1,384	1,087	0,296	0,000	1,440
DyE	0,903	0,290	0,000	1,136	0,930	0,266	0,000	1,127
Lenguaje	267,497	49,262	112,320	372,780	264,976	50,180	119,210	376,960

Variable	Forma L (n=40.099)				Forma M (n=38.446)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Números	0,991	0,337	0,000	1,421	1,021	0,329	0,000	1,407
Algebra	0,761	0,319	0,000	1,100	0,803	0,302	0,000	1,102
FyE	1,062	0,289	0,000	1,402	1,090	0,287	0,000	1,422
DyE	0,855	0,300	0,000	1,160	0,874	0,317	0,000	1,197

Lenguaje 262,646 50,623 118,000 376,960                    265,678 50,812 123,110 379,350

## ANEXOG: REPRESENTATIVIDAD

```
. use "Definitivas\base_2007_2011_representatividad8vo_GSE.dta", clear
. encode GRUPO_8vo, generate(GSE_8vo)
```

```
. label list GSE_8vo
```

```
GSE_8vo:
  1 A
  2 B
  3 C
  4 D
  5 E
```

```
. encode DDCIA_8vo, generate(DCIA_8vo)
```

```
. label list DCIA_8vo
```

```
DCIA_8vo:
  1 MC
  2 MD
  3 PP
  4 PS
```

```
. drop if genero_8vo==" "
(0 observations deleted)
```

```
. encode genero_8vo, generate(FEMALE)
```

```
. label list FEMALE
```

```
FEMALE:
  1 F
  2 M
```

```
. describe IDREGION DCIA_8vo ptje_lect_8vo ptje_mat_8vo
```

```
storage display value
variable name type format label variable label
-----
IDREGION double %10.0g IDREGION
DCIA_8vo long %8.0g DCIA_8vo DDCIA
ptje_lect_8vo double %10.0g ptje_lect
ptje_mat_8vo double %10.0g ptje_mat
```

```
. tab IDREGION
```

```
IDREGION | Freq. Percent Cum.
-----+-----
  1 | 4,122 1.94 1.94
  2 | 7,880 3.72 5.66
  3 | 3,667 1.73 7.39
  4 | 8,999 4.24 11.63
  5 | 21,641 10.21 21.84
  6 | 11,916 5.62 27.46
  7 | 12,985 6.12 33.59
  8 | 25,509 12.03 45.62
```

9	12,667	5.97	51.59
10	10,551	4.98	56.57
11	1,315	0.62	57.19
12	2,042	0.96	58.15
13	81,077	38.24	96.39
14	4,892	2.31	98.70
15	2,755	1.30	100.00
-----			
Total	212,018	100.00	

```
. tab DCIA_8vo
```

DDCIA	Freq.	Percent	Cum.
MC	31,596	14.90	14.90
MD	60,781	28.67	43.57
PP	16,533	7.80	51.37
PS	103,108	48.63	100.00
-----			
Total	212,018	100.00	

```
. tab FEMALE
```

genero	Freq.	Percent	Cum.
F	106,201	50.09	50.09
M	105,817	49.91	100.00
-----			
Total	212,018	100.00	

```
. ttest FEMALE, by (sample)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	49342	1.562665	.0022332	.4960626	1.558288	1.567042
1	162676	1.479813	.0012387	.4995938	1.477385	1.48224
-----						
combined	212018	1.499094	.0010859	.5000004	1.496966	1.501223
-----						
diff		.082852	.0025634	.0778278	.0878763	
-----						
diff = mean(0) - mean(1)				t = 32.3209		
Ho: diff = 0				degrees of freedom = 212016		
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 1.0000		Pr( T  >  t ) = 0.0000		Pr(T > t) = 0.0000		

```
. ttest IDREGION, by (sample)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	49342	9.313789	.0171329	3.805747	9.280209	9.34737
1	162676	9.247105	.0094462	3.809941	9.22859	9.265619
-----						
combined	212018	9.262624	.0082724	3.80906	9.24641	9.278838
-----						
diff		.0666848	.019576	.0283164	.1050532	

```

diff = mean(0) - mean(1)          t = 3.4065
Ho: diff = 0                      degrees of freedom = 212016

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.9997    Pr(T |> |t|) = 0.0007    Pr(T > t) = 0.0003

```

```
. ttest DCIA_8vo, by (sample)
```

```
Two-sample t test with equal variances
```

```

-----+-----
Group | Obs   Mean  Std. Err.  Std. Dev.  [95% Conf. Interval]
-----+-----
0 | 49342  2.792631 .0053971  1.198861  2.782053  2.803209
1 | 162676  2.934637 .0028605  1.153742  2.92903  2.940244
-----+-----
combined | 212018  2.901589 .0025322  1.165941  2.896626  2.906552
-----+-----
diff |      -1.420059 .0059844      -1.537351  -.1302767
-----+-----

```

```

diff = mean(0) - mean(1)          t = -23.7295
Ho: diff = 0                      degrees of freedom = 212016

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.0000    Pr(T |> |t|) = 0.0000    Pr(T > t) = 1.0000

```

```
. ttest ptje_lect_8vo, by (sample)
```

```
Two-sample t test with equal variances
```

```

-----+-----
Group | Obs   Mean  Std. Err.  Std. Dev.  [95% Conf. Interval]
-----+-----
0 | 47595  238.3952 .2246208  49.00391  237.9549  238.8355
1 | 159610  260.4728 .1226627  49.00525  260.2324  260.7132
-----+-----
combined | 207205  255.4016 .1095722  49.87701  255.1868  255.6164
-----+-----
diff |      -22.07764 .2559348      -22.57926  -21.57601
-----+-----

```

```

diff = mean(0) - mean(1)          t = -86.2628
Ho: diff = 0                      degrees of freedom = 207203

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.0000    Pr(T |> |t|) = 0.0000    Pr(T > t) = 1.0000

```

```
. ttest ptje_mat_8vo, by (sample)
```

```
Two-sample t test with equal variances
```

```

-----+-----
Group | Obs   Mean  Std. Err.  Std. Dev.  [95% Conf. Interval]
-----+-----
0 | 49290  243.5554 .2057577  45.681  243.1521  243.9586
1 | 162676  264.7823 .1205761  48.6321  264.546  265.0187
-----+-----
combined | 211966  259.8463 .1059803  48.79306  259.6386  260.054
-----+-----
diff |      -21.22698 .2465982      -21.71031  -20.74366
-----+-----

```

```

diff = mean(0) - mean(1)          t = -86.0792
Ho: diff = 0                      degrees of freedom = 211964

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0
Pr(T < t) = 0.0000    Pr(T |> |t|) = 0.0000    Pr(T > t) = 1.0000

```

**ANEXO H: CARTA DE RECEPCIÓN DE PAPER**

Se enviará antes de la defensa la versión en inglés.

## **ANEXO I: PAPER POR ENVIAR**

**Nota: Se enviará la versión en inglés antes de la defensa**

### **Factores que explican el desempeño académico en matemáticas: un estudio longitudinal a escala nacional y en contexto curricular.**

Macarena Santana, Miguel Nussbaum, Tania Cabello, Magdalena Claro, Luis Maldonado

#### **1. Introducción**

Parte de la discusión sobre la calidad de la educación tiene relación con la pertinencia de los fines y se refiere a las habilidades, conocimientos y valores en los cuales es más adecuado y deseable formar a los individuos para integrarse a la sociedad. Es una discusión amplia que involucra aspectos culturales, políticos, sociales y económicos y que en el sistema educativo queda expresada en el marco curricular (UNESCO, 2005). Históricamente, el objetivo de la educación formal ha ido evolucionando de acuerdo a las necesidades de la sociedad. Partió con el interés de la iglesia en la Edad Media por instruir a las nuevas generaciones en la doctrina religiosa, y siguió luego de la formación de los estados nacionales con el interés creciente del estado por formar la conciencia ciudadana como base del orden político (Ramirez & Meyer, 1980). Gran parte de los sistemas modernos de educación se formaron hacia finales del siglo XIX cuando la educación primaria pasó a ser universal y obligatoria (Boyd, 1966). La revolución industrial significó otro impulso al demandar mano de obra más calificada y sobre todo, a partir de la segunda mitad del siglo XX, la educación pasó a ser vista como un capital y por lo tanto, motor fundamental del desarrollo económico de los países (Becker, 2009; Schultz, 1961).

Desde fines del siglo XX se ha comenzado a hablar de las habilidades del siglo XXI, demandadas por una sociedad donde la producción del conocimiento se transforma en el motor del desarrollo económico y social, reemplazando las materias primas y la fuerza de trabajo de la sociedad industrial (Anderson, 2008; Castells, 2002; Griffin, Care, & McGaw, 2012). En este contexto, las propuestas siglo XXI plantean la necesidad de disminuir la carga de aprendizaje de contenidos para orientarse al desarrollo de las distintas alfabetizaciones o *literacies* (lectura, escritura, matemáticas y TIC) y habilidades de pensamiento de orden superior (creatividad, pensamiento crítico, resolución de problemas, toma de decisiones, entre otros (Binkley et al., 2012; OECD, 2010; UNESCO, 2005). Lo anterior implica un cambio en la noción del aprender, desde *el aprender qué al aprender cómo*, lo que sentaría las bases del aprendizaje

permanente (lifelong learning) en contextos que son cada vez más cambiantes y complejos (Anderson, 2008). En consecuencia, en términos curriculares se plantea la necesidad de organizar el tiempo de tal manera de tener mayor espacio para ejercitar y aplicar los contenidos, investigar y construir el propio conocimiento por medio del trabajo de proyectos (Trilling & Fadel, 2009).

Lo anterior se vuelve más importante si se considera que en muchos sistemas escolares no se logran impartir los contenidos que se propone el marco curricular formal para cada año académico. Por ejemplo, un estudio realizado en Estados Unidos mostró que no se estaban trabajando los contenidos y habilidades correspondientes a cada nivel (Schmidt, et.al. 2011). En Chile los resultados son similares, en el sentido de que ninguno de los niveles de enseñanza analizados alcanza un 100% de la cobertura curricular (Ministerio de Educación, 2004, 2013), variando de un 41% a un 73% de cobertura en matemáticas dependiendo de la fuente de información consultada (Ministerio de Educación, 2013).

En este contexto de baja implementación curricular y la necesidad de dejar espacio para desarrollar en los estudiantes las habilidades básicas para pensar de forma autónoma, resulta relevante entender qué es más importante reforzar y priorizar en el currículum de cada asignatura. En la presente investigación interesa responder a esta pregunta para el caso de matemáticas.

Desde el ámbito de la investigación cognitiva, existen estudios que encuentran asociaciones entre la alfabetización numérica temprana y el rendimiento posterior en matemáticas (Geary, 2011; Gersten, Jordan, & Flojo, 2005; Jordan, Glutting, & Ramineni, 2010; Passolunghi & Lanfranchi, 2012; Träff, 2013). Específicamente, se encuentra que las habilidades tempranas de números son una precondition para el pensamiento de segundo orden matemático (Chard et al., 2005; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Krajewski & Schneider, 2009; Passolunghi & Lanfranchi, 2012). Sin embargo, la gran mayoría de estudios encontrados en este ámbito se realizan en base a aplicaciones de pequeña y mediana escala, con un número de casos entre 70 y 411 estudiantes en un número bajo de escuelas. Además, todos los estudios encontrados abordan el tema desde una mirada cognitiva, que no permite orientar de forma directa a las políticas públicas. Considerando lo anterior, la presente investigación, por una parte, analiza una muestra representativa nacional de gran escala y, segundo, estudia el tema desde una perspectiva curricular, específicamente, de cómo afecta el dominio de contenidos curriculares previos en el desempeño curricular futuro en matemáticas. Específicamente, la primera pregunta de investigación es: ¿qué conocimientos previos en los ejes de la asignatura de matemáticas son más importantes para explicar el desempeño general futuro de los estudiantes en matemáticas?

Para poder entender bien el aporte específico de los conocimientos previos en los ejes de la asignatura de matemáticas, es necesario considerar también otras variables que son relevantes para explicar el desempeño en matemáticas de los estudiantes. Existe evidencia que muestra que el nivel socioeconómico, el dominio del lenguaje y el género son variables importantes para explicar el desempeño académico en general y las matemáticas en particular. Específicamente, en relación al nivel socioeconómico, se ha encontrado amplia evidencia de su importancia en el rendimiento en general (Bradley & Corwyn, 2002; Buchmann, 2002; Dahl & Lochner, 2005; OECD, 2013a, 2013b; Sirin, 2005; Sullivan, 2007) y específicamente en matemáticas (Jordan, Kaplan, Olah, & Locuniak, 2006; Sirin, 2005). Además, hay evidencia de la asociación entre el nivel socioeconómico inicial en el rendimiento posterior en matemáticas (Krajewski & Schneider, 2009).

En relación a la alfabetización lectora y escrita, se ha encontrado que su conocimiento es fundamental en otros desempeños académicos como las matemáticas y las ciencias (Mullis, Martin, & Foy, 2013; Norris & Phillips, 2003; Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008; Vilenius - Tuohimaa, Aunola, & Nurmi, 2008). Específicamente con respecto al rendimiento en matemáticas, la evidencia muestra que los estudiantes con peor desempeño en lectura demuestran consistentemente un peor desempeño en matemáticas (Mullis et al., 2013) y que la habilidad de comprensión lectora está fuertemente asociada al desempeño para resolver problemas matemáticos escritos (Vilenius - Tuohimaa et al., 2008). Además, con respecto al progreso, se ha encontrado una asociación entre las dificultades en lectura y un progreso lento en superar dificultades en matemáticas (Gersten et al., 2005).

Finalmente, también hay mucha investigación sobre las diferencias de género en el rendimiento en matemáticas, aunque estas varían entre países. En la prueba PISA 2012, en los países de la OECD los hombres obtienen en promedio 11 puntos más que las mujeres, sin embargo, solo en 37 de los 65 países los hombres tienen un mejor desempeño que las mujeres (OECD, 2013b). En la misma línea, (Else-Quest, Hyde, & Linn, 2010) analizan las diferencias de género en las pruebas TIMSS y PISA del año 2003 y encuentran que el tamaño promedio del efecto de género es muy bajo, aunque este varía bastante entre países. Por otro lado, se ha encontrado que desde edades tempranas, los hombres tienden a tener mejor desempeño que las mujeres y que estas diferencias pueden ser incluso mayores en los grupos de padres con alta educación (Penner & Paret, 2008). En Chile, los resultados de TIMSS 2003 y 2011 indican que, en promedio, los hombres tienen significativamente un mejor resultado que las mujeres (Mullis, Martin, Foy, & Arora, 2012).

Considerando la evidencia anterior, la segunda pregunta de investigación es ¿qué factores sociodemográficos y de desempeño académico son más importantes para determinar el desempeño curricular futuro de los estudiantes en matemáticas?

Responder a las dos preguntas de investigación busca entregar evidencia a la política pública sobre aquello que es más importante priorizar en el contexto de un currículum nacional de matemáticas, considerando las características sociodemográficas y de desempeño académico de los estudiantes.

## **2. Metodología**

El Sistema de Medición de la Calidad Educativa (SIMCE) en Chile surge en 1988 con el propósito de mejorar la calidad y equidad de la educación, mediante la provisión de información relevante para el quehacer de los distintos actores del sistema educativo. En particular, se comunican los logros de aprendizaje de los estudiantes en diferentes áreas de aprendizaje del currículum nacional, y se relacionan con el contexto escolar y social en el que estos aprenden (Agencia de la Calidad de la Educación, 2013).

Para responder a las preguntas de investigación se analizará la asociación entre un conjunto de variables individuales medidas el año 2007 cuando los estudiantes cursaban 4to año básico y su desempeño en la prueba SIMCE Matemáticas que rindieron el año 2011 cuando cursaban 8vo básico. Las co-variables individuales son: el desempeño de los estudiantes en los cuatro distintos ejes curriculares (Números, Álgebra, Formas y Espacio, Datos y Azar) evaluados en la prueba SIMCE Matemáticas de 4to básico, el desempeño en la prueba SIMCE Lenguaje de 4to básico, el grupo socioeconómico al que pertenecen los estudiantes y su género.

Los ejes curriculares definidos a evaluar en la prueba SIMCE de cuarto básico 2007 en la asignatura de matemáticas fueron: Números, Operaciones Aritméticas, y Formas y Espacio (descritos en Anexo 1). Sin embargo, el año 2010 se incluyen los ejes de Geometría y Datos y Azar, por lo que los ítems evaluados el año 2007 se reclasificaron según estas nuevas categorías. En consecuencia, los resultados reportados en este estudio incluyen el eje curricular de Datos y Azar.

Es importante mencionar que la prueba nacional SIMCE no ha sido construida con la finalidad de evaluar ni reportar estos cuatro ejes de forma separada, sino que busca entregar información integrada de los logros de aprendizaje alcanzados por los estudiantes en los distintos niveles de enseñanza (Agencia de la Calidad de la Educación, 2013) y representar esto a partir de un solo puntaje por alumno. Sin embargo, este estudio tiene un enfoque curricular que busca entender, a partir de los

contenidos evaluados en los diferentes ítems de la prueba, la importancia relativa de cada uno de los ejes curriculares.

## **2.1. Muestra**

La prueba SIMCE tiene un diseño rotado con 4 formas de prueba diferentes, donde cada estudiante rinde una de las cuatro formas. En consecuencia, la muestra efectiva de este estudio está constituida por cuatro submuestras de estudiantes que rindieron diferentes formas de la prueba SIMCE de Matemáticas el año 2007 (n1: 42.789, n2: 41.342, n3: 40.099, n4: 38.446). En total, la muestra efectiva corresponde a 162.676 estudiantes pertenecientes a 5.827 colegios, sobre un universo de 226.785 alumnos registrados para rendir la prueba de Matemáticas analizada (Agencia de la Calidad de la Educación, 2014). Solo se consideraron aquellos alumnos que habían rendido la prueba en 4° (2007) y en 8° (2011). Considerando la pérdida de casos, para asegurar representatividad nacional de la muestra en cuanto a cobertura país y type of school (public, subsidized or private), se calcularon pesos muestrales utilizando la misma metodología que la usada en una prueba nacional de carácter muestral para evaluar habilidades ICT, SIMCE TIC (Centro de Educación y Tecnología - ENLACES, 2011).

## **2.2. Modelo y definición de variables utilizadas**

Para responder a las preguntas de investigación se ajustan modelos de regresiones lineales (OLS) con cluster-robust standard errors using sampling weights, donde la variable dependiente corresponde al puntaje SIMCE Matemáticas obtenido el año 2011 cuando los estudiantes cursaban en 8vo básico. Este es un puntaje con rango entre 135,35 y 395,66 y una media 264,78 (en el total de la muestra). Esta prueba evalúa cuatro ejes curriculares de matemáticas: Números, Álgebra, Geometría y, Datos y azar.

Los modelos de regresión consideran los pesos muestrales calculados y el cálculo del error estándar se realiza por colegio (cluster). El cálculo convencional de los errores se vuelve más sesgado en la medida que aumenta el tamaño de la muestra. En cambio, el cálculo de cluster-robust standard errors es más conservador y considera la agrupación de estudiantes por colegio, en términos de que relaja la suposición de que los errores estándar son independientes entre sí y permite la correlación arbitraria entre los errores dentro de una escuela (Angrist & Pischke, 2008). Todos los análisis fueron realizados utilizando Stata 12.

Como se mencionó anteriormente, la prueba SIMCE se aplica en cuatro formas diferentes con un conjunto de ítems comunes a partir de los cuales se realiza un procedimiento de equating para obtener un sólo puntaje por estudiante comparable

entre formas a nivel de escuela. Sin embargo, el interés de este estudio está en el rendimiento de los estudiantes en cada eje curricular evaluado en la prueba, por lo que se generaron scores por subconjuntos de ítems de la prueba (para cada eje curricular), por lo que se debe ajustar un modelo de regresión lineal para cada submuestra que rindió cada una de las cuatro formas del test. Cabe destacar que realizar un procedimiento de equating fue descartado dado el bajo número de ítems comunes entre formas por cada eje.

Las variables independientes incluidas en el modelo son los scores de los cuatro ejes curriculares evaluados en la prueba SIMCE Matemáticas rendida por los estudiantes el año 2007 cuando cursaban 4to básico; Numbers, Algebra, Geometry, and Statistics and Probability; el puntaje SIMCE Lenguaje también rendido por los estudiantes el año 2007 cuando cursaban 4to básico, su género y grupo socioeconómico (SES). Cabe destacar que los puntajes de los cuatro ejes curriculares de Matemáticas evaluados el año 2007 se componen de diferentes cantidades de ítems, que además varían por forma de la prueba. Los estadísticos descriptivos se encuentran en el Anexo 2. A continuación se describe cada una de las covariables:

- a) **Números:** Corresponde al puntaje obtenido en las preguntas sobre el eje curricular de Números en el SIMCE Matemáticas en 4to básico aplicado el año 2007 (detalle en Anexo 1). Para obtener un puntaje sobre el desempeño de los estudiantes en este eje, se realizó un análisis factorial exploratorio a partir de correlaciones tetracóricas de todos los ítems binarios (entre 15 y 17 por forma) donde se estimó un puntaje a través del método Barlett (DiStefano et al., 2009) del primer componente (eigenvalue del primer factor [5,84 a 6,81]; Cronbach's alpha= [0,81 a 0,84]). Finalmente, para luego poder comparar los resultados, los puntajes se estandarizaron dividiendo la variable generada en cada forma por dos de sus desviaciones estándar ([0,324 a 0,338]). Este método se llevó a cabo independientemente para cada una de las cuatro sub muestras (según forma rendida) de la prueba aplicada durante el 2007.
- b) **Álgebra:** Corresponde al puntaje obtenido en las preguntas sobre el eje curricular de Álgebra en el SIMCE Matemáticas en 4to básico aplicado el año 2007 (detalle en Anexo 1). Se llevó a cabo la misma metodología que en a) para generar un puntaje a partir de los ítems (entre 4 y 5 por forma) (eigenvalue primer factor = [1,5 a 1,90]; Cronbach's alpha= [0,53 a 0,61]). Finalmente, se estandariza dividiendo por dos desviaciones estándar ([0,301 a 0,324]) en cada forma.
- c) **Formas y espacio:** Corresponde al puntaje obtenido en las preguntas sobre el eje curricular de Geometría en el SIMCE Matemáticas en 4to básico aplicado el año 2007 (detalle en Anexo 1). Se llevó a cabo la misma metodología que en a) para generar un puntaje a partir de los ítems (entre 9 y 11 por forma), (eigenvalue primer factor = [2,84 a 3,59]; Cronbach's alpha= [0,65 a 0,71]). Finalmente, se

- estandariza dividiendo por dos desviaciones estándar ([0,287 a 0,301]) en cada forma.
- d) Datos y azar: Corresponde al puntaje obtenido en las preguntas sobre el eje Datos y Azar en el SIMCE Matemáticas en 4to básico aplicado el año 2007 (detalle en Anexo 1). Se llevó a cabo la misma metodología que en a) para generar un puntaje a partir de los ítems (entre 4 y 6 por forma), (eigenvalue primer factor = [2,20 a 2,56]; Cronbach's alpha= [0,45 a 0,67]). Finalmente, se estandariza dividiendo por dos desviaciones estándar ([0,301, 0,324]) en cada forma.
- e) Puntaje SIMCE Lenguaje 4to básico (2007): Corresponde al puntaje obtenido en la prueba SIMCE Lenguaje en 4to básico aplicado el año 2007. Esta prueba evalúa comprensión lectora a través de tareas de localización, interpretación relación y reflexión de información en textos literarios y no literarios (Ministerio de Educación, 2007). El rango de puntajes está entre 112,32 y 379,35 con media 265,23 (calculados sobre el total de la muestra). Este puntaje se estandariza a partir de la división por dos desviaciones estándar de la variable (50,23).
- f) Género: autodeclaración obtenida a partir de los cuestionarios de contexto SIMCE. Corresponde a una variable binaria donde 1 representa femenino y 0 a masculino (agregar % de hombres y mujeres en la muestra efectiva).
- g) Grupo socioeconómico: corresponde a la clasificación por colegio realizada por el Ministerio de Educación y sus valores son nominales (A, B, C, D y E) donde A corresponde al grupo más bajo y E corresponde al más alto. Esta variable se construye en base al nivel educacional de la madre, nivel educacional del padre, ingreso mensual del hogar e índice de vulnerabilidad de la escuela (Ministerio de Educación, 2012).

It is worth mentioning that, in general, the scores estimated have acceptable to good internal consistency (referencia) with alphas varying from 0,62 to 0,84 (see Appendix 4). Nevertheless, in four cases these levels of internal consistency are not reached, which means that findings regarding Algebra in sub-samples J ( $\alpha=0.53$ ), L ( $\alpha=0.55$ ) and M ( $\alpha=0.55$ ) and in Probability and Statistics in sub-sample J (0.45) should be interpreted with caution.

Se revisaron posibles problemas de multicolinealidad entre las covariables mediante un análisis VIF que arrojó valores entre 1,04 y 3,21 (ver Anexo 3). Estos resultados permiten descartar problemas de multicolinealidad ya que son bajo los valores de referencia ( $VIF < 4$  o  $VIF < 10$ ) (Mandel, 1985; O'brien, 2007).

### 3. Resultados

En la **Tabla 4-6Resultados** modelos de regresión (OLS) por submuestra se presentan los resultados de los modelos de regresión ajustados para cada submuestra. El ajuste de

los modelos es adecuado con un  $R^2$  de 0.52, por lo que se puede decir que los modelos se ajustan bien a los datos (referencia).

**Tabla 0-1: Resultados modelos de regresión OLS por submuestra**

VARIABLES	Forma 4to J Puntaje Matemáticas 8vo	Forma 4to K Puntaje Matemáticas 8vo	Forma 4to L Puntaje Matemáticas 8vo	Forma 4to M Puntaje Matemáticas 8vo
Números*	26.55*** (2.286)	26.92*** (1.165)	29.01*** (1.339)	25.09*** (0.708)
Álgebra*	16.32*** (0.855)	17.10*** (1.152)	16.14*** (0.565)	15.56*** (0.971)
Formas y Espacio*	9.622*** (0.0769)	10.50*** (0.476)	10.19*** (0.871)	9.086** (1.776)
Datos y Azar*	1.784 (0.914)	1.418 (1.415)	5.410*** (0.235)	4.859*** (0.399)
Puntaje SIMCE Lenguaje 4to básico	15.95*** (1.421)	13.74*** (1.318)	12.00** (2.332)	15.85*** (0.900)
Género (mujer)	-5.100** (0.976)	-5.166** (0.634)	-3.242* (0.948)	-5.083** (0.815)
2.Grupo D (GSE)	-16.34*** (0.977)	-17.29*** (0.337)	-15.30** (1.937)	-18.05** (2.161)
3.Grupo C (GSE)	-27.17** (3.338)	-27.64** (3.715)	-22.60** (3.068)	-25.88** (4.226)
4.Grupo B (GSE)	-30.18*** (2.865)	-31.20*** (2.926)	-25.32** (2.741)	-29.88** (3.663)
5.Grupo A (GSE)	-28.33*** (1.679)	-28.66*** (0.788)	-23.91*** (0.405)	-28.65** (3.435)
Constante	166.5*** (8.176)	166.2*** (5.881)	164.7*** (8.506)	163.8*** (8.177)
Observaciones	42,789	41,342	40,099	38,446
R-cuadrado	0.520	0.518	0.523	0.521

\*Las variables indicadas se construyen a partir de distintas muestras correspondientes a las cuatro formas de la prueba.

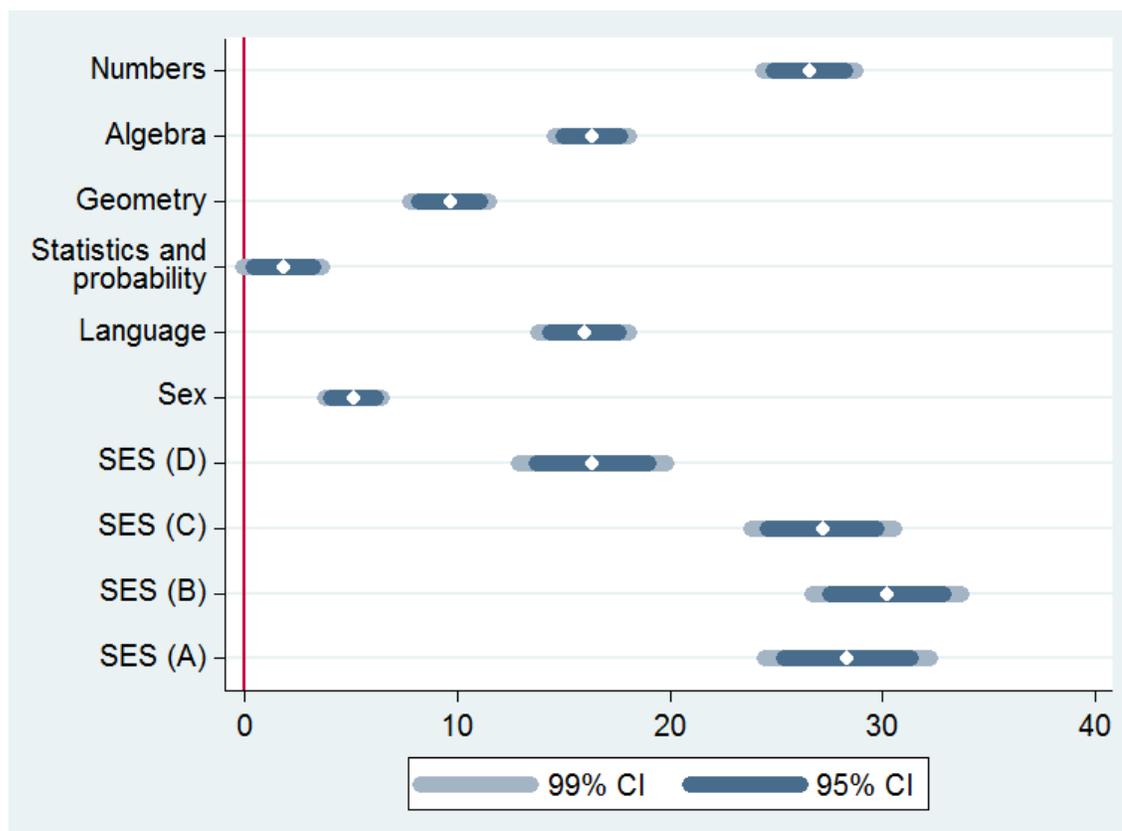
Robust standard errors in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$  two tailed test

Para la prueba de Matemáticas analizada los resultados indican que la gran mayoría de las variables incluidas en el modelo (números, álgebra, formas y espacio, puntaje en SIMCE Lenguaje, género y grupo socioeconómico) tienen un efecto estadísticamente significativo ( $p < 0.1$ ), y consistente entre formas. El score obtenido en el eje curricular

de Datos y Azar no tiene un efecto estadísticamente significativo en el rendimiento en la prueba de Matemáticas analizada en dos de las submuestras (forma J y K).

Gráfico 1: Coeficientes\* de regresión para sub-sample J (al 95% y 99% de confianza)



\*Todos los coeficientes se presentan en valor absoluto para facilitar la interpretación del gráfico.

En términos comparativos (ver gráfico 1 para sub-sample J) revisando los intervalos de confianza al 95% y mediante Test F (ver Anexo 4 y 5 para todas las sub-samples) se puede observar que entre las distintas variables de desempeño de los estudiantes en 4to básico, el efecto marginal del eje curricular de Números es el que aparece como más importante y se diferencia significativamente del efecto de los demás scores por eje y del rendimiento en lenguaje, en todas las submuestras. Sin embargo, no es posible afirmar que exista una diferencia significativa entre el efecto absoluto de números con respecto a los grupos socioeconómicos A, B ni C. En segundo lugar se encuentra que el puntaje en el eje de Álgebra y en la prueba SIMCE de Lenguaje, tienen una importancia similar. Específicamente, las diferencias entre ambos coeficientes son estadísticamente significativas en sólo dos de las submuestras (formas K y L) (intervalos de confianza al 95%, ver Anexo 2). En tercer lugar, el rendimiento en lenguaje aparece en tres de las submuestras, como más importante que el puntaje en el eje curricular de Formas y Espacio. Específicamente, se observan diferencias estadísticamente

significativas (al 95% de confianza) entre ambos coeficientes en todas las submuestras excepto en quienes rindieron la forma L (ver Anexo 4 y 5).

Resulta interesante que, controlando por las demás covariables, el efecto marginal del score en el eje de Números es similar en magnitud al efecto marginal de pertenecer a los grupos socioeconómicos A, B o C en comparación a pertenecer al grupo E (el más aventajado). Esto puede significar que el aumento de dos desviaciones estándar en el eje de Números permitiría “compensar”, en términos de puntaje Simce, el hecho de pertenecer a los grupos socioeconómicos más bajos.

Adicionalmente, se puede ver que controlando por las demás covariables, las mujeres obtienen en todas las submuestras un puntaje significativamente peor que los hombres (entre 3.2 y 5.2 puntos menos). El grupo socioeconómico se presenta en comparación al grupo E (más alto) y se puede observar como al disminuir el grupo al que pertenece el estudiante, el puntaje de Matemáticas va disminuyendo significativamente.

#### **4. Discusión y conclusiones**

Este artículo presenta un estudio que buscó responder dos preguntas de investigación. Primero, qué conocimientos previos en los ejes de la asignatura de matemáticas son más importantes para determinar el desempeño general futuro en este ámbito. Y segundo, cuáles son los factores sociodemográficos y de desempeño más importantes para determinar el desempeño curricular futuro de los estudiantes en matemáticas. El abordar el tema en una muestra representativa nacional y desde una perspectiva curricular, permite que las conclusiones obtenidas orienten el diseño de las políticas públicas en este ámbito.

En primer lugar, la importancia del eje curricular de números para el desempeño futuro en matemáticas, muestra que los resultados a nivel de desempeño curricular son consistentes con los resultados obtenidos desde la investigación cognitiva. En un contexto de baja implementación curricular, este resultado aporta evidencia a la política educativa sobre la importancia de trabajar las habilidades numéricas tempranas para un mejor desempeño general en matemática de los estudiantes.

En segundo lugar, la importancia de números es reforzada al considerar otro hallazgo importante de este estudio, en relación a que un buen manejo del eje de números permite en alguna medida compensar las desventajas relacionadas con el nivel socioeconómico de los estudiantes. Este resultado es particularmente relevante en un país como Chile, donde las diferencias socioeconómicas y culturales en los resultados

de aprendizaje de los estudiantes es una de las más grandes, tal como muestran los resultados de PISA (OECD, 2013a) .

En tercer lugar, los resultados muestran la importancia también de lenguaje en el desempeño general en matemáticas. Aún más, tener un buen desempeño en lenguaje resultó ser casi tan importante como manejar álgebra y más importante que manejar los ejes de formas y espacio, y datos y azar. Este resultado es consistente con otras investigaciones que muestran la importancia del lenguaje para el desempeño en matemáticas (Mullis et al., 2013; Norris & Phillips, 2003; Savolainen et al., 2008; Vilenius - Tuohimaa et al., 2008). Por otra parte, entrega indicios sobre la relevancia de afianzar las alfabetizaciones básicas como números, lectura y escritura para un buen desempeño académico en matemáticas. Considerando este resultado, sería relevante estudiar si este es el caso sólo para matemáticas o también para otras asignaturas como ciencias.

Finalmente, los resultados son consistentes con investigaciones previas que muestran diferencias de género en el desempeño de los estudiantes chilenos en matemáticas (Mullis et al., 2012; OECD, 2013b)(Mullis et al., 2012). Este resultado en conjunto con los demás hallazgos, señala la importancia de reforzar el trabajo en el eje de Números focalizándose especialmente en las estudiantes mujeres.

En síntesis, los resultados de este estudio entregan evidencia a la política educativa sobre la importancia de reforzar los contenidos curriculares de números y de lenguaje para el desempeño futuro en matemáticas, y en especial, de hacerlo con estudiantes mujeres y de nivel socioeconómico más bajo. Finalmente, para saber si estos resultados trascienden a la realidad chilena, es relevante poder estudiar otros países en investigaciones futuras.

## 5. Referencias

- Anderson, R. (2008). Implications of the information and knowledge society for education. In J. Voogt & G. Knezek (Eds.), *International handbook of information technology in primary and secondary education* (pp. 5-22). New York: Springer.
- Becker, G. S. (2009). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*: University of Chicago Press.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (Eds.). (2012). *Defining twenty-first century skills*: Springer.
- Boyd, W. (1966). *The History of Western Education*. New York: Barnes & Noble.

- Bradley, R. H., & Corwyn, R. F. (2002). SOCIOECONOMIC STATUS AND CHILD DEVELOPMENT. *Annual Review of Psychology, 53*(1).
- Buchmann, C. (2002). Measuring Family Background in International Studies of Education: Conceptual Issues and Methodological Challenges *Methodological Advances in Cross-National Surveys of Educational Achievement* (pp. 150-197). Washington, DC: The National Academies Press.
- Castells, M. (2002). *The Internet Galaxy: Reflections on the Internet, Business, and Society*. OUP Oxford.
- Centro de Educación y Tecnología - ENLACES. (2011). Documentación Técnica SIMCE TIC 2° medio 2011. Santiago de Chile: Ministerio de Educación.
- Chard, D., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). Using Measures of Number Sense to Screen for Difficulties in Mathematics: Preliminary Findings. *Assessment for Effective Intervention, 30*(2), 3-14.
- Dahl, G. B., & Lochner, L. (2005). The Impact of Family Income on Child Achievement: National Bureau of Economic Research.
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological Bulletin, 136*(1), 103-127.
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539-1552. doi: 10.1037/a0025510
- Gersten, R. J., Jordan, N. C., & Flojo, J. R. (2005). Early Identification and Interventions for Students With Mathematics Difficulties. *Journal of Learning Disabilities, 38*(4).
- Griffin, P. E., Care, E., & McGaw, E. (2012). The Changing Role of Education and Schools. . In P. E. Griffin, E. McGaw, & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 1-17). New York London: Springer.
- Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences, 20*(2), 82-88. doi: <http://dx.doi.org/10.1016/j.lindif.2009.07.004>
- Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number Sense Growth in Kindergarten: A Longitudinal Investigation of Children at Risk for Mathematics Difficulties. *Child Development, 77*(1), 153-175.
- Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850-867. doi: 10.1037/a0014939
- Krajewski, K., & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction, 19*(6), 513-526. doi: 10.1016/j.learninstruc.2008.10.002

- Ministerio de Educación. (2004). Implementación Curricular en el Aula Matemáticas Primer Ciclo Básico (NB1 y NB2): Unidad de Currículum y Evaluación.
- Ministerio de Educación. (2007). SIMCE. Orientaciones para la Medición. 4° Educación Básica 2007. Santiago, Chile.: Unidad de Currículum y Evaluación.
- Ministerio de Educación. (2012). Metodología de construcción de grupos socioeconómicos SIMCE 2011: Unidad de Currículum y Evaluación.
- Ministerio de Educación. (2013). Implementación del Currículum de Educación Media en Chile. *Serie de Evidencias, Año 2(21)*.
- Mullis, I. V. S., Martin, M. O., & Foy, P. (2013). The impact of reading ability on TIMSS Mathematics and Science achievement at fourth: an analysis by item reading demands grade. In M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS and PIRLS 2011: Relationships Among Reading, Mathematics, and Science Achievement at the Fourth Grade—Implications for Early Learning*. TIMSS and PIRLS International Study Center: Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*: International Association for the Evaluation of Educational Achievement.
- Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87(2)*, 224-240. doi: 10.1002/sce.10066
- OECD. (2010). The OECD Programme for the International Assessment of Adult Competencies (PIACC). Paris.
- OECD. (2013a). PISA 2012 results: Excellence through equity: Giving every student the chance to succeed (Vol. II). PISA: OECD Publishing.
- OECD. (2013b). PISA 2012 results: What students know and can do e Student performance in mathematics, reading and science (Vol. I). PISA: OECD Publishing.
- Passolunghi, M. C., & Lanfranchi, S. (2012). Domain-specific and domain-general precursors of mathematical achievement: A longitudinal study from kindergarten to first grade. *British Journal of Educational Psychology, 82(1)*, 42-63. doi: 10.1111/j.2044-8279.2011.02039.x
- Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research, 37(1)*, 239-253. doi: 10.1016/j.ssresearch.2007.06.012
- Ramirez, F. O., & Meyer, J. W. (1980). Comparative Education: The Social Construction of the Modern World System. *Annual Review of Sociology, 6*, 369-399.
- Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., & Holopainen, L. (2008). Reading comprehension, word reading and spelling as predictors of school achievement and choice of secondary education. *Learning and Instruction, 18(2)*, 201-210. doi: 10.1016/j.learninstruc.2007.09.017

- Schultz, T. W. (1961). Investment in Human Capital. *The American Economic Review*, 51(1), 1-17.
- Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417-453. doi: 10.3102/00346543075003417
- Sullivan, A. (2007). Cultural Capital, Cultural Knowledge and Ability. *Sociological Research Online*, 12(6). doi: 10.5153/sro.1596
- Träff, U. (2013). The contribution of general cognitive abilities and number abilities to different aspects of mathematics in children. *Journal of Experimental Child Psychology*, 116(2), 139-156. doi: <http://dx.doi.org/10.1016/j.jecp.2013.04.007>
- UNESCO. (2005). Education for All. The Quality Imperative. *EFA Global Monitoring Report*. Paris.
- Vilenius - Tuohimaa, P. M., Aunola, K., & Nurmi, J. E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology*, 28(4), 409-426. doi: 10.1080/01443410701708228

## ANEXO

### 2. Contenidos evaluados en SIMCE 4to Básico 2007

Se evalúan tres ejes temáticos a través de los cuales se evalúa de manera integrada la resolución de problemas. Los ejes son:

Números: En este eje se evaluará la capacidad de los estudiantes para **leer y escribir números** naturales y establecer relaciones entre estos, **ordenándolos**, buscando regularidades sencillas en secuencias o completando igualdades que tienen una incógnita en distintos lugares. También se contemplará la comprensión de **fracciones** a partir de su identificación en contextos gráficos o su relación con las partes de un todo. Finalmente, se evaluará el uso de los números para leer, interpretar y organizar información en tablas y gráficos. La resolución de problemas se evaluará en tareas relativas a la formación de números a partir de sus dígitos. En estos problemas, así como en la capacidad para descomponer y ordenar números se evaluará implícitamente la comprensión del **valor posicional de los dígitos que forman un número**.

Operaciones aritméticas: En este eje se evaluará la capacidad de los estudiantes para usar la suma, resta, multiplicación y división de números naturales en distintas situaciones, y el uso de algoritmos convencionales para realizar operaciones. Los estudiantes enfrentarán una gama amplia de problemas que requieren seleccionar datos y operaciones, realizar cálculos y verificar resultados. Al mismo tiempo enfrentarán problemas que demandan idear estrategias simples para su resolución.

Forma y espacio: En este eje se evaluará la orientación espacial, la capacidad para describir e interpretar trayectorias y ubicaciones en planos esquemáticos y la capacidad de relacionar el punto desde donde se observa un objeto con la representación gráfica de este. En este eje también se evaluará el reconocimiento, la comparación y la clasificación de figuras y cuerpos geométricos. La resolución de problemas se evaluará en tareas que requieren componer y descomponer formas geométricas de dos o tres dimensiones.

(Ministerio de Educación, 2007, pp. 14-15).

### 3. Tabla 0-2: Estadísticos descriptivos de las covariables estandarizadas por submuestra

	Forma J (n=42,789)	Forma K (n=41,342)	Forma L (n=40,099)	Forma M (n=38,446)
--	--------------------	--------------------	--------------------	--------------------

Variable	Alpha	Mean	Std. Dev.	Min	Max	Alpha	Mean	Std. Dev.	Min	Max	Alpha	Mean	Std. Dev.	Min	Max	Alpha	Mean	Std. Dev.	Min	Max
Números	0.82	1,5	0,5	0,0	2,1	0.81	1,6	0,5	0,0	2,3	0.84	1,5	0,5	0,0	2,1	0.82	1,6	0,5	0,0	2,1
Algebra	0.53	1,3	0,5	0,0	1,9	0.61	1,3	0,5	0,0	1,8	0.55	1,2	0,5	0,0	1,7	0.55	1,3	0,5	0,0	1,8
Geometry	0.65	1,7	0,5	0,0	2,3	0.69	1,8	0,5	0,0	2,4	0.71	1,8	0,5	0,0	2,4	0.67	1,9	0,5	0,0	2,5
Probability and Statistics	0.45	1,6	0,5	0,0	2,0	0.62	1,7	0,5	0,0	2,1	0.62	1,4	0,5	0,0	1,9	0.67	1,4	0,5	0,0	1,9
Lenguaje	-	2,7	0,5	1,1	3,8	-	2,6	0,5	1,2	3,8	-	2,6	0,5	1,2	3,7	-	2,6	0,5	1,2	3,7
GSE	-	3,1	1,1	1,0	5,0	-	3,1	1,1	1,0	5,0	-	3,1	1,1	1,0	5,0	-	3,1	1,1	1,0	5,0
Género	-	0,5	0,5	0,0	1,0	-	0,5	0,5	0,0	1,0	-	0,5	0,5	0,0	1,0	-	0,5	0,5	0,0	1,0

#### 4. Análisis VIF

	VIF			
	Forma J	Forma K	Forma L	Forma M
Numeros*	2.93	2.92	3.21	2.90
Algebra*	1.88	1.87	1.88	1.77
FyE*	2.02	2.15	2.17	2.15
DyA*	1.98	1.84	2.17	2.26
Lenguaje	2.29	2.23	2.29	2.29
GSE	1.04	1.05	1.05	1.05
Género	1.16	1.18	1.19	1.17
Mean VIF	1.90	1.89	2.00	1.94
Condition number	27.55	26.14	26.75	26.74
Det (correlation matrix)	0.0626	0.0643	0.0510	0.0560

\*Las variables se construyen a partir de distintas submuestras

## 5. Gráfico 2: Coeficientes\* de regresión para cada sub-sample (95% y 99% de confianza)



\*Todos los coeficientes se presentan en valor absoluto para facilitar la interpretación del gráfico.

## 6. Test F comparing coefficients

### a. Sub-sample J

F-ratio (Sub-sample J)					
	Alg.	FyE	DyA	Leng.	Fem.
Num.	63,64 (0,0000)	187,41 (0,0000)	398,93 (0,0000)	60,75 (0,0000)	1063,94 (0,0000)
Alg.		41,27 (0,0000)	218,72 (0,0000)	0,11 (0,7349)	668,7 (0,0000)
FyE			48,22 (0,0000)	27,57 (0,0000)	282,06 (0,0000)
DyA				133,44	59,84

		(0,0000)	(0,0000)
Leng.			423,86
			(0,0000)

Prob > F in parenthesis  
 between-group degrees of freedom = 1  
 within-group degrees of freedom = 6798

F-ratio (Sub-sample J)

	GSE B	GSE C	GSE D	GSE E
	146,68	434,19	490,36	350,59
GSE A	(0,0000)	(0,0000)	(0,0000)	(0,0000)
GSE B		168,52	247,23	120,25
		(0,0000)	(0,0000)	(0,0000)
GSE C			17,81	1,44
			(0,0000)	(0,2295)
GSE D				3,66
				(0,0557)

Prob > F in parenthesis  
 between-group degrees of freedom = 1  
 within-group degrees of freedom = 6798

F-ratio (Sub-sample J)

	Numeros
GSE A	965,29
	(0,0000)
GSE B	717,96
	(0,0000)
GSE C	1195,21
	(0,0000)
GSE D	1254,15
	(0,0000)
GSE E	1055,77
	(0,0000)

Prob > F in parenthesis  
 between-group degrees of freedom = 1  
 within-group degrees of freedom = 6798

**b. Sub-sample K**

F-ratio (Sub-sample K)

	Alg.	FyE	DyA	Leng.	Fem.
Num.	53,7	163,2	395,82	85,39	894,24
	(0,0000)	(0,0000)	(0,0000)	(0,0000)	(0,0000)
Alg.		32,53	203,66	8,23	629,99

	(0,0000)	(0,0000)	(0,0041)	(0,0000)
FyE		70,45	7,89	299,26
		(0,0000)	(0,0050)	(0,0000)
DyA			108,43	51,71
			(0,0000)	(0,0000)
Leng.				298,24
				(0,0000)

Prob > F in parenthesis  
 between-group degrees of freedom = 1  
 within-group degrees of freedom = 6451

F-ratio (Sub-sample K)				
	GSE B	GSE C	GSE D	GSE E
	107,7	294,43	354,93	254,36
GSE A	(0,0000)	(0,0000)	(0,0000)	(0,0000)
		135,22	217,47	94,49
GSE B		(0,0000)	(0,0000)	(0,0000)
			22,02	1,02
GSE C			(0,0000)	(0,3129)
				354,93
GSE D				(0,0134)

Prob > F in parenthesis  
 between-group degrees of freedom = 1  
 within-group degrees of freedom = 6451

F-ratio (Sub-sample K)	
	Numeros
GSE A	883,52 (0,0000)
GSE B	540,91 (0,0000)
GSE C	892,94 (0,0000)
GSE D	992,2 (0,0000)
GSE E	800,12 (0,0000)

Prob > F in parenthesis  
 between-group degrees of freedom = 1  
 within-group degrees of freedom = 6451

c. **Sub-sample L**

F-ratio (Sub-sample L)				
Alg.	FyE	DyA	Leng.	Fem.

Num.	100,12 (0,0000)	27,35 (0,0000)	97,4 (0,0000)	14,23 (0,0000)	956,35 (0,0000)
Alg.		27,35 (0,0000)	14,4 (0,0000)	2,1 (0,0002)	610 (0,0000)
FyE			14,4 (0,0001)	2,1 (0,1476)	171,42 (0,0000)
DyA				28,37 (0,0000)	70,02 (0,0000)
Leng.					232,97 (0,0000)

Prob > F in parenthesis  
between-group degrees of freedom = 1  
within-group degrees of freedom = 6190

F-ratio (Sub-sample L)				
	GSE B	GSE C	GSE D	GSE E
	56,94 (0,0000)	137,47 (0,0000)	164,16 (0,0000)	124,58 (0,0000)
GSE A		52,7 (0,0000)	96,38 (0,0000)	44,82 (0,0000)
GSE B			16,67 (0,0000)	1,55 (0,2128)
GSE C				1,85 (0,1738)
GSE D				

Prob > F in parenthesis  
between-group degrees of freedom = 1  
within-group degrees of freedom = 6190

F-ratio (Sub-sample L)	
	Numbers
GSE A	1044,05 (0,0000)
GSE B	446,62 (0,0000)
GSE C	682,94 (0,0000)
GSE D	720,1 (0,0000)
GSE E	591,56 (0,0000)

Prob > F in parenthesis  
between-group degrees of freedom = 1  
within-group degrees of freedom = 6190

## d. Sub-sample M

F-ratio (Sub-sample M)					
	Alg.	FyE	DyA	Leng.	Fem.
Num.	59,47 (0,0000)	139,93 (0,0000)	235,21 (0,0000)	50,78 (0,0000)	883,78 (0,0000)
Alg.		35,24 (0,0000)	96,62 (0,0000)	0,08 (0,7707)	650,1 (0,0000)
FyE			13,04 (0,0003)	34,11 (0,0000)	240,99 (0,0000)
Leng.					451,32 (0,0000)

Prob > F in parenthesis  
 between-group degrees of freedom = 1  
 within-group degrees of freedom = 5844

F-ratio (Sub-sample M)				
	GSE B	GSE C	GSE D	GSE E
GSE A	107,73 (0,0000)	241,41 (0,0000)	315,36 (0,0000)	250,63 (0,0000)
GSE B		81,59 (0,0000)	174,4 (0,0000)	90,1 (0,0000)
GSE C			35,81 (0,0000)	8,79 (0,0030)
GSE D				315,36 (0,1925)

Prob > F in parenthesis  
 between-group degrees of freedom = 1  
 within-group degrees of freedom = 5844

F-ratio (Sub-sample M)	
	Numeros
GSE A	816,22 (0,0000)
GSE B	430,2 (0,0000)
GSE C	644,2 (0,0000)
GSE D	748,16 (0,0000)
GSE E	650,17 (0,0000)

Prob > F in parenthesis  
between-group degrees of freedom = 1  
within-group degrees of freedom = 5844

## ANEXO J: CUARTO CRITERIO DE CLASIFICACIÓN (PAPER)

**Title: Curricular achievement under a sequential learning system**

**Authors:** Miguel Nussbaum<sup>1,2\*</sup>, Magdalena Claro<sup>1,2\*</sup>, Raimundo Frohlich<sup>1</sup>, Francisco Claro<sup>3</sup>, Tania Cabello<sup>2</sup>, Macarena Santana<sup>1</sup>, Ernesto San Martín<sup>2,3,4</sup>, Claudia Matus<sup>5</sup>.

### **Affiliations:**

<sup>1</sup> School of Engineering, Computer Science Department, Pontificia Universidad Católica de Chile, Santiago, Chile ().

<sup>2</sup> Center for the Study of Educational Policy and Practice (CEPPE), Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Campus San Joaquín, Macul, Santiago, Chile (\*).

<sup>3</sup> Faculty of Education, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile

<sup>4</sup> Faculty of Mathematics, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, Santiago, Chile

<sup>5</sup> Education Quality Assurance Agency, Government of Chile.

\*Correspondence to: [mn@ing.puc.cl](mailto:mn@ing.puc.cl), [magdalena.claro@gmail.com](mailto:magdalena.claro@gmail.com)

**Abstract:** Educational systems deal with delivering a standard curriculum to a heterogeneous student population following the principle of sequencing in learning. This paper shows how analyzing the relationship between students' curricular age (i.e. the knowledge and skills they have effectively acquired) and their actual grade level, reveals the extent to which education systems respect the curricular sequencing principle. It also highlights the consequences of failing to respect this principle, by analyzing Chile as a case study, where the majority of students never manage to catch up. Finally, an index is proposed to allow for objective analysis and comparisons between different education systems and groups.

### **One Sentence Summary:**

A method for analyzing if an educational system respects the sequencing learning principle was proposed, and applied to the Chilean case.

**Main Text:**

A country's school curriculum defines the knowledge, skills and values intended for its students (1,2). It does so by following the principles of sequencing and progression in learning (3,4). These principles dictate that every learning experience be based on the preceding ones (5-7). The sequences are set by experts and based on each subject's own logic to provide an appropriate order in which to present the concepts to learners (8). More recently, these sequences have been validated empirically using so-called learning progressions (9-11). A curriculum therefore provides schools and teachers with a pace and sequence for their students to follow, organized by age-grading (12,13). However, students vary in terms of cultural capital (14), prior knowledge (15), cognitive development (16), and individual traits (16,17). This leads students to tackle the curriculum with different styles and at different paces (16).

Educational systems deal with delivering a standard curriculum to a heterogeneous student population in different ways. There are two main models: the *separation model* and the *comprehensive school model* (18). In the former, students follow different paths after their primary education (19,20); in the latter they follow a single curriculum until aged 14 or 16 (19). Within the comprehensive system, Mons identifies three models for addressing learner differences (18). In the *à-la-carte integration model* adopted by countries like Canada, USA and Australia, students are grouped by ability using streaming and setting. In the *individualized integration model*, popular in Scandinavia, personalized teaching and support is offered to students so they can follow the curriculum at a similar pace. Finally, in the *uniform integration model* popular in Latin countries, students follow the same learning patterns as a group and the pace of learning is regulated by having failing students repeat the school year (21).

Questions therefore arise regarding the effectiveness of different education systems in following a standard curriculum. The present study analyses this question by considering effectiveness as a percentage of students whose curricular age (i.e. the knowledge and skills effectively acquired) coincides with their grade level, revealing the extent to which different systems respect the sequencing principle. From this perspective, the uniform integration model employed in Chile (19,21) was taken as a case study, where students are particularly heterogeneous, both academically and socio-culturally (22,23). The study therefore looks to answer the following research questions: Do curricular age and grade level coincide for students finishing 8<sup>th</sup> grade in Chile? What are the consequences of the two failing to coincide? Can the alignment of effective curricular age and grade level be compared objectively for different school administration types (e.g. public vs. private) or groups (e.g. gender, socioeconomic status)? To answer the first two questions, a new method of analysis is proposed. To answer the third, an index based on statistical dispersion is developed.

To deal with the first question on effective curricular age and grade level, data from the 2011 8<sup>th</sup> grade Chilean standardized test (SIMCE) for Mathematics and Language was analysed. This test is administered to every student in Chile<sup>4</sup> and measures fundamental knowledge and skills for 5<sup>th</sup> to 8<sup>th</sup> grade in Math and for 4<sup>th</sup> to 8<sup>th</sup> grade in Language (24). The final sample comprised 183,243 students who sat the 4<sup>th</sup> and 8<sup>th</sup> grade SIMCE in 2007 and 2011, respectively. The 4<sup>th</sup> grade data is needed to answer our second research question.

Curriculum experts classified the questions on the 8<sup>th</sup> grade SIMCE by grade level (5<sup>th</sup> to 8<sup>th</sup> grade). While questions could relate to knowledge and skills acquired in previous grades, the context in which they were applied was appropriate for 8<sup>th</sup> grade. Experts also judged that at least six questions were required in order to test a grade level. It was therefore necessary to combine grade levels with fewer than six questions on the test.

The *Satisfactory* level taken from the national learning standards (25,26) was used to determine the minimum requirement for passing each grade. This level is defined as having “acquired the basic knowledge and skills stipulated in the curriculum for the evaluated period”<sup>5</sup>.

Classifying students by curricular age was based on the understanding that a student who fails a level will also struggle with the following levels. This was true for 87.3% of students in Mathematics and 76.4% in Language.

Given that not all students met the conditions for sequencing, the following rules were set to determine a student’s curricular age (5<sup>th</sup> to 8<sup>th</sup> grade):

1. If a student failed two or more grades, they were considered to have failed from the first grade in which they did not meet the necessary requirements.
2. If a student passed two or more consecutive grades, they were considered to have reached the last grade they passed.

Figure 1 shows the distribution of curricular age for Mathematics (Figure 1a) and Language (Figure 1b). Only 1 in 4 students in Chile were at the expected curricular age in Mathematics, and less than half in Language. Furthermore, 64% of students in Mathematics and 39% of students in Language did not pass any of the grades and therefore had an effective curricular age of lower than 5<sup>th</sup> grade and 4<sup>th</sup>-5<sup>th</sup> grade, respectively. Grades were categorized differently for Language because of the grouping that was required in order to meet the six-question minimum described previously.

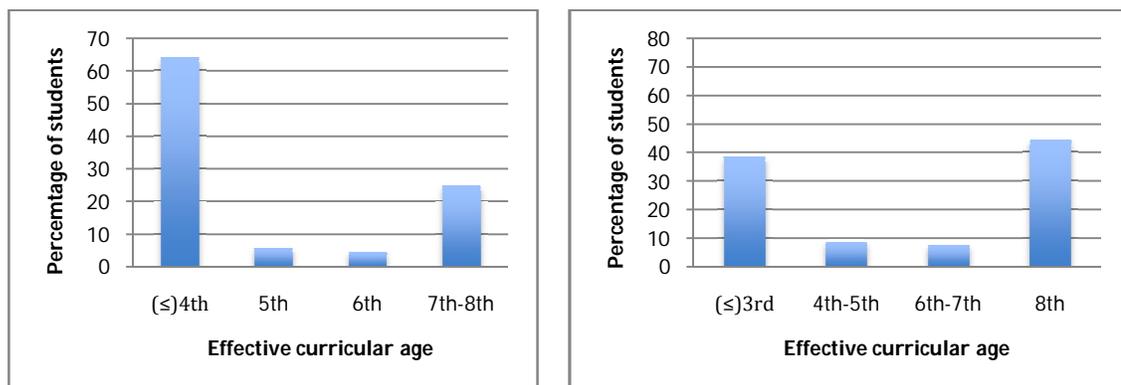
**Figure 1 Distribution of effective curricular age on the 2011 8th grade Mathematics (1a) and Language SIMCE (1b)**

**Figure 1a**

**Figure 1b**

<sup>4</sup><http://www.agenciaeducacion.cl/simce/que-es-el-simce/>

<sup>5</sup>Learning Standards for 8<sup>th</sup> grade – Decree 129\_2013



To answer the second question regarding the consequences of curricular age and grade level not coinciding, the students' 4<sup>th</sup> and 8<sup>th</sup> grade SIMCE scores were compared.

Table 1 shows the pass-fail rate for students on the 8<sup>th</sup> grade SIMCE versus their performance on the 4<sup>th</sup> grade test. In this case, a pass was considered as having met the *Satisfactory* level defined previously. For most students, failing to acquire the necessary knowledge at 4<sup>th</sup> grade hindered their future performance (66.3% for Mathematics and 46.5% for Language). As with previous studies (27), these results show that failing to respect the sequencing principle leads students to fall behind, with only the minority managing to catch up.

**Table 1: Pass rates for the 8<sup>th</sup> grade Mathematics and Language SIMCE based on 4<sup>th</sup> grade performance**  
Table 1

	Mathematics %	Language %
4th PASS-8th PASS	15.7	27.7
4th FAIL-8th PASS	9.4	17.1
4th PASS-8th FAIL	8.7	8.7
4th FAIL-8th FAIL	66.3	46.5

To answer the third question, whether it is possible to objectively compare the alignment of curricular age and grade level across education systems, an index based on the sequencing principle was created. We will call it the Curriculum Achievement Index (CAI), described as follows.

Consider a group of  $N$  children being tested for curricular achievement. Suppose they are grouped in  $S$  sequential segments according to the curriculum level they have reached, so that:

$$\begin{array}{lll}
 N_1 & \text{reached level} & 1 \\
 N_2 & \text{"} & 2 \\
 N_3 & \text{"} & 3
 \end{array}$$

$$\begin{array}{ccc} \cdot & & \cdot \\ \cdot & \text{“} & \cdot \\ N_S & & S \end{array}$$

The average achievement level of the group is then

$$\langle l \rangle = \frac{1}{N} \{N_1 + 2N_2 + 3N_3 \dots + SN_S\} \quad (1)$$

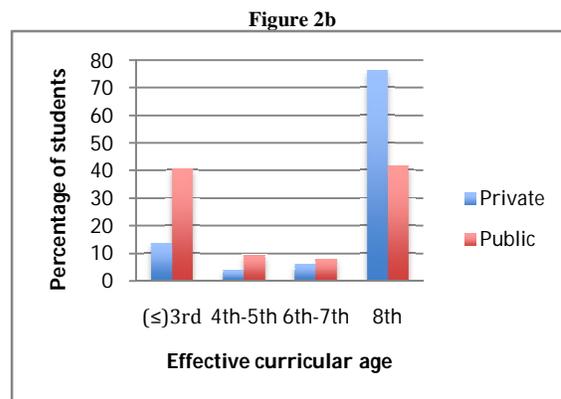
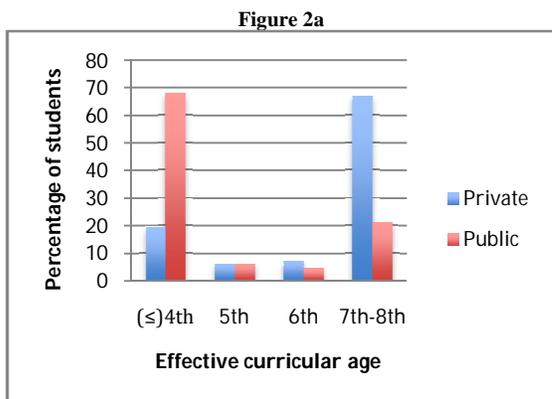
where  $N = N_1 + N_2 + \dots + N_S$ . Notice that in the best case, where everyone in the group has reached the highest level,  $N_1 = N_2 = \dots = N_{S-1} = 0$ ,  $N_S = N$  and  $\langle l \rangle = S$ , while in the worst case, where the group has not progressed beyond the lowest level,  $N_1 = N$ ,  $N_2 = \dots = N_S = 0$  and  $\langle l \rangle = 1$ . We define the Curriculum Achievement Index as the normalized average level

$$CAI = \frac{\langle l \rangle}{S} \quad , \quad (2)$$

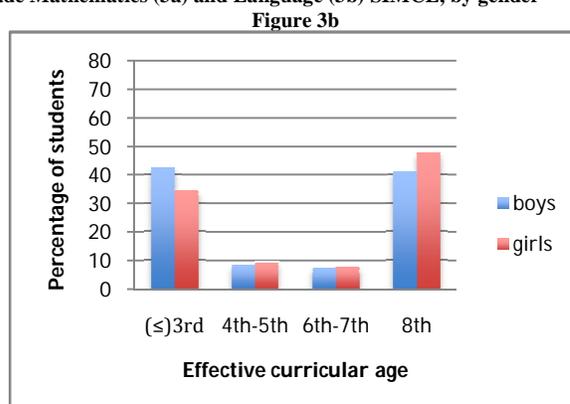
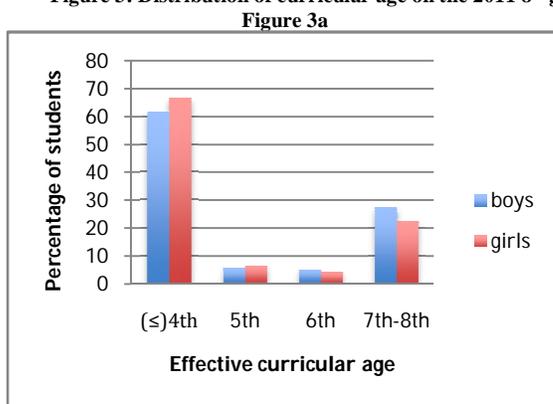
a number that conveniently covers the range  $1/S \leq CAI \leq 1$ , so that  $(CAI)_{best} = 1$  while  $(CAI)_{worst} = 1/S$ . The advantage of this measurement is that the optimal figure has a universal character and is not dependent on the number of levels involved, as is the simple average level defined in Eq. (1).

As an example, distributions for the 8<sup>th</sup> grade Mathematics and Language SIMCE scores were analysed and their respective CAI calculated. Figures 2 compares the distributions according to school type (publicly or privately funded), Figure 3 compares students' gender and Figure 4 compares families' socioeconomic status. Figures labelled (a) correspond to Mathematics and those labelled (b), to Language. Table 2 shows the CAI for each of these cases, where  $S = 4$ , with a minimum CAI of 0.25 and maximum of 1.

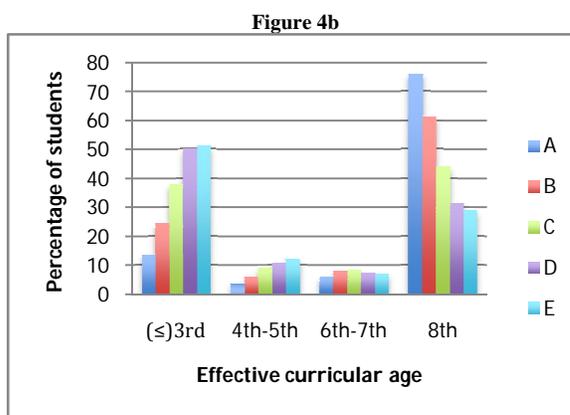
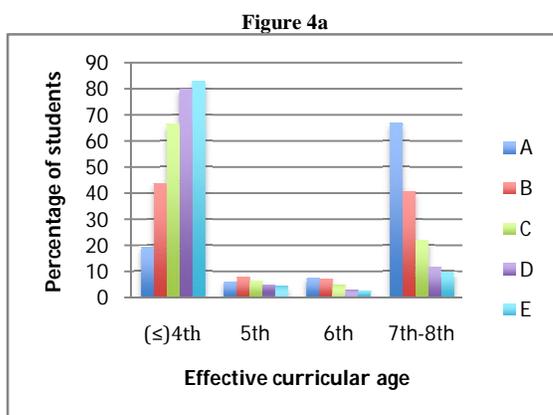
**Figure 2: Distribution of curricular age on the 2011 8<sup>th</sup> grade Mathematics (2a) and Language (2b) SIMCE, by school type**



**Figure 3: Distribution of curricular age on the 2011 8<sup>th</sup> grade Mathematics (3a) and Language (3b) SIMCE, by gender**



**Figure 4: Distribution of curricular age on the 2011 8<sup>th</sup> grade Mathematics (4a) and Language (4b) SIMCE, by socioeconomic status. A: First quintile, B: Second quintile, C: Third quintile, D: Fourth quintile, E: Fifth quintile**



**Table 2: Curricular Achievement Index by school type, gender and socioeconomic status on the 2011 Mathematics and Language SIMCE test**

	National		School Type				Gender				Socioeconomic Status			
			Public		Private		Female		Male		First quintile		Fifth quintile	
2011 SIMCE	Math	Lang	Math	Lang	Math	Lang	Math	Lang	Math	Lang	Math	Lang	Math	Lang
CAI	0.48	0.65	0.45	0.63	0.81	0.86	0.46	0.67	0.5	0.62	0.81	0.86	0.35	0.53

Chilean system demonstrates significant differences in curricular achievement between students attending publicly and privately funded schools, as well as belonging to different socioeconomic groups. Although the difference between genders is not as great, girls tend to outperform boys in Language, with the opposite being true in Mathematics. This is also the case in other countries (23). The CAI also allows comparisons to be made on a national level. For example, results are better in Language than Mathematics and the difference between the public and private systems is greater in Mathematics than in Language (Table 2).

Analysing the results of an education system based on the alignment of curricular age and grade level indicates how effectively the system delivers a standard curriculum by respecting the sequencing principle. In Chile, the results show that the education system tends to pass students without ensuring they have acquired the necessary knowledge and skills. The comparison between groups and school types using the proposed index also reveals that the public system (attended by 92.4% of students) is outperformed by the private system, with the same being true for a comparison between the lowest and highest socioeconomic groups. This is partly because the system does not have mechanisms for dealing with student heterogeneity, unlike countries such as Finland (28,29) or Norway (30). This observation is particularly relevant, as there are large differences in cultural capital when students enter first grade and which are reinforced over time by the education system (31). It is important that educational policies address this problem.

This study introduces a new assessment index (CAI) and proposes its universal use. This index is based on national definitions of curricular sequencing and represents achievement as a measurement of statistical dispersion. It therefore allows results to be compared between different segments within a sample, or different education systems and groups, regardless of their particular objectives. As defined previously, the index is a number between  $1/S$  and 1, where  $S$  is the number of grade levels being considered. The ideal achievement corresponds universally to unity, while the lower bound could be standardized by setting  $S=8$ , for instance, and distributing the available data proportionally to cover the whole range.

Finally, it is important to highlight that this study used standardized tests, which were not intended to measure curricular achievement by grade level or sequence. In this sense, specific tests for measuring sequencing achievements would provide more accurate information.

#### References and Notes:

1. *Education for all. The quality imperative* (UNESCO, Paris, France, 2004).
2. J. Boli, F.O. Ramirez, W. J. Meyer, Explaining the origins and expansion of mass education. *Comparative Education Review*, **29** (2), 145-170 (1985).
3. M. Hughes, *Progression in Learning* (Multilingual Matters, England, 1996).
4. T. Prieto, A. Blanco, V. Brero, La progresión en el aprendizaje de dominios específicos: una propuesta para la investigación. *Enseñanza de las Ciencias*.**20**(1), 3-14 (2002).
5. D.B. Tyack, *The one best system: a history of American urban education* (Harvard Univ. Press, Cambridge, Mass., 1974).
6. R.W. Tyler, *Basic principles of curriculum and instruction* (University of Chicago Press, Chicago, London, 1971).
7. R.W. Tyler, The curriculum- then and now. *The Elementary School Journal*, **57**(7) (1973).
8. V. Talanquer, Progresiones de aprendizaje: promesa y potencial. *Educ. quím.*, **24**(4), 362-364 (2013).
9. R. Duschl, S. Maenga, A. Sezenb, Learning progressions and teaching sequences: a review and analysis. *Studies in Science Education*. **47** (2), 123–182 (2011).
10. T. Corcoran, F.A. Mosher, A. Rogat, Learning progressions in science. An evidence-based approach to reform. CPRE Research Report # RR-63 (2009).
11. D. Fortus, J. Krajcik, in *Second International Handbook of Science Education*, B.J. Fraser, K.G. Tobin, C.J. McRobbie, Eds. (Springer International Handbooks of Education, 2012), chap. 52.
12. L. Cuban, *A successful reform: The age-graded school*. (2012). Retrieved february 2013 from: [http://larrycuban.wordpress.com/2012/02/01/a-successful-school-reform-the-age-graded-school-part-1-2/#\\_ednref](http://larrycuban.wordpress.com/2012/02/01/a-successful-school-reform-the-age-graded-school-part-1-2/#_ednref)
13. D. Tyack, W. Tobin, The “grammar” of schooling: Why has it been so hard to change? *American Educational Research Journal*, **31**(3), 453-47 (1994).
14. P. Bourdieu, J-C, Passeron, *Reproduction in education, society and culture* (Sage Publications, London, 1977).
15. J. Bransford, National Research Council (U.S.), Committee on Developments in the Science of Learning, National Research Council (U.S.) & Committee on Learning Research and Educational Practice. *How people learn: brain, mind, experience, and school* (National Academy Press: Washington, DC, 2000).
16. J.S. Bruner, *Toward a theory of instruction* (Belknap Press of Harvard Univ., Cambridge, Mass, 1966).

17. B.S Bloom, *Human characteristics and school learning* (McGraw-Hill, New York, 1976).
18. N. Mons, *Les nouvelles politiques éducatives: La France fait-elle les bons choix?* (Presses Universitaires de France, Paris, 2007).
19. V. Dupriez, X. Dumay, A. Vause, How do school systems manage pupils' heterogeneity? *Comparative Education Review*, **52**(2), 245-273 (2008).
20. M. Goos, B.M Schreier, H.M.E Knipprath, B.D Fraine, J.V Damme, U. Trautwein, How can cross-country differences in the practice of grade retention be explained? A closer look at national educational policy factors. *Comparative Education Review*, **57**(1), 54-84 (2013).
21. V. Dupriez, *Methods for grouping learners at school. Fundamentals of educational planning*. N°93 (IIEP/UNESCO, Paris, 2010).
22. *Equity and quality in education: supporting disadvantaged students and schools*, (OECD, OECD Publishing, 2012).
23. *PISA 2012 results: excellence through equity: giving every student the chance to succeed (Volume II)* (OECD, OECD Publishing, 2013).
24. "Informe de resultados nacionales 2011" (Ministerio de Educación-Unidad de Currículum y Evaluación-SIMCE, Santiago, 2011).
25. "Estándares de Aprendizaje, Lectura, 8vo Básico" (Ministerio de Educación, Unidad de Currículum y Evaluación, Santiago 2013a).
26. "Estándares de Aprendizaje, Matemática, 8vo Básico" (Ministerio de Educación, Unidad de Currículum y Evaluación, Santiago 2013b).
27. L. Pritchett, A. Beatty, The negative consequences of overambitious curricula in developing countries (Faculty Research Working Paper Series, Harvard Kennedy School, 2012; <http://ssrn.com/abstract=2235869>).
28. I. Halinen, J. Ritva, Towards inclusive education: the case of Finland. *Prospects*, **38**(1), 77-97 (2008).
29. P. Sahlberg, A model lesson. Finland shows us what equal opportunity looks like. *American Educator* (Spring 2012).
30. D. Nusche, L.Earl, W.Maxwell, C. Shewbridg, *OECD Reviewsof Evaluation and Assessment in EducationNorway*. (OECD, Paris, 2011).
31. R. Rosas, C. Santa Cruz, *Dime en qué colegio estudiaste y te diré qué CI tienes. Radiografía al desigual acceso al capital cognitivo en Chile* (Ediciones Universidad Católica de Chile, Santiago, 2013).

**Acknowledgments:** Research supported by the Center for the Study of Educational Policy and Practice (CEPPE-UC), Project CIE01-CONICYT

For more information about Chilean standardized mathematics and language tests, see: <http://www.agenciaeducacion.cl/simce/>

### ANEXO K: DESCRIPCIÓN SUBSCORES (VARIABLES INDEPENDIENTES MODELO)

- Variables no estandarizadas

```
. summ tnumerosJ talgebraJ tFyEJ tdatosyazarJ ptje_lect_4to female if
idforma_4to=="J"
```

Variable	Obs	Mean	Std. Dev.	Min	Max
tnumerosJ	42789	1.008751	.3381758	0	1.450151
talgebraJ	42789	.8081332	.3195052	0	1.185423
tFyEJ	42789	1.014906	.3018661	0	1.3844
tdatosyazarJ	42789	.9031978	.2901641	0	1.135659
ptje_lect~to	42789	267.4974	49.26181	112.32	372.78
female	42789	.520975	.4995657	0	1

```
. summ tnumerosK talgebraK tFyEK tdatosyazarK ptje_lect_4to female if
idforma_4to=="K"
```

Variable	Obs	Mean	Std. Dev.	Min	Max
tnumerok	41342	1.061513	.3248705	0	1.465988
talgebraK	41342	.8200041	.3247271	0	1.185963
tFyEK	41342	1.08723	.2960705	0	1.439871
tdatosyazarK	41342	.9298162	.2663393	0	1.126967
ptje_lect~to	41342	264.9758	50.17964	119.21	376.96
female	41342	.5203183	.499593	0	1

```
. summ tnumerok talgebraK tFyEK tdatosyazarK ptje_lect_4to female if
idforma_4to=="L"
```

Variable	Obs	Mean	Std. Dev.	Min	Max
tnumerokL	40099	.9907462	.3365077	0	1.421396
talgebraL	40099	.7614254	.3186611	0	1.100128
tFyEL	40099	1.061845	.2891251	0	1.401744
tdatosyazarL	40099	.8552217	.2998558	0	1.160107
ptje_lect~to	40099	262.6463	50.62286	118	376.96
female	40099	.5214344	.4995466	0	1

```
. summ tnumerokM talgebraM tFyEM tdatosyazarM ptje_lect_4to female if
idforma_4to=="M"
```

Variable	Obs	Mean	Std. Dev.	Min	Max
tnumerokM	38446	1.021474	.3291201	0	1.407477
talgebraM	38446	.8032501	.3017359	0	1.101757
tFyEM	38446	1.090304	.2871208	0	1.422475
tdatosyazarM	38446	.8743477	.3167485	0	1.196595
ptje_lect~to	38446	265.6775	50.81222	123.11	379.35
female	38446	.5178692	.4996871	0	1

- Variables estandarizadas

```
. summ z2_tnumerosJ z2_talgebraJ z2_tFyEJ z2_tdatosyazarJ z2_ptje_lect_4to
      GSE2_4to female if idforma_4to=="J"
```

Variable	Obs	Mean	Std. Dev.	Min	Max
z2_tnumerosJ	42789	1.49146	.5	0	2.144078
z2_talgebraJ	42789	1.264664	.5	0	1.855092
z2_tFyEJ	42789	1.681054	.5	0	2.293069
z2_tdatosy~J	42789	1.556357	.5	0	1.956925
z2_ptje_le~o	42789	2.715059	.5	1.140031	3.783661
GSE2_4to	42789	3.101685	1.087726	1	5
female	42789	.520975	.4995657	0	1

```
. summ z2_tnumerosK z2_talgebraK z2_tFyEK z2_tdatosyazarK z2_ptje_lect_4to
      GSE2_4to female if idforma_4to=="K"
```

Variable	Obs	Mean	Std. Dev.	Min	Max
z2_tnumerosK	41342	1.633748	.5000001	0	2.256266
z2_talgebraK	41342	1.262605	.5	0	1.826092
z2_tFyEK	41342	1.836101	.5	0	2.431635
z2_tdatosy~K	41342	1.745548	.5	0	2.11566
z2_ptje_le~o	41342	2.640272	.5	1.187832	3.756105
GSE2_4to	41342	3.085192	1.074587	1	5
female	41342	.5203183	.499593	0	1

```
. summ z2_tnumerosL z2_talgebraL z2_tFyEL z2_tdatosyazarL z2_ptje_lect_4to
      GSE2_4to female if idforma_4to=="L"
```

Variable	Obs	Mean	Std. Dev.	Min	Max
z2_tnumerosL	40099	1.4721	.5	0	2.111981
z2_talgebraL	40099	1.194726	.5	0	1.726173
z2_tFyEL	40099	1.836308	.5000001	0	2.424113
z2_tdatosy~L	40099	1.426055	.5	0	1.934442
z2_ptje_le~o	40099	2.594147	.5	1.165481	3.723219
GSE2_4to	40099	3.071099	1.061953	1	5
female	40099	.5214344	.4995466	0	1

```
. summ z2_tnumerosM z2_talgebraM z2_tFyEM z2_tdatosyazarM z2_ptje_lect_4to
      GSE2_4to female if idforma_4to=="M"
```

Variable	Obs	Mean	Std. Dev.	Min	Max
z2_tnumerosM	38446	1.551826	.5000001	0	2.138242
z2_talgebraM	38446	1.331048	.5000001	0	1.825697
z2_tFyEM	38446	1.898686	.5	0	2.477137
z2_tdatosy~M	38446	1.380192	.5	0	1.888873
z2_ptje_le~o	38446	2.614307	.5	1.211421	3.732862
GSE2_4to	38446	3.057379	1.054398	1	5
female	38446	.5178692	.4996871	0	1

**ANEXO L: RESULTADOS MODELO SIN PESOS Y ERROR ESTANDAR TRADICIONAL**

- Resultado modelo sin pesos y errores estándar por colegio

VARIABLES	Forma 4to J ptje_mat_8vo	Forma 4to K ptje_mat_8vo	Forma 4to L ptje_mat_8vo	Forma 4to M ptje_mat_8vo
z_tnumeros*	27.68*** (2.281)	27.44*** (2.040)	28.82*** (0.562)	26.65*** (1.144)
z_talgebra*	17.42*** (0.607)	17.42*** (1.318)	16.98*** (0.522)	15.77*** (0.878)
z_tFyE*	9.600*** (0.578)	11.76*** (0.789)	11.19*** (0.649)	10.09** (1.433)
z_tdatosyazar*	0.679 (1.432)	-0.372 (1.839)	5.522*** (0.535)	4.138** (0.775)
z_ptje_lect_4to	16.89*** (0.883)	15.46*** (0.877)	11.66** (1.692)	16.67*** (1.326)
1.female	-4.882*** (0.153)	-5.554** (0.616)	-3.865*** (0.222)	-5.329*** (0.347)
2.GSE2_4to	-18.37** (1.938)	-17.69*** (1.726)	-17.87** (2.352)	-19.16** (2.719)
3.GSE2_4to	-28.70** (4.742)	-28.00** (4.970)	-26.65** (4.406)	-27.93** (5.794)
4.GSE2_4to	-33.11** (3.848)	-32.36** (4.866)	-30.35** (4.438)	-32.42** (5.080)
5.GSE2_4to	-31.80*** (3.147)	-30.27** (3.098)	-29.57** (3.037)	-32.24** (4.594)
Constant	166.7*** (6.897)	164.0*** (7.646)	169.2*** (8.194)	162.3*** (9.231)
Observations	42,789	41,342	40,099	38,446
R-squared	0.535	0.537	0.547	0.538

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

- Resultado modelo sin pesos y errores estándar robustos

VARIABLES	Forma 4to J ptje_mat_8vo	Forma 4to K ptje_mat_8vo	Forma 4to L ptje_mat_8vo	Forma 4to M ptje_mat_8vo
z2_tnumeros*	26.55***	26.92***	29.01***	25.09***

	(0.837)	(0.895)	(0.895)	(0.873 )
z2_talgebra*	16.32***	17.10***	16.14***	15.56***
	(0.681)	(0.779)	(0.689)	(0.666)
z2_tFyE*	9.622***	10.50***	10.19***	9.086***
	(0.727)	(0.707)	(0.805)	(0.758)
z2_tdatosyazar*	1.784**	1.418**	5.410***	4.859***
	(0.725)	(0.709)	(0.787)	(0.782)
z2_ptje_lect_4to	15.95***	13.74***	12.00***	15.85***
	(0.802)	(0.798)	(0.794)	(0.740)
1.female	-5.100***	-5.166***	-3.242***	-5.083***
	(0.526)	(0.576)	(0.582)	(0.523)
2.GSE2_4to	-16.34***	-17.29***	-15.30***	-18.05***
	(1.173)	(1.515)	(1.905)	(1.610)
3.GSE2_4to	-27.17***	-27.64***	-22.60***	-25.88***
	(1.161)	(1.486)	(1.813)	(1.545)
4.GSE2_4to	-30.18***	-31.20***	-25.32***	-29.88***
	(1.227)	(1.543)	(1.871)	(1.564)
5.GSE2_4to	-28.33***	-28.66***	-23.91***	-28.65***
	(1.391)	(1.691)	(2.034)	(1.701)
Constant	166.5***	166.2***	164.7***	163.8***
	(2.300)	(2.241)	(2.870)	(2.264)
Observations	42,789	41,342	40,099	38,446
R-squared	0.520	0.518	0.523	0.521

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

- Resultado modelo sin pesos y errores estándar individuales.

VARIABLES	Forma 4to J ptje_mat_8vo	Forma 4to K ptje_mat_8vo	Forma 4to L ptje_mat_8vo	Forma 4to M ptje_mat_8vo
z_tnumeros*	27.68***	27.44***	28.82***	40.49***
	(0.552)	(0.554)	(0.586)	(0.872)
z_talgebra*	17.42***	17.42***	16.98***	26.13***
	(0.442)	(0.443)	(0.449)	(0.742)
z_tFyE*	9.600***	11.76***	11.19***	17.57***
	(0.458)	(0.476)	(0.482)	(0.861)
z_tdatosyazar*	0.679	-0.372	5.522***	6.532***
	(0.453)	(0.440)	(0.482)	(0.800)
z_ptje_lect_4to	16.89***	15.46***	11.66***	0.164***
	(0.489)	(0.485)	(0.497)	(0.00504)
1.female	-4.882***	-5.554***	-3.865***	-5.329***
	(0.330)	(0.332)	(0.336)	(0.346)

2.GSE2_4to	-18.37*** (0.659)	-17.69*** (0.664)	-17.87*** (0.670)	-19.16*** (0.687)
3.GSE2_4to	-28.70*** (0.618)	-28.00*** (0.624)	-26.65*** (0.632)	-27.93*** (0.645)
4.GSE2_4to	-33.11*** (0.649)	-32.36*** (0.657)	-30.35*** (0.667)	-32.42*** (0.680)
5.GSE2_4to	-31.80*** (0.787)	-30.27*** (0.812)	-29.57*** (0.844)	-32.24*** (0.880)
Constant	166.7*** (1.206)	164.0*** (1.197)	169.2*** (1.223)	162.3*** (1.258)
Observations	42,789	41,342	40,099	38,446
R-squared	0.535	0.537	0.547	0.538

---

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1