

# Relación inconsistente entre puntajes del Sistema de Medición de Calidad de la Educación y el promedio de notas en estudiantes de octavo básico en Chile

Vergara Díaz, Gastón; Peredo López, Hernán

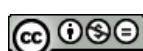
Relación inconsistente entre puntajes del Sistema de Medición de Calidad de la Educación y el promedio de notas en estudiantes de octavo básico en Chile

Revista Educación, vol. 44, núm. 1, 2020

Universidad de Costa Rica, Costa Rica

Disponible en: <https://www.redalyc.org/articulo.oa?id=44060092027>

DOI: <https://doi.org/10.15517/revedu.v44i1.37569>



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivar 3.0 Internacional.

## Relación inconsistente entre puntajes del Sistema de Medición de Calidad de la Educación y el promedio de notas en estudiantes de octavo básico en Chile

The Inconsistent Relationship between Education Quality Measurement System Scores and Grade Point Averages of 8th Grade Students in Chile

*Gastón Vergara Díaz*

*Instituto de Estadística, Universidad Austral de Chile,  
Chile*

*gastonvergara@uach.cl*

 <http://orcid.org/0000-0003-0109-7214>

DOI: <https://doi.org/10.15517/revedu.v44i1.37569>

Redalyc: <https://www.redalyc.org/articulo.oa?id=44060092027>

*Hernán Peredo López*

*Universidad de La Frontera, Chile*

*hperedolopez@gmail.com*

 <http://orcid.org/0000-0001-8850-8076>

Recepción: 18 Junio 2019

Aprobación: 25 Noviembre 2019

### RESUMEN:

El Sistema de Medición de Calidad de la Educación [SIMCE] tiene entre sus propósitos institucionales contribuir al mejoramiento de la calidad de la educación e informar a la ciudadanía de los logros de aprendizaje del estudiantado chileno en diferentes áreas, pero existen pocas investigaciones concretas en este último ámbito. El objetivo principal de la investigación es establecer la correlación entre las notas del estudiantado al finalizar el 8º básico y los puntajes SIMCE en las pruebas de Lenguaje y Comunicación, Matemáticas y Ciencias Naturales. Se utilizaron los datos de octavo año básico, proporcionados por la Agencia de Calidad de la Educación, perteneciente al Ministerio de Educación de Chile. La información original incluyó 12.586 alumnos, nominados como zona sur, por pertenecer al censo de las cinco ciudades más importantes del sur de Chile, esto es, Temuco, Valdivia, Osorno, Puerto Montt y el caso extremo de Ancud. Para completar el análisis correlacional y despejar el enmascaramiento de los resultados al considerar solo los datos Censales, se elaboró un muestreo dirigido, seleccionando las ciudades de Temuco, por ser la ciudad más grande. Ancud por sus características de ciudad pequeña y extrema y finalmente un colegio seleccionado al azar entre las cinco ciudades. Se determinó que existe una baja relación entre las pruebas SIMCE y el promedio de notas, demostrando que el instrumento utilizado a nivel nacional para medir los aprendizajes en estudiantes de enseñanza básica, difiere notablemente de los utilizados por el profesorado en los colegios.

**PALABRAS CLAVE:** Pruebas nacionales, Calidad de la educación, Sistema de Medición de Calidad de la Educación, SIMCE.

### ABSTRACT:

The Educational Quality Measurement System [SIMCE] has the institutional objective of improving the quality of education and keeping citizens abreast of the different types of learning achievements of the student body in Chile. However, not much tangible research has been done in this regard. The main purpose of this study is to establish a correlation between end of the year grades for 8th grade students and SIMCE scores attained by the students in Language and Communication, Mathematics and Natural Sciences tests. The data was collected from 8<sup>th</sup> grade students by the Quality Education Agency which pertains to Chile's Ministry of Education. The original information included 12,586 students from the Southern Area of the country since it comprises the census of the five most important cities in Southern Chile which include Temuco, Valdivia, Osorno, Puerto Montt and Ancud which represents an extreme scenario. In order to complete this correlational analysis and unmask the results which had only took census data into account, a targeted sample was developed using the city of Temuco, the largest city, Ancud, the smallest and most remote city and a high school selected randomly among the five cities. It was determined that there is a weak relationship between SIMCE test scores and student grade point averages which proves that the national instrument used to gauge student learning in Chile's Basic Education Cycle, differs significantly from those used by school faculty.

**KEYWORDS:** National tests, Quality of Education, Educational Quality Measurement System, SIMCE.

## INTRODUCCIÓN

El gobierno chileno, después de asumir una lógica de mercado en la educación, crea en 1988 el Sistema de Medición de la Calidad de la Educación (SIMCE), encargado de medir la calidad de la educación y preparar el *ranking* nacional de calidad de los colegios del país, que guía a las familias en la selección de la mejor educación para los educandos (Rojas, 2008).

El origen institucional y unilateral de la prueba sigue generando resistencia social y gremial del profesorado, que se agudiza con motivo de la entrega anual de resultados de la prueba (Inzunza, y Campos, 2017), además también pudiera explicar la ideologización de la investigación y discusión respecto a sus resultados, que en general ha abordado temas propios de la contingencia social o efectos sociales indeseados, descuidando el fondo del problema, que es su inconsistencia para medir la calidad de los aprendizajes en la educación chilena y con ello asegurar a las familias chilenas que el *ranking* de colegios generado con los puntajes SIMCE, es confiable. Actualmente, Chile es uno de los países latinoamericanos que posee más pruebas estandarizadas nacionales y extranjeras, aplicando once a su alumnado de enseñanza básica y media, sin resolver aún el fondo del problema que es si se está midiendo con certeza la calidad de la educación (San Martín, Paukner y Sanhueza, 2015), puesto que los puntajes SIMCE se mantienen casi sin variación durante la última década. La comparación de resultados SIMCE con sistemas estandarizados internacionales como el Programa para la Evaluación Internacional de Estudiantes (PISA) no son todavía concluyentes y en un primer intento de establecer equivalencias de puntajes en ambas pruebas, se encontraron criterios promisorios los cuales permiten profundizar estudios posteriores en este ámbito (Fernández, 2009). Comparando detalladamente los resultados obtenidos por estudiantes de 2º Medio en la prueba SIMCE de Matemáticas, con los equivalentes de PISA del mismo año, se encontró que los contenidos incluidos en la prueba PISA equivalen a los del Marco Curricular chileno, aun cuando las exigencias de esta última van más allá del mero conocimiento conceptual y procedimental, avanzando en la aplicación de las Matemáticas para la solución de problemas contextualizados (Alfaro y Gormaz, 2009).

En este escenario de incertidumbre acerca de la baja capacidad predictiva de la prueba SIMCE en función del rendimiento académico del estudiantado, el objetivo principal de la investigación es establecer si existe una relación consistente entre las notas del alumnado al finalizar la enseñanza básica (8º básico) y los puntajes SIMCE en las pruebas de Lenguaje y Comunicación, Matemáticas y Ciencias Naturales. Se inicia el análisis con un agrupamiento zona sur de los resultados, considerando el conjunto de datos de las tres pruebas en las ciudades Temuco, Valdivia, Osorno, Puerto Montt y Ancud. Luego se avanza en un análisis más detallado, considerando los datos de las tres pruebas para las ciudades de Temuco y Ancud por separado. Se finaliza la profundización del análisis, examinando los datos de un colegio elegido al azar en una de las cinco ciudades seleccionadas en la investigación.

## MARCO TEÓRICO

En relación con los orígenes del SIMCE, se puede rastrear desde la preocupación del Estado chileno por medir los logros de aprendizaje en la educación básica que se manifiesta el año 1960 con la Prueba Nacional, bajo la lógica de un Estado docente responsable de la educación pública obligatoria (Bravo, 2011) y se mantiene vigente hasta que en 1980, la dictadura militar la cambia drástica y arbitrariamente por una lógica de mercado, que transforma a la educación chilena en un *commodity*, mediante una Directiva Presidencial (Falabella, 2015), en que la subsidiaridad del estado traspasa la decisión de dónde matricular a los educandos, a sus padres, quienes como consumidores deciden el tipo de colegio, de acuerdo a su condición económica y al *ranking* de calidad de colegios que se elaboraría de acuerdo a los resultados de la prueba SIMCE (Ruiz, 2010), asumiendo que es una herramienta válida para generarla, validez que este artículo cuestiona.

Los efectos prácticos del drástico cambio de paradigma fueron entre otros, que el financiamiento de la educación pública deja de ser centralizado y se cambia por otro mediante *vouchers*, que premian el desempeño del colegio en la prueba SIMCE; la administración de la educación pública se traspasa a los municipios; se promueve el ingreso a la *industria de la educación* (reconocida así por autoridades actuales del país) de los privados, con prácticamente casi ninguna regulación en los primeros años; se crea una prueba nacional estandarizada (SIMCE), que supuestamente mide anualmente los conocimientos adquiridos por el estudiantado de los distintos tipos de colegios (Falabella, 2015) y con sus resultados se asigna un *voucher* de acuerdo al puntaje SIMCE de cada colegio; estos mismos puntajes permitirían clasificar anualmente a los colegios en una escala de Categoría de Desempeño y además, que es lo más importante y riesgoso, proponer el cierre de un colegio clasificado como **insuficiente** durante cuatro años seguidos, en virtud del nuevo Sistema Nacional de Aseguramiento de la Calidad de la Educación Escolar (SAC) (Agencia de Calidad de la Educación, 2019).

**La evolución de los rendimientos en la prueba SIMCE**, desde la implantación arbitraria del sistema, muestran que las mejoras son poco significativas luego de tres décadas de la imposición del sistema, a pesar de los denodados esfuerzos de gobiernos de diferente orientación política por demostrar algún grado de mejora interna (Bravo, 2011), ratificando que la discusión ideológica entre Estado docente y Estado subsidiario para los efectos de la calidad de la educación chilena es inconducente si no se agregan pruebas tangibles al respecto, que es el propósito de este artículo. Respecto a estándares internacionales la situación es más crítica, pues la autocomplacencia nacional no es suficiente para contrarrestar que la mejora en puntaje en tres de las principales pruebas (Lenguaje y Comunicación, Matemática y Ciencias Naturales) desde la implementación de la prueba SIMCE, es casi la misma que aún nos separa del promedio de aprendizaje de las pruebas internacionales (Bravo, 2011). Si a lo anterior agregamos que la calidad de la educación básica chilena como un todo, es la peor de los países OCDE, ubicándose en la posición 89 de su *ranking* y que los conocimientos de Matemáticas y Ciencias ubican a Chile en el penúltimo lugar del *ranking* OCDE, se configura un panorama complicado para seguir apostando por el SIMCE como una herramienta válida para medir la calidad de los aprendizajes del estudiantado chileno de educación básica y secundaria (Riveros y Báez, 2014).

**¿Qué evalúa el SIMCE?**, es entonces una pregunta perfectamente válida, dado que los gobiernos posdictadura mantuvieron en general el diseño del estado subsidiario, descentralizado y basado en el mercado de la educación (San Martín, et al. 2015). Solo se incorporaron algunos avances tímidos en regulación y se mantuvieron en la norma, las evaluaciones estandarizadas y los incentivos por desempeño (Falabella, 2015), que en muchos casos justificaron la creación de *colegios de excelencia*, con base en su desempeño SIMCE. Por otro lado, se postula que la repetición anual de la prueba hace que la preocupación primordial del profesorado en sus colegios, sea la preparación para la prueba y se dejen con muy poca atención los aprendizajes reales en sus cursos, ello porque la hipervigilancia de la educación de mercado impone actualmente a los colegios, además de responder a indicadores de desempeño no validados objetivamente hasta el momento, la obligación de presentar una opción atractiva para padres en busca de educación de calidad para sus educandos (Falabella, 2018).

A nivel internacional se ha estudiado bastante el fenómeno de la educación chilena y algunas propuestas concretas de mejora van en primer lugar por eliminar el sistema de *vouchers*, salvo que se mejoren drásticamente los mecanismos de asignación, de manera que realmente apoyen y no cierren a los colegios más necesitados o *insuficientes*, pues hay evidencia concreta en el propio país, de que con incentivos bien aplicados estos colegios mejoran significativamente sus aprendizajes (Murnane, Waldman, Willett, Bos y

Vegas, 2017). Se postula, además, que la retroalimentación desde las instancias administrativas de la prueba debiese centrarse más en temas pedagógicos, que en sugerencias de mercadotecnia y finalmente, que los aprendizajes deben ser recíprocos, pues hasta el momento pareciera que la institucionalidad aprende menos que el estudiantado (Vegas, 2018).

**La validez del SIMCE** ha sido hasta hoy, más discutida desde el punto de vista socio político y justamente ése es el talón de Aquiles para su reformulación (San Martín, et al. 2015; Falabella, 2015). Existen deficiencias estructurales de la prueba, que de ser modificadas mejorarían significativamente su validez y a la vez bajarían el nivel de rechazo ciudadano y el del profesorado que la sufren. La deficiencia de fondo radica en la falta de definiciones claras y objetivas sobre la calidad de los aprendizajes, que es su objetivo principal de medición; le sigue en importancia el antagonismo de dos de sus macro propósitos, como son la rendición de cuentas y al mismo tiempo el nivel, el apoyo a las prácticas pedagógicas; luego se menciona como una incoherencia de diseño que se le hayan contabilizado 17 propósitos diferentes, lo que parece a todas luces un exceso o al menos un despropósito (Flórez, 2013). En la dimensión de constructo y contenidos de la prueba se critica que la declaración documental, muchas veces no coincide con lo que la prueba mide y se cuestiona, además, que constructo y contenidos varíen de un año a otro, dificultando su comparabilidad (Flórez, 2013). En la dimensión de criterio, el profesorado ha planteado la falta de coincidencia entre lo enseñado en los colegios y lo medido por la prueba (Flórez, 2013). Los resultados de este artículo refuerzan el planteamiento del profesorado, mostrando la baja consistencia de la prueba SIMCE para medir aprendizajes.

## METODOLOGÍA

### Fuente de datos y selección de la muestra

Los datos utilizados en esta investigación corresponden a estudiantes de ambos sexos de octavo año básico, proporcionados por la Agencia de Calidad de la Educación, del Ministerio de Educación de Chile. Los datos proporcionados correspondieron a los puntajes de las pruebas SIMCE de Lenguaje y Comunicación (en adelante LC), Matemáticas (M), y Ciencias Naturales (CN) y el promedio general de notas, todos del año 2015 (Agencia de Calidad de la Educación, 2015). La información original incluyó 12.586 estudiantes, nominados como zona sur, por pertenecer al censo de las cinco ciudades más importantes del sur de Chile, esto es, Temuco, Valdivia, Osorno, Puerto Montt y el caso extremo de Ancud. De los 12.586 estudiantes, 11.016 rindieron LC, 11.138 M y 11.176 CN. Esta investigación analiza la correlación entre las pruebas SIMCE y el promedio general de notas. Para completar el análisis correlacional y despejar el enmascaramiento de los resultados al considerar solo los datos Censales, se desarrolló un muestreo dirigido, seleccionando las ciudades de Temuco, por ser la ciudad más grande, con 3.925 estudiantes. Ancud por sus características de ciudad pequeña y extrema con 637 estudiantes y finalmente un colegio seleccionado al azar entre las cinco ciudades, con un total de 313 estudiantes.

### VARIABLES DE ESTUDIO

Las Pruebas SIMCE de LC, M y CN, son variables de tipo numéricas continuas y corresponden al puntaje obtenido por el alumnado en la evaluación de aprendizaje.

El Promedio General, es una variable de tipo numérica continua y corresponde al promedio general anual del estudiantado que tiene en el curso. El promedio varía entre 1 y 7, siendo 4 la calificación mínima de aprobación. Esta variable se correlaciona con cada una de las pruebas SIMCE.

## Análisis de correlación y gráficas

Los análisis estadísticos se ejecutaron empleando el programa estadístico de acceso libre R Studio versión 1.1.442 (RStudio, 2016).

Para establecer correlaciones lineales entre el promedio de notas y el rendimiento en puntaje de las diferentes pruebas, se aplicó el estadístico coeficiente de correlaciones simples producto momento de Pearson (Lewis-Beck y Lewis-Beck, 2016),

Para la representación gráfica de la relación entre el promedio general notas y el puntaje de las pruebas SIMCE, se construyeron gráficos de cajas o diagrama de Tukey (Kloke y McKean, 2015), para diferentes rangos de notas, con una amplitud de clase de 0,5 puntos.

Las comparaciones entre grupos de estudiantes según cohortes de notas se elaboraron a través del test no paramétrico de Kruskal-Wallis, debido a que los datos no cumplieron los supuestos de normalidad y homogeneidad de varianza (Conover, 1999).

## ANÁLISIS Y DISCUSIÓN DE RESULTADOS

Un primer análisis de resultados para todas las muestras, indica una baja correlación entre el promedio general de notas y los correspondientes puntajes de las tres pruebas SIMCE, salvo raras situaciones que superan levemente el nivel de 0,5 (Tabla 1).

**TABLA 1**  
Relación entre el Promedio General del alumnado y las pruebas  
del sistema de medición de la calidad de la educación (SIMCE)

Localidad	Lenguaje y Comunicación	Matemáticas	Ciencias Naturales
Zona Sur	0,476 (**)	0,483 (**)	0,492 (**)
Valdivia	0,512 (**)	0,533 (**)	0,531 (**)
Temuco	0,487 (**)	0,509 (**)	0,515 (**)
Osorno	0,484 (**)	0,493 (**)	0,485 (**)
Puerto Montt	0,446 (**)	0,434 (**)	0,465 (**)
Ancud	0,450 (**)	0,438 (**)	0,448 (**)
Colegio	0,458 (**)	0,477 (**)	0,447 (**)

Nota: \* p < 0,05; \*\* p < 0,01

Fuente: Elaboración propia

La baja relación entre el rendimiento académico del alumnado en todas las muestras y las tres pruebas SIMCE, se ve ratificada en la distribución de amplitud y anomalías de los puntajes en los diagramas de cajas de Tukey que se presentan para los diferentes rangos de notas (Figura 1). Lo anterior no permite relacionar con algún grado de certeza que una determinada nota o rango de esta, logre explicar un puntaje en las pruebas SIMCE. Situación que deja en evidencia la inconsistencia instrumental para evaluar los aprendizajes del alumnado, entre lo que aplica el Estado de Chile a través de la prueba SIMCE en diferentes áreas de conocimiento y lo que aplican los colegios a través del profesorado. Esta inconsistencia ha llevado a cuestionar en el país, la validez de esta prueba para evaluar los aprendizajes en estudiantes de enseñanza básica y media (Meckes, 2003).

A una década de la implantación de la prueba SIMCE, se critica formalmente la dificultad para interpretar sus resultados, pues sus niveles de dificultad varían según sus validaciones anuales, se plantea además la

debilidad de su validez al bajar demasiado su nivel de dificultad, lo cual se reflejaría en que su complejidad correspondería a dos y tres cursos menores a los que se aplica (Eyzaguirre y Fontaine, 1999).

**TABLA 2**  
Estadísticos descriptivos del rendimiento de las pruebas SIMCE para diferentes localidades

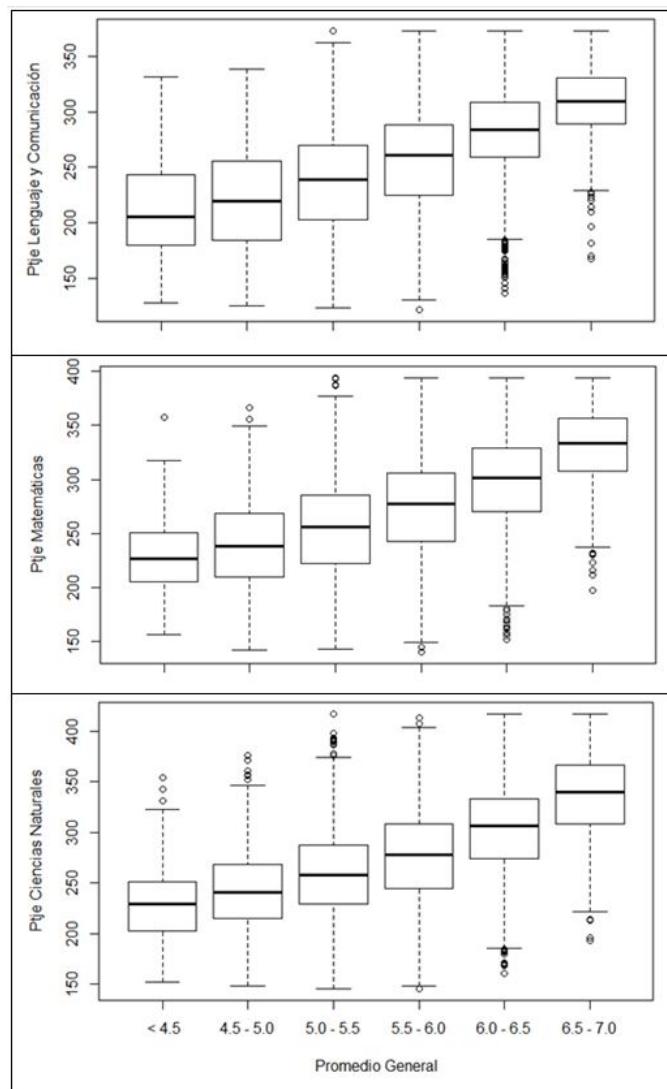
Localidad	Lenguaje y Comunicación			Matemáticas			Ciencias Naturales		
	Media	D.E.	N	Media	D.E.	N	Media	D.E.	N
Zona Sur	249,0	49,0	11.016	267,3	48,5	11.138	270,7	48,9	11.176
Temuco	247,3	49,6	3.439	268,6	48,2	3.481	272,6	49,6	3.489
Valdivia	252,7	48,1	1.819	269,7	48,0	1.858	274,7	48,6	1.866
Osorno	247,3	49,6	3.439	268,6	48,2	3.481	272,6	49,6	3.489
Puerto Montt	247,8	49,4	3.108	264,4	50,0	3.152	268,0	49,3	3.161
Ancud	241,6	47,0	559	257,9	48,6	563	258,2	45,6	561
Colegio	273,4	40,7	296	301,9	33,86	303	305,0	41,59	303

Nota: D.E = Desviación estándar; N = Número total de estudiantes

Fuente: Elaboración propia

Al desarrollarse un análisis macro a nivel de la zona sur, se obtienen correlaciones bajas entre el promedio de notas y los puntajes para cada una de las pruebas, situación similar ocurre a nivel de ciudad e incluso considerando solo un colegio en particular (Tabla 1), a pesar de que este último si tiene promedios de puntajes más altos en las tres pruebas, esto es, el colegio seleccionado con el 2.7% de representación sobre el total del Sur de Chile, obtiene promedios superiores en 24 puntos en LC, 34 puntos en M y 34 en CN (Tabla 2). Para efecto de revelar de manera más clara las inconsistencias entre el rendimiento del estudiantado en el colegio y los resultados obtenidos del examen de la prueba SIMCE, es recomendable elaborar un análisis correlacional entre notas de enseñanza media y los resultados de la prueba SIMCE a nivel de colegio y no agrupados por ciudades, regiones o país.

## Zona Sur



**FIGURA 1**

Diagramas de Tukey para puntajes SIMCE en las diferentes categorías de promedio de notas, Zona Sur

Fuente: Elaboración propia

Las variaciones del puntaje SIMCE en las tres pruebas y analizado por cada rango de nota, presentan una gran dispersión, destacando para las tres pruebas la mayor variabilidad en los rangos de notas 5,0 – 5,5 y 5,5 – 6,0 (Figura 1, Tabla 3).

Al ejecutar un análisis comparativo a nivel de puntajes medios de las pruebas SIMCE y para cada uno de los rangos de notas, todos presentan diferencias significativas excepto en LyC para los rangos de notas < 4,5 y 4,5 – 5,0, sin embargo, el análisis de correlación per se efectuado (Tabla 1), invisibiliza la variabilidad de puntajes SIMCE para estudiantes que demuestran el mismo rendimiento en el colegio (Tabla 3). Esta inconsistencia apoya las críticas respecto a la validez de la prueba SIMCE para medir aprendizajes, pues la otra interpretación posible para ello, es que la calificación del profesorado estuviera errada, opción menos aceptable dado que implicaría una deficiencia a nivel docente en todo el país.

**TABLA 3**  
**Diferencias estadísticamente significativas de los puntajes**  
**SIMCE para los rangos del promedio general de notas, Zona Sur**

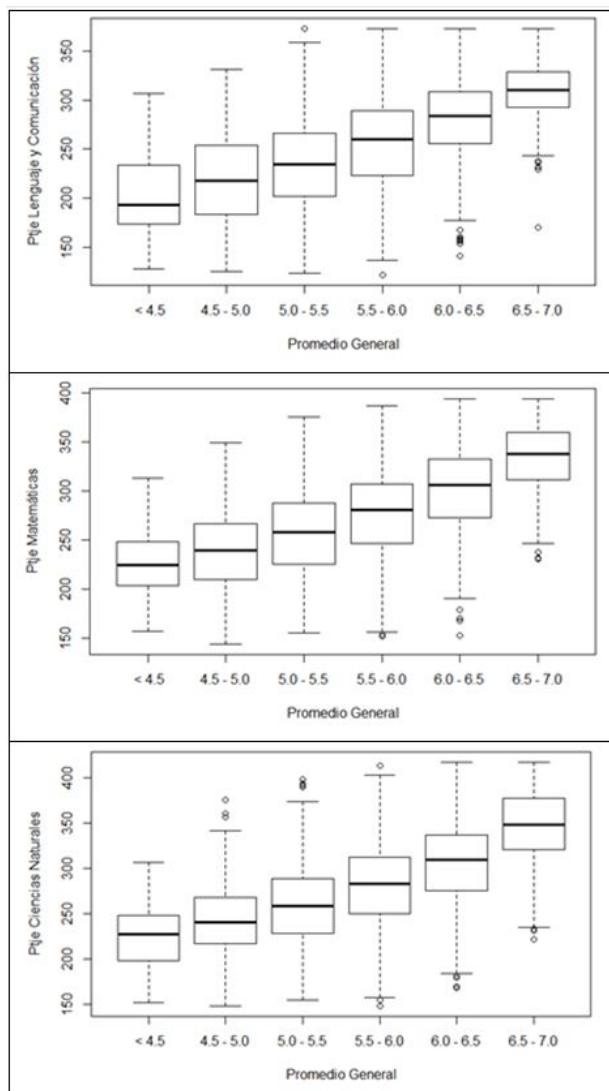
Pruebas SIMCE	Promedio General	N	Media	Desviación Estándar	Mínimo	Máximo	K-W
Lenguaje y Comunicación	< 4,5	301	211,7 <sup>3<sup>a</sup></sup>	42,50	127,68	331,44	
	4,5 - 5,0	1.761	220,6 <sup>7<sup>a</sup></sup>	44,08	125,23	338,64	
	5,0 - 5,5	3.582	236,8 <sup>1<sup>b</sup></sup>	44,90	123,64	372,93	2.398 <sup>:gl=5;</sup>
	5,5 - 6,0	3.111	255,9 <sup>3<sup>c</sup></sup>	43,94	121,66	373,16	<sup>p&lt;0,</sup> <sup>01</sup>
	6,0 - 6,5	1.821	280,5 <sup>9<sup>d</sup></sup>	40,01	136,76	376,16	
	6,5 - 7,0	440	306,7 <sup>7<sup>e</sup></sup>	36,05	160,87	373,16	
Matemáticas	< 4,5	304	228,7 <sup>6<sup>a</sup></sup>	36,58	156,30	357,26	
	4,5 - 5,0	1.776	239,7 <sup>5<sup>b</sup></sup>	39,87	141,90	366,69	
	5,0 - 5,5	3.622	255,0 <sup>8<sup>c</sup></sup>	42,59	143,40	394,39	2.451 <sup>:gl=6;</sup>
	5,5 - 6,0	3.146	273,6 <sup>3<sup>d</sup></sup>	44,75	140,17	394,39	<sup>p&lt;0,</sup> <sup>01</sup>
	6,0 - 6,5	1.845	298,1 <sup>9<sup>e</sup></sup>	44,15	152,20	394,39	
	6,5 - 7,0	445	329,4 <sup>1<sup>f</sup></sup>	36,74	157,69	394,39	
Ciencias Naturales	< 4,5	304	229,8 <sup>4<sup>a</sup></sup>	35,88	152,31	354,24	
	4,5 - 5,0	1.792	242,7 <sup>1<sup>b</sup></sup>	38,42	148,66	375,77	
	5,0 - 5,5	3.632	258,7 <sup>3<sup>c</sup></sup>	42,33	145,58	417,13	2.473 <sup>:gl=5;</sup>
	5,5 - 6,0	3.156	276,5 <sup>4<sup>d</sup></sup>	44,72	145,24	412,89	<sup>p&lt;0,</sup> <sup>01</sup>
	6,0 - 6,5	1.847	302,7 <sup>0<sup>e</sup></sup>	44,58	160,85	417,13	
	6,5 - 7,0	445	335,0 <sup>3<sup>f</sup></sup>	44,53	193,06	417,13	

Nota: Para las columnas de promedios, las coincidencias de las letras minúsculas en el superíndice indican la no existencia de diferencias significativas (prueba K-W,  $p<0.01$ ).

Fuente: Elaboración propia

Comparando los puntajes medios entre las tres pruebas se observa que LC es 18 puntos inferior a M y 22 respecto a CN (Tabla 3). La interrogante que se plantea y que forma parte de la discusión, es entonces ¿cómo el alumnado comprende lo que leen en matemáticas y ciencias naturales? (Marchant, Lucchini, y Cuadrado, 2007).

## Ciudad de Temuco



**FIGURA 2**

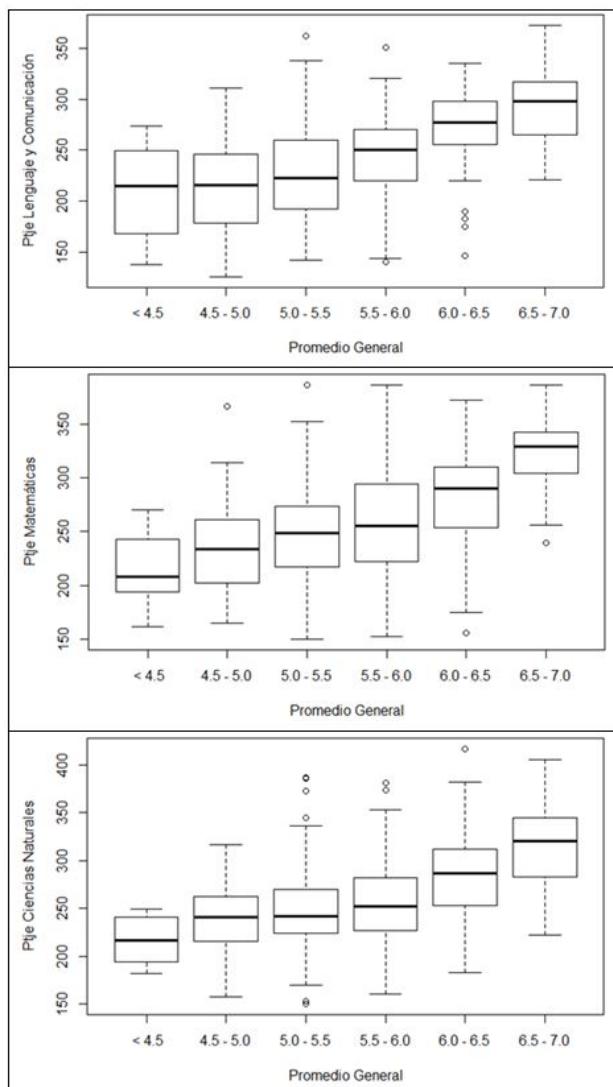
Diagramas de Tukey para los puntajes SIMCE en las diferentes categorías de promedio de notas, Temuco

Fuente: Elaboración propia

Al analizar las tres pruebas SIMCE y su relación con las notas de colegio, consolidando todos los colegios de la ciudad de Temuco, se observa que para todas las pruebas en los rangos de notas 5,0 y 6,5 se encuentran las mayores dispersiones en puntajes (Figura 2).

La diferencia de puntajes entre las pruebas mantiene la tendencia de la Zona Sur, pues el mínimo de LyC es 21 puntos menores al de M y 26 puntos menor al de CN. En los puntajes máximos la diferencia de LyC es 21 puntos menores al puntaje de M y 44 puntos menores al de CN (Tabla 4, Temuco).

## Ciudad de Ancud



**FIGURA 3**

Diagramas de Tukey para los puntajes SIMCE en las diferentes categorías de promedio de notas, Ancud

Fuente: Elaboración propia

En la ciudad de Ancud podemos observar que la dispersión de los puntajes en las tres pruebas es notoriamente menor que en la Zona Sur y la ciudad de Temuco, en al menos un 15 %. En LyC las mayores dispersiones de puntaje se observan en los rangos 4,5 – 5,0 y 6,0 – 6,5, M muestra igualmente que los rangos de mayor dispersión son desde 4,0 – 5,0 y 6,0 – 6,5. Las dispersiones de puntajes en CN se dan entre los rangos de notas 5,0 – 5,5 y 6,0 – 6,5 (Figura 3, Tabla 4, Ancud).

**TABLA 4**  
Diferencias estadísticamente significativas entre el promedio general de notas y los puntajes SIMCE

Temuco							
Pruebas SIMCE	Promedio General	N	Media	Desviación Estándar	Mínimo	Máximo	K- W
Lenguaje y Comunicación	< 4,5	118	204,83 <sup>a</sup>	43,02	127,68	306,38	
	4,5 - 5,0	547	219,24 <sup>a</sup>	43,79	125,34	331,52	782; gl=5;
	5,0 - 5,5	1.071	234,21 <sup>b</sup>	44,66	123,64	372,93	;
	5,5 - 6,0	1.011	255,87 <sup>c</sup>	44,56	121,66	373,16	p<0, 01
	6,0 - 6,5	563	278,87 <sup>d</sup>	41,92	141,24	372,93	
	6,5 - 7,0	129	307,88 <sup>e</sup>	31,68	170,63	372,93	
Matemáticas	< 4,5	115	226,31 <sup>a</sup>	33,10	156,30	313,09	
	4,5 - 5,0	554	238,70 <sup>a</sup>	39,23	143,34	349,42	850;
	5,0 - 5,5	1.087	256,31 <sup>b</sup>	41,38	155,32	375,39	gl=5;
	5,5 - 6,0	1.028	276,52 <sup>c</sup>	42,73	151,71	386,64	;
	6,0 - 6,5	569	300,96 <sup>d</sup>	44,98	152,2	394,39	p<0, 01
	6,5 - 7,0	128	332,63 <sup>e</sup>	37,19	230,52	394,39	
Ciencias Naturales	< 4,5	116	226,90 <sup>a</sup>	34,60	152,31	306,51	
	4,5 - 5,0	558	243,04 <sup>a</sup>	38,07	148,66	375,77	839;
	5,0 - 5,5	1.085	259,71 <sup>b</sup>	41,99	154,95	397,69	gl=5;
	5,5 - 6,0	1.031	280,88 <sup>c</sup>	44,56	148,40	412,89	;
	6,0 - 6,5	570	304,89 <sup>d</sup>	46,61	168,08	417,13	p<0, 01
	6,5 - 7,0	129	341,37 <sup>e</sup>	44,38	221,96	416,90	
Ancud							
Lenguaje y Comunicación	< 4,5	14	211,58 <sup>a</sup>	43,68	137,43	273,95	
	4,5 - 5,0	85	214,58 <sup>a</sup>	45,74	125,23	310,83	
	5,0 - 5,5	166	228,06 <sup>a</sup>	43,51	141,93	362,63	128; gl=5;
	5,5 - 6,0	173	244,41 <sup>a</sup>	40,40	140,34	350,84	;
	6,0 - 6,5	87	272,47 <sup>b</sup>	35,39	145,93	335,66	p<0, 01
	6,5 - 7,0	34	294,25 <sup>c</sup>	38,36	220,86	373,16	
Matemáticas	< 4,5	12	216,83 <sup>a</sup>	35,18	161,64	270,43	
	4,5 - 5,0	85	234,83 <sup>a</sup>	40,81	164,55	366,69	
	5,0 - 5,5	165	246,87 <sup>a</sup>	39,96	149,61	386,82	115; gl=6;
	5,5 - 6,0	177	257,21 <sup>a</sup>	48,32	152,20	386,82	;
	6,0 - 6,5	89	281,32 <sup>b</sup>	45,00	155,33	372,67	p<0, 01
	6,5 - 7,0	35	323,48 <sup>c</sup>	32,41	239,81	386,82	
Ciencias Naturales	< 4,5	12	217,45 <sup>a</sup>	24,30	182,06	249,24	
	4,5 - 5,0	84	238,37 <sup>a</sup>	33,73	157,09	317,04	
	5,0 - 5,5	166	246,38 <sup>a</sup>	39,82	149,86	386,81	113; gl=5;
	5,5 - 6,0	176	256,28 <sup>a</sup>	41,23	160,30	381,60	;
	6,0 - 6,5	88	285,34 <sup>b</sup>	43,11	182,51	417,13	p<0, 01
	6,5 - 7,0	35	317,60 <sup>b</sup>	47,12	222,46	405,87	

Nota: Para las columnas de promedios, las coincidencias de las letras minúsculas en el superíndice indican la no existencia de diferencias significativas (prueba K-W,  $p<0.01$ ).

Fuente: Elaboración propia

En el caso de Temuco, el análisis comparativo a nivel medias de los puntajes en todos los rangos de notas muestra que solo los dos más bajos no presentan diferencias estadísticamente significativas en ninguna de las tres pruebas (Tabla 4, Temuco). Se justificaría así, la separación que se hizo al analizar la muestra por el número de colegios en cada una de ellas, dado que en 344 colegios de la Zona Sur se enmascara con mayor facilidad el comportamiento de los puntajes SIMCE respecto a 100 colegios de Temuco y las respectivas notas obtenidas por el alumnado en sus colegios.

Considerando la significancia de las medias de los puntajes en relación con los rangos de notas, el comportamiento en Ancud es notoriamente diferente a la Zona Sur y la ciudad de Temuco, pues en ninguna de las tres pruebas existen diferencias significativas desde el rango  $< 4,5$  hasta 6,0. Los rangos 6,0 – 6,5 y 6,5 – 7,0 muestran significancia entre sí y con el resto de los rangos inferiores en las pruebas LyC y M, no así en la prueba de CN en la que no presentan diferencias significativas entre sí, pero ambos la presentan respecto al resto de notas inferiores.

La diferencia de puntajes SIMCE mínimos entre las pruebas en Ancud, es también diferente a la Zona Sur y Temuco, pues el de LyC es 25 puntos menores al de M, pero M y CN no muestran diferencia de puntaje entre ellas. En los puntajes máximos se aprecia un comportamiento acorde con las tendencias anteriores, pues LyC es 14 puntos menos que M y esta a su vez es 30 puntos menos que CN.

Esta situación ratifica uno de los propósitos del artículo, comprobar la tendencia a la mayor inconsistencia de los puntajes respecto a las notas del colegio, en la medida que disminuye el tamaño de la muestra, que en el caso de Ancud es de 32 colegios.

## Colegio seleccionado al azar

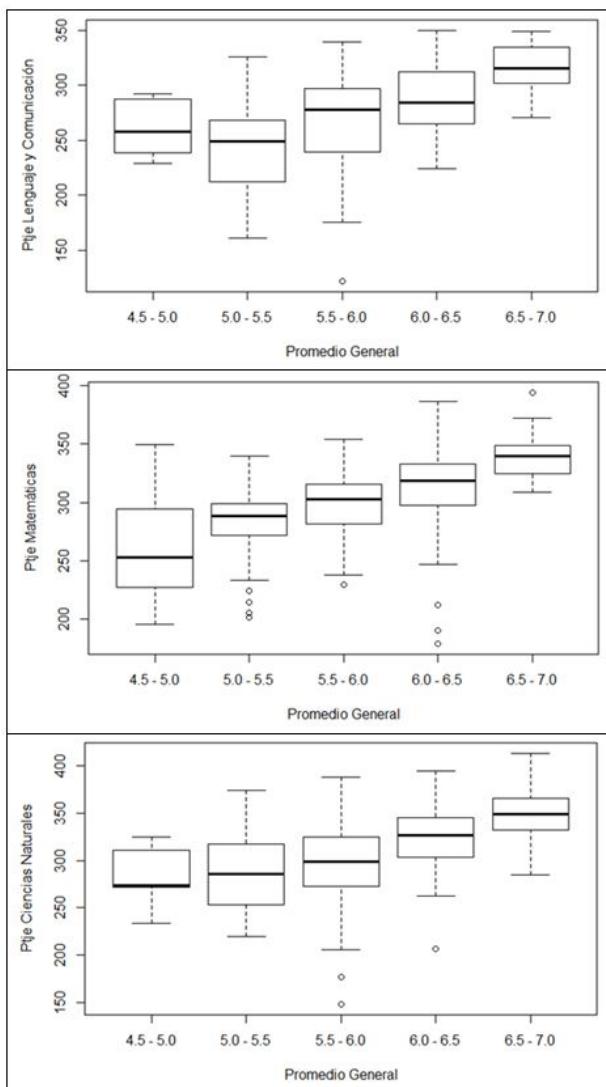


FIGURA 4

Diagramas de Tukey para los puntajes SIMCE en las diferentes categorías de promedio de notas, Colegio  
Fuente: Elaboración propia

El comportamiento del colegio seleccionado es el más diferente de todos los análisis elaborados, pues no existen estudiantes con notas menores a 4,5, desapareciendo esta categoría en todas las pruebas y las dispersiones de puntajes son menos amplias que todas las analizadas previamente en al menos un 25% (Figura 4).

La prueba de LyC muestra la mayor dispersión en los rangos 5,0 – 6,0, y al mismo tiempo la menor dispersión de todos los análisis desarrollados en el rango 4,5 – 5,0. La prueba de M muestra igualmente dispersiones más acotadas y las mayores dispersiones se encuentran en los rangos 4,5 – 5,0 y 6,0 – 6,5. En la prueba de CN se aprecia un comportamiento semejante en la menor dispersión de los puntajes, encontrándose las mayores dispersiones en los rangos 5,5 – 6,0 y 6,0 – 6,5.

**TABLA 5**  
**Diferencias estadísticamente significativas encontradas entre el promedio general de notas y los puntajes SIMCE, colegio seleccionado al azar**

Pruebas SIMCE	Promedio General	N	Media	Desviación Estándar	Mínimo	Máximo	K-W
Lenguaje y Comunicación	4,5 - 5,0	6	260,4 7a	26,74	228,95	292,66	
	5,0 - 5,5	59	246,7 4a	37,28	160,80	326,22	65;
	5,5 - 6,0	119	267,7 7ab	41,21	121,66	339,29	gl=4; p<0, 01
	6,0 - 6,5	90	288,7 6abc	30,59	224,10	350,28	
	6,5 - 7,0	22	315,8 5ac	24,13	270,99	349,31	
	4,5 - 5,0	7	263,1 7a	53,63	195,67	349,42	
Matemáticas	5,0 - 5,5	63	283,2 1ab	29,09	201,25	339,59	73;
	5,5 - 6,0	121	298,8 9ab	27,17	229,81	354,26	gl=4; p<0, 01
	6,0 - 6,5	91	313,3 0ac	34,06	178,83	386,82	
	6,5 - 7,0	21	339,3 0c	20,96	308,84	394,39	
	4,5 - 5,0	7	285,4 5a	31,70	233,66	324,96	
	5,0 - 5,5	62	285,8 7ab	40,11	219,94	373,92	64;
Ciencias Naturales	5,5 - 6,0	122	295,0 7ab	40,70	148,40	387,75	gl=5; p<0, 01
	6,0 - 6,5	91	322,4 1ac	31,95	207,18	393,65	
	6,5 - 7,0	21	349,5 7c	31,01	284,49	412,89	

Nota: Para las columnas de promedios, las coincidencias de las letras minúsculas en el superíndice indican la no existencia de diferencias significativas (prueba K-W,  $p<0.01$ ).

Fuente: Elaboración propia

La conducta del colegio en relación con todos los análisis previos, se diferencia también en la significancia de los puntajes respecto de las notas que no siguen un mismo patrón en las tres pruebas. La prueba de LyC muestra que ningún rango de notas tiene significancia para los puntajes, y que además tampoco la hay entre los rangos 5.5 a 6.5 y 6.0 – 7.0 respectivamente. En las pruebas de M y CN no hay significancia desde el rango 4.5 – 5.0 hasta 6.0 – 6.5, así como tampoco existe entre los rangos 5.0 a 6.0 y 6.0 a 7.0 (Tabla 5).

La diferencia de puntajes mínimos entre pruebas sigue el patrón (no así los rangos de notas) de la Zona Sur y Temuco, pues la prueba de M es superior en 56 puntos a la de LyC y la prueba de CN es superior en 30 puntos a la de M. En el caso de los puntajes máximos el patrón se mantiene, dado que la prueba de M es 44 puntos superior a la de LyC y la prueba de CN es superior en 18 puntos a la de M.

Considerando todos los análisis previos, es preocupante que los puntajes en LyC en todas las muestras en promedio sean al menos 16 puntos menores que el puntaje medio de M y al menos 16 puntos menores que CN. En el colegio, las diferencias medias se amplían aún más, es así como LyC es 29 puntos menor que M y 32 puntos más bajo que CN. Esta constatación es preocupante dado que la *comprensión lectora* a nivel internacional es considerada el requisito básico para entender instrucciones escritas, entender problemas de

la vida real, interactuar en grupos sociales, aprender diferentes temas, darse a entender y lógicamente entender las instrucciones adecuadamente en las propias pruebas SIMCE (Mullis y Martin, 2017). Las evaluaciones de la prueba en cada oportunidad que se aplica aportan un punto de vista diferente, dado que la construcción de las preguntas, varían en cada ocasión que se ejecuta y además el contenido de las preguntas equivale a los conocimientos de dos años previos al nivel de curso del estudiantado que la rinde (Eyzaguirre y Fontaine, 1999; Meckes, 2003).

En la ola social de críticas a la prueba SIMCE, es posible encontrar explicaciones a prácticamente todas las interrogantes discutidas y la primera razón para la irrelevancia de la prueba estaría en las tensiones de propósitos contrapuestos desde su creación: rendición de cuentas (mercado) y apoyo a prácticas pedagógicas (Waissbluth, 2013; Campos y Guerrero, 2016) y la más mencionada por su importancia, es la falta de definición previa del significado de calidad para interpretar los resultados de la prueba. Se plantea además entre las críticas más importantes, la cantidad de propósitos, completando en la actualidad 17, lo cual la hacen muy poco práctica en la evaluación de sus resultados (Flórez, 2013).

La inconsistencia de los resultados discutidos y las críticas revisadas al respecto, indican con claridad que al menos se debieran reformular con claridad y objetividad los propósitos de la prueba, quitarle su orientación de solo comparar colegios para inducir calidad en educación como un bien de consumo, revisar profundamente la calidad y cobertura de las preguntas y finalmente tender hacia menos evaluaciones y mejorar decididamente los aprendizajes (Schinske y Tanner, 2014).

Lo que se observa anualmente a nivel institucional es que la Agencia de Calidad de la Educación (ACE), además de entregar los resultados SIMCE, analiza principalmente las variaciones de los puntajes en las respectivas pruebas y colegios, en una forma muy coherente con una educación de mercado, pero se deja de lado el análisis pedagógico de los puntajes y la consistencia de las respectivas notas, que sería el ideal para identificar fehacientemente las deficiencias de fondo del sistema. El análisis que hace la presente investigación pretende demostrar la inconsistencia de los puntajes SIMCE respecto a las notas del estudiantado que la rinde anualmente y con ello orientar la búsqueda pedagógica de las mejoras necesarias, dejando de lado las discusiones ideológicas, que hasta ahora han producido pocos resultados concretos.

Parece conveniente entonces, que las instancias institucionales responsables del SIMCE hagan un análisis técnico pedagógico, con prescindencia ideológica como base para la obtención de logros de calidad, de la validez de la prueba y mientras ello no suceda, al menos abstenerse de modificar anualmente la prueba, basándose como hasta ahora, solamente en la variación de los puntajes respecto a la aplicación previa de la prueba. Más importante aún, por las repercusiones sociales de la medida, las autoridades responsables del SIMCE debieran abstenerse de ejercer la potestad de cerrar colegios *insuficientes* de acuerdo con los puntajes de la prueba, hasta que se demuestre objetivamente, con datos, dicha insuficiencia, que por lo demás está demostrado, con datos, que esta se subsana con apoyo presupuestario y pedagógico (Murnane, et al. 2017)

## CONCLUSIONES

Los resultados obtenidos en la investigación dejan en evidencia la existencia de una baja relación entre los puntajes obtenidos en la prueba SIMCE y el promedio de notas del alumnado en el colegio, lo cual está revelando que la metodología empleada en la prueba SIMCE, para medir los aprendizajes, difiere considerablemente de los métodos aplicados por los colegios para medir los mismos logros. Otro factor que está influyendo gravemente, son los múltiples propósitos que tiene la prueba SIMCE, no permitiendo evaluar lo que realmente se enseña en los planes de estudio de los colegios. Las múltiples modificaciones que ha experimentado la prueba SIMCE en el tiempo y velando solo por el propósito de que los puntajes obtenidos permiten generar escenarios con *ranking* entre colegios, ha desnaturalizado los contenidos que se evalúan respecto al nivel en que se aplican y el apoyo pedagógico que pueda aportar a los colegios. Si el profesorado conociera y pudiera analizar esta baja relación entre el sistema nacional de evaluación SIMCE y sus propios

métodos para evaluar el aprendizaje del estudiantado, entonces aparecería la interrogante de cuáles serían las estrategias pedagógicas que tendrían que implementar para que sus estudiantes aprendan y que ambos instrumentos converjan en una evaluación.

## AGRADECIMIENTOS

Esta investigación utilizó como fuente de información las bases de datos de la Agencia de Calidad de la Educación. Los autores agradecen a la Agencia de Calidad de la Educación el acceso a la información. Todos los resultados del estudio son de responsabilidad de los autores y en nada comprometen a dicha Institución.

## REFERENCIAS

- Agencia de Calidad de la Educación. (2015). *Bases de Datos de la Agencia de Calidad de la Educación*. Santiago, Chile. Recuperado de <https://www.agenciaeducacion.cl/>
- Agencia de Calidad de la Educación. (2019). *Sistema Nacional de Aseguramiento de la Calidad de la Educación Escolar (SAC)*. Chile. Recuperado de <https://www.agenciaeducacion.cl/orientacion/sac/>
- Alfaro, L. y Gormaz, R. (2009). Análisis comparativo de los resultados chilenos en las pruebas de Matemática SIMCE y PISA. En *¿Qué nos dice PISA sobre la educación de los jóvenes en Chile? Nuevos análisis y perspectivas sobre los resultados en PISA 2006* (pp. 239-260). Santiago, Chile: Gráfica 7 Ltda.
- Bravo, J. (2011). SIMCE: Pasado, presente y futuro del Sistema Nacional de Evaluación. *Estudios Públicos, Estudios Públicos* (123), 189-211
- Campos, J. y Guerrero P. (2016). Efectos indeseados de la medición de la calidad educativa en Chile. La respuesta de la sociedad civil. *Cad. Cedes*, 36(100), 355-374
- Conover, W. (1999). *Practical nonparametric statistics*. United States of America: John Wiley & Sons
- Eyzaguirre, B. y Fontaine L. (1999). ¿Qué mide realmente el SIMCE? *Estudios Públicos*, (75), 110 - 166
- Falabella, A. (2015). El mercado escolar en Chile y el surgimiento de la nueva gestión pública: el tejido de la política entre la dictadura neoliberal y los gobiernos de la centroizquierda (1979 a 2009). *Educ. Soc.*, 36(132), 699-722.
- Falabella, A. (2018). La seducción por la hipervigilancia: el caso de la educación escolar chilena (1973 – 2011). En C. Ruiz, L. Reyes y F. Herrera (Ed.). *Privatización de lo público en el sistema escolar Chile y la agenda global de educación* (pp. 163- 187), Santiago: LOM.
- Fernández, C. (2009). Distribución de puntajes SIMCE en PISA y SIMCE 2006, en perspectiva comparada En *¿Qué nos dice PISA sobre la educación de los jóvenes en Chile?* (pp. 315-326). Santiago, Chile: Gráfica 7 Ltda.
- Flórez M. (2013). *Análisis crítico de la validez del sistema de medición de la calidad de la educación (SIMCE). Informe final*. Oxford, Inglaterra: Oxford University, Centre for Educational Assessment. Recuperado de <https://bit.ly/2lGB0SQ>
- Inzunza, J. y Campos J. (2017). El SIMCE en Chile: historia, problematización y resistencia. En J. Alarid (Presidencia), *XI Seminario Internacional de la Red Estrado. Movimientos Pedagógicos y Trabajo Docente en tiempos de estandarización*. Congreso llevado a cabo en la ciudad de México, México.
- Kloke J., y McKean j. (2015). *Nonparametric Statistical Methods Using R*. Estados Unidos: Chapman & Hall.
- Lewis-Beck, C. y Lewis-Beck, M. (2016). *Applied Regression and Introduction*. Washintong DC: SAGE Publications.
- Marchant, T., Lucchini, G., y Cuadrado, B. (2007). ¿Por qué leer bien es importante?: asociación del dominio lector con otros aprendizajes. *Psykhe (Santiago)*, 16(2), 3-16.
- Meckes, L. (2003). El Simce: Su desarrollo y desafíos actuales. *Pensamiento Educativo*, 33, 160-178.
- Mullis, I. y Martin, M. (2017). *PIRLS 2016 Assessment Framework*. Boston, Estados Unidos: TIMSS & PIRLS International Study Center.

- Murnane, R., Waldman, M., Willett, J., Bos, M. y Vegas, E. (2017). *The consequences of educational voucher reform in Chile. Massachusetts: National Bureau of Economic Research*. Recuperado de <https://www.nber.org/papers/w23550.pdf>
- Riveros, L. y Báez, G. (2014). Chile y la OCDE. La dicotomía entre macroeconómico y el desarrollo humano. *Estudios Internacionales*, 46(179), 9-34
- RStudio Team (2016). *RStudio: Integrated Development for R*. Boston: RStudio, Inc. Recuperado de <http://www.rstudio.com/>
- Rojas, A. (2008). *Situación actual y perspectivas de la educación chilena desde el enfoque de derechos humanos*. Santiago, Chile: Oficina Regional de Educación de la UNESCO para América Latina y el Caribe.
- Ruiz, C. (2010). *De la República al mercado. Ideas educacionales y política en Chile*. Santiago, Chile: Lom Ediciones.
- San Martin, V., Paukner F. y Sanhueza, S. (2015). ¿Cuál cable primero? El desarme del sistema educativo en Chile. *Latinoamérica*, 60(1), 159-185
- Schinske, J., y Tanner, K. (2014). Teaching more by grading less (or differently). *CBE—Life Sciences Education*, 13(2), 159-166.
- Vegas, E. (2018). 5 lecciones de las reformas educativas recientes en Chile. *Brookings*. Recuperado de <https://brook.gs/2lGUbfI>
- Waissbluth, M. (21 de julio de 2013). Las perversiones del SIMCE. *La Tercera*. Recuperado de <https://www.laterceraroma.com/voces/las-perversiones-del-simce/>