



## Artículo

# Capacidad predictiva de las notas en enseñanza media sobre el rendimiento en pruebas de selección universitaria: el caso chileno



Carlos Rodríguez Garcés y Carmen Gloria Jarpa Arriagada\*

Departamento de Ciencias Sociales, Facultad de Educación y Humanidades, Universidad del Bío-Bío, Chillán, Chile

## INFORMACIÓN DEL ARTÍCULO

### Historia del artículo:

Recibido el 17 de octubre de 2014

Aceptado el 18 de marzo de 2015

On-line el 19 de mayo de 2015

### Palabras clave:

Rendimiento escolar

Aprovechamiento educativo

Sistema de selección universitaria

Enseñanza media

Evaluación

## R E S U M E N

El objetivo de este trabajo es realizar una aproximación a la cuantificación de la capacidad predictiva que evidencian las notas de enseñanza media (NEM) en su relación con el puntaje obtenido por estudiantes de Chile en las pruebas de selección universitaria (PSU). Se indaga sobre los puntajes promedio alcanzados por los estudiantes de cada establecimiento educacional en cada una de las pruebas de selección universitaria rendidas en el proceso de admisión 2012, según rama de enseñanza y tipo de establecimiento. Presenta además una matriz de correlaciones (bivariadas y parciales) a fin de establecer la intensidad con que se vinculan los resultados de cada prueba con el rendimiento escolar. Dentro de los principales hallazgos podemos constatar que la segmentación del sistema educativo no configura tan solo niveles de logro diferenciados por tipo de colegio y capital social del alumnado, sino que además da cuenta de sistemas evaluativos internos escasamente válidos y confiables, reflejo de divergentes políticas de asignación de notas y niveles de exigencia. En los establecimientos educativos, en especial los municipales y los que atienden alumnos de alta vulnerabilidad social, los sistemas de evaluación internos evidencian una reducida capacidad predictiva NEM-PSU.

© 2014 Instituto de Ciencias de la Educación de la Universidad de Oviedo. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Predictive ability of high school grades on performance at the national university selection test: The Chilean case

### A B S T R A C T

The aim of this research is to quantify the predictive ability of high school grades (NEM) in relation to score obtained by Chilean students in the national university selection test (PSU). It implies a revision of the average scores achieved by students coming from every High School in the country, by specific area, rendered in the 2012 national university selection test accordingly to types of teaching and school. It also includes a correlation matrix (both bivariate and partial) to determine a relation between each test output and school performance (NEM). Main findings are: The educational system in Chile segregates and sets different levels of achievement according to types of school and students' social capital. It also provides evidence of the existence of poor internal evaluation systems, with little validity and reliability levels, that reflect diverging policies for school grading and academic demand. In municipal high schools or schools that attend to students at high social and economic risk, their internal evaluation systems show a reduced predictive capacity for PSU scores.

© 2014 Instituto de Ciencias de la Educación de la Universidad de Oviedo. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### Keywords:

Academic performance

Educational achievement

National university selection test

High school

Evaluation

## Introducción

El Estado de Chile definió la educación como una estrategia central para el desarrollo socioeconómico de territorios, organizaciones y personas, en la que la instalación de talentos, habilidades y

\* Autor para correspondencia: Avenida La Castilla sin número, Chillán, Chile, Tel.: +56 42 2463556. Correo electrónico: [carmen.jarpa@gmail.com](mailto:carmen.jarpa@gmail.com) (C.G. Jarpa Arriagada).

conocimientos son competencias cruciales en este proceso (OCDE, 2004). Para Guttman (2001) esto supone que la sociedad debe educar a todos los niños y jóvenes para que sean capaces de participar en la definición colectiva de sociedad. Una evidente prueba de esta prioridad nacional es que el gasto público en Educación aumentó de 2,4% del PIB en 1990 a 4,5% en 2009 (UNESCO, 2012). De igual manera, el gasto privado aumentó fuertemente en cada uno de los niveles de enseñanza, se registró una expansión en el número anual de horas de clase, se mejoraron los niveles de cobertura, retención y asistencia del alumnado, en especial de los pertenecientes a los sectores más vulnerables del país; se implementaron nuevos modelos de enseñanza, programas de perfeccionamiento docente, y se hizo una fuerte inversión en tecnologías para la enseñanza, las cuales se vislumbran como factores de relevancia en el proceso de modernización de la educación (OCDE, 2004).

Sin embargo, estos esfuerzos y la alta inversión realizada se han visto escasamente asociados con los niveles de logro escolar definidos en los Planes y Programas del Ministerio de Educación Chileno (MINEDUC). En efecto, el Sistema Nacional de Medición de la Calidad de la Educación así como las pruebas internacionales (TIMSS-1999; TIMSS-2003; PISA-2000; PISA-2006, PISA-2009 y PISA-2012 [OCDE, 2014]) revelan, en el mejor de los casos, solo modestas mejoras. El sistema educacional chileno presenta profundas falencias en cuanto a su capacidad de instalar competencias cognitivas y procedimentales, en especial en los sectores de Lenguaje y Matemática y, especialmente, en la población socioeconómicamente más desfavorecida. De esta manera, se configura como un sistema escasamente eficiente en sus logros de aprendizaje y de calidad fuertemente segmentada según el capital social de su alumnado.

Brunner y Elacqua (2004:1) afirman «[...] mientras más pobre el país, más alto es el porcentaje de la varianza (refiriéndose al rendimiento) explicado por los resultados de la escuela; mientras más rico el país, se atribuye un mayor peso a la familia». Si aplicamos este razonamiento a Chile, los resultados de los estudiantes pobres se explican más por las estrategias pedagógicas de la escuela. Esta debilidad que tienen las escuelas para lograr que sus alumnos alcancen metas educativas de calidad y con independencia de su origen social se ve reflejada también en la escasa pertinencia y capacidad predictiva que evidencian los procesos evaluativos internos. En consecuencia, el sistema educativo nacional no solo se caracteriza por estar segmentado en su efectividad escolar por la vía del perfil sociocultural y económico de sus alumnos, sino que además también lo estaría respecto de los sistemas evaluativos internos que cada establecimiento implementa. Los modelos educativos actuales han privilegiado prácticas evaluativas de carácter reproductivo, de control instrumental y memorístico (Stiggins, 2004), configurándose como un proceso desvinculado a la enseñanza (Celman, 2005).

En este contexto, las notas de la enseñanza media (NEM) en Chile, a pesar de contener información relevante pues, por un lado, resumen en un único valor las apreciaciones evaluativas de un alumno hechas por sus distintos profesores; y por otra, son un componente importante dentro del sistema de selección (CRUC, 2008; Donoso, 1998; Koljatic y Silva, 2006), actualmente están en tela de juicio respecto de su real capacidad predictiva de la trayectoria académica del estudiante y su idoneidad como factor de ponderación en los sistemas de selección universitaria. De esta manera, aunque las NEM han evidenciado ser un complemento relevante en el proceso de admisión en la educación superior (Rodríguez, Fita, y Torrado, 2004), tienden a ser el reflejo de distintas políticas de asignación de notas y de diferenciados niveles de exigencia e inflación de tales notas (Contreras, Gallegos y Meneses, 2009). En este sentido, Prieto y Contreras (2008:246), afirman:

No es posible considerar la enseñanza y la evaluación como procesos neutros o inocuos, no solo porque influyen en la calidad de

los procesos formativos y producen efectos en los estudiantes que les impactan de manera crítica sino que, dado que es una actividad esencialmente subjetiva y valorativa, constituyen un reflejo de las concepciones de los profesores, entre las que se encuentran las creencias o conocimientos especializados.

En esta línea de argumentación, es posible advertir la presencia del concepto de currículum cuando intentamos aproximarnos a medir de qué manera lo que se enseña en la escuela tiene consecuencias posteriores en procesos de selección universitaria. Al respecto, Loureiro (2009:11) analiza las relaciones entre la evaluación en el aula, el currículum y las evaluaciones externas, y señala:

[...] Resulta necesario, en primera instancia, distinguir entre currículum prescrito y currículum implementado. A partir de dicha distinción, cabe preguntarse qué relación existe entre el currículum prescrito por los Ministerios de Educación y el currículum realmente enseñado en las aulas. Inevitablemente, siempre existe una distancia entre ambos, dado que la tarea de enseñar exige siempre hacer algún tipo de selección de contenidos y actividades, que puede ser de mayor o menor magnitud.

Como podemos apreciar, las NEM están vinculadas al currículum y a la efectividad en el aula, donde el papel del profesor resulta un factor relevante (Brunner y Elacqua, 2004). Reafirmando lo anterior, Meckes (2007) señala que hoy existe abundante literatura y experiencias que demuestran cómo la evaluación de aula apropiada e integrada en el proceso de enseñanza y aprendizaje da cuenta de progresos notables en los resultados alcanzados por los alumnos. No obstante, desde una visión más crítica, Ravela (2010:6), plantea:

No siempre los profesores evalúan de manera apropiada y justa. Asimismo, en un país puede existir una gran heterogeneidad en los niveles de exigencia de las evaluaciones elaboradas a nivel local. Por eso muchos países optan por establecer exámenes nacionales al final de algunos niveles clave de la enseñanza, en especial al final de la educación media.

Si nos detenemos en la idea expresada anteriormente por Ravela, resulta concordante con las reflexiones de Vargas (2010), quien señala que las evaluaciones intraaula son muy distintas en naturaleza, extensión, contexto y finalidad, en comparación con una evaluación estandarizada como la prueba de selección universitaria (PSU). Vargas (2010) afirma que las evaluaciones que realizan los profesores tienen como ámbito de acción el aula y, como sujeto de evaluación, al estudiante individual, mientras que las evaluaciones estandarizadas se hacen a escala mayor (comunal, regional o nacional) y tienen como sujeto de evaluación el sistema educativo en su conjunto. No obstante, afirma el autor, si el profesor ha calificado los logros de cada alumno en diversos tramos, aspectos o contenidos del currículum implementado, debería haber una mínima correspondencia con el currículum prescrito y, por lo tanto, con los puntajes obtenidos en la PSU que mide los contenidos del respectivo subsector.

Lo expresado nos permite coincidir con Manzi, Bosch, Bravo, Pino, Donoso y Pizarro (2010:47) que explicitan en forma categórica su preocupación por las NEM, señalando que es necesario «[...] continuar homogeneizando las distribuciones de los puntajes de los factores de selección, para lo cual se requiere abordar el tema del promedio de notas de la enseñanza media». La afirmación precedente se funda en estudios acerca del decrecimiento de los índices de rendimiento en el logro de los objetivos de las diferentes asignaturas de la enseñanza media. Para Donoso (1998), la reducción del rango o recorrido de los promedios implica una disminución de la varianza, con los consiguientes efectos sobre la capacidad predictiva que pueda tener la variable. En efecto, según el autor, se ha logrado advertir una progresiva reducción del rango de las calificaciones por elevación de los valores inferiores, lo que

**Tabla 1**  
Descripción de establecimientos participantes

	N.º	%
<i>Dependencia administrativa</i>		
Municipal	1.550	37,8
Particular subvencionado	2.150	52,4
Particular pagado	404	9,8
<i>Rama de enseñanza</i>		
Humanista científica diurna	2.001	48,8
Humanista científica nocturna	664	16,2
Técnico profesional	1.439	35,1
<i>Grupo socioeconómico</i>		
A (bajo)	979	27,7
B (medio-bajo)	1.123	31,7
C (medio)	680	19,2
D (medio-alto)	405	11,4
E (alto)	351	9,9

Fuente: DEMRE (2012). Elaboración propia de los autores.

ha impactado sobre la capacidad predictiva en el caso de las NEM. Manzi et al. (2010) explican este fenómeno por la incorporación al proceso de selección universitaria de estudiantes provenientes de los más bajos quintiles de ingreso.

Sobre los actuales resultados en los procesos de enseñanza en la educación media en Chile, Molina (2008:120), plantea:

En este contexto de análisis surge la siguiente paradoja: si la enseñanza media tiene como principal sentido el «prepararse para el futuro», sus actuales niveles de logro académico amenazan seriamente el cumplimiento de dicho sentido primario destacado por los jóvenes estudiantes de liceos municipales. En la perspectiva de superar esta paradoja, el liceo municipal tendría entre sus principales desafíos la tarea de (re)construir el carisma y la fuerza movilizadora que tuvo la enseñanza pública en momentos claves del desarrollo de la sociedad chilena.

En suma, Chile instala un sistema estandarizado fundado en la necesidad de asegurar que los alumnos que ingresan en la universidad sean los mejor dotados, esto es, con las capacidades requeridas para aprovechar adecuadamente la formación a recibir y para obtener buenos resultados académicos (Donoso y Schiefelbein, 2007). Sin embargo, el aumento artificioso de las calificaciones en la enseñanza media, la reducción del recorrido de los promedios, con la consiguiente disminución de la varianza, tendría efectos sobre la capacidad predictiva que puede ser atribuida a las NEM. Estos efectos se observan en la posibilidad de determinar una probabilidad anticipada de los puntajes a obtener en la PSU, tanto como en el futuro éxito académico que dicho estudiante registrará en la educación superior. En consecuencia y en específica observancia a este primer efecto enunciado de las NEM, en esta investigación se busca configurar la capacidad predictiva que tiene la evaluación escolar docente en la determinación del nivel de logro en pruebas estandarizadas en el sistema chileno de selección universitaria.

## Método

### Participantes

La muestra está constituida por un total de 4.104 establecimientos educacionales de enseñanza media cuyo alumnado rindió la PSU durante el proceso de admisión 2012, con independencia de sus diversas ramas y tipo de establecimiento (tabla 1).

El sistema educativo chileno distingue 3 tipos de establecimientos según el tipo de establecimiento. Los establecimientos municipales son entidades públicas en su propiedad, infraestructura y financiamiento, cuya administración ha sido delegada, a partir de la reforma de los años 80, casi en su totalidad al municipio de cada comuna. Los establecimientos particulares subvencionados son centros educativos de naturaleza privada pero que reciben

financiamiento estatal a través de una subvención o «voucher» por alumno atendido. Por último, están los centros educativos particulares pagados que se caracterizan por no percibir aportes del Estado, funcionando solo con los recursos proporcionados por los padres y apoderados a través del pago de aranceles. Con base a esta tipología administrativa, la muestra se caracteriza por la mayoritaria presencia de establecimientos particulares subvencionados: uno de cada 2 pertenecen a esta categoría (52,4%), seguidos de los de administración municipal (pública) con el 37,8%. La presencia de establecimientos particulares pagados alcanza una marginal presencia del 9,8%.

Por otra parte, en lo referente a la rama de enseñanza, el marco regulatorio distingue 3 modalidades principales de enseñanza en la educación media: la humanista científica diurna (HC-diurna), la humanista científica nocturna (HC-nocturna) y la técnico-profesional (TP). Los establecimientos de educación HC-diurna (48,8% de la muestra) tienen como propósito la preparación del estudiante para la continuidad de estudios en la educación superior, mientras que los centros de educación HC-nocturna (16,2% de la muestra) centran su accionar en la nivelación de estudios cuyo objetivo final es obtener la licencia de la enseñanza media. Por último, la educación TP tiene como finalidad la formación para el mundo del trabajo, proporcionando un título profesional técnico de nivel medio, aunque no es excluyente la continuidad de estudios superiores. Estos centros representan el 35,1% de la muestra.

Atendiendo al criterio de segmentación socioeducativa, los centros educativos se estratifican en 5 grupos con base en el perfil socioeconómico del alumnado, tipología que por lo demás es la utilizada por el Ministerio de Educación de Chile para diferenciar a los colegios. Esta estratificación en lo específico contempla los ingresos familiares, el nivel educativo de los padres y el índice de vulnerabilidad escolar del establecimiento. De acuerdo a esta clasificación, la mayor parte de los colegios de la muestra pertenecen a grupos A y B (59,4%), es decir, sus padres tienen menos de 10 años de escolaridad, poseen bajos e inestables ingresos familiares y al menos el 40% de los estudiantes se encuentran en condición de vulnerabilidad social. Por otra parte, los colegios que atienden al alumnado de mayor nivel socioeconómico, y con independencia de su tipología administrativa, alcanzan el 9,9%.

### Instrumentos

Para el desarrollo de la investigación se hizo uso de la base de datos del Proceso de Admisión 2012, proporcionada por el Departamento de Evaluación, Medición y Registro Académico de la Universidad de Chile, organismo público a cargo de la administración de la PSU. Esta prueba es el mecanismo de selección que se rinde desde el año 2003 en Chile y que vino a reemplazar a la antigua prueba de aptitud académica (PAA). La PSU es una prueba que pretende constatar los niveles de conocimiento vinculados a los contenidos de los planes y programas definidos por el Ministerio de Educación y que configuran el marco curricular vigente. Consta de 4 pruebas, 2 de ellas obligatorias (Lenguaje y Comunicación; Matemáticas) y 2 optativas (Historia y Ciencias Sociales; Ciencias). Los puntajes de estas pruebas, una vez corregido el azar, son normalizados a una escala común de promedio de 500 puntos y desviación estándar de 100 puntos. En general, los puntajes transitan en una escala que va desde 150 hasta 850 puntos.

Las NEM corresponden al promedio de las notas de enseñanza media y son expresión del rendimiento académico del alumno durante los últimos 4 años de estudio obligatorios en Chile. Estas notas son convertidas, dependiendo de la rama de enseñanza, a puntuación estándar PSU. Ambos componentes (puntajes PSU y NEM) son factores que, en ponderaciones variables, dependiendo de la universidad o carrera, definen el puntaje de postulación a una determinada carrera de educación superior. En la actualidad,

este sistema de selección es utilizado por las 25 universidades que conforman el Consejo de Rectores (universidades tradicionales) y, desde 2012, por 8 universidades privadas o no tradicionales.

### Procedimiento

La información utilizada corresponde a bases de datos de acceso público amparado en la Ley de Transparencia que rige este tipo de requerimientos y disponible en: [www.demre.cl/estadisticasP2012.htm](http://www.demre.cl/estadisticasP2012.htm). Por tanto, se trata de información oficial que contempla la totalidad de los establecimientos cuyos estudiantes participan del proceso de admisión universitaria a nivel nacional. Esta base de datos registra información de los resultados de las pruebas de admisión universitaria del año 2012 expresada en puntaje estándar (normalizado de media 500 y desviación estándar 100), específicamente, los de las pruebas de Matemática, Lenguaje, Historia y Ciencias. Por tratarse de pruebas externas nacionales, son test sometidos a procesos de validación o medidas de consistencia interna, que arrojan un alfa de Cronbach de 0,94 y un error de medida promedio de 4,3 (Donoso y Contreras, 2009).

Además de los puntajes PSU, la base de datos consigna información respecto del rendimiento escolar del alumno (NEM) que expresa la trayectoria del alumno homologado a puntaje PSU mediante una escala de conversión, a fin de hacerlo funcionalmente equivalente.

A esta información se incorpora el tipo de colegio según la rama de enseñanza (HC-diurno, HC-nocturno, TP), tipo de establecimiento (municipal, particular subvencionado y particular pagado) y estratificación del establecimiento en razón al perfil socioeducativo del alumnado, distinguiéndose 5 grupos jerárquicos.

### Análisis de datos

El estudio sigue un diseño estadístico de tipo correlacional en la medida que busca vincular la relación estadística existente entre el rendimiento escolar (NEM) y el puntaje PSU de cada una de las pruebas. La búsqueda de esta relación se basa en el supuesto de que estos test estandarizados aplicados a nivel nacional están teóricamente alineados con el currículum escolar, donde las NEM expresarían el nivel de aprovechamiento que de dicho currículum ha hecho el alumno en su trayectoria escolar. Este análisis se segmenta por rama de enseñanza, tipo de establecimiento y estratificación socioeducativa del establecimiento, a fin de observar el comportamiento en los índices de correlación en el interior de cada conglomerado o segmento generado, así como comparar sus diferenciales interconglomerados.

**Tabla 2**  
Puntajes de la prueba de selección universitaria y notas de enseñanza media según rama de enseñanza y dependencia administrativa, proceso de admisión 2012

Rama	Dependencia	N	Lenguaje		Matemáticas		Historia		Ciencias		NEM	
			Media	D.T.	Media	D.T.	Media	D.T.	Media	D.T.	Media	D.T.
HC-diurna	Municipal	494	448,3 <sup>a</sup>	54,6	453,4 <sup>a</sup>	53,2	457,2 <sup>a</sup>	56,3	440,6 <sup>a</sup>	55,3	527,4	46,9
	Part. subv.	1.135	508,5 <sup>a</sup>	55,1	511,0 <sup>a</sup>	57,9	514,4 <sup>a</sup>	54,6	497,0 <sup>a</sup>	56,3	537,1	45,1
	Part. pagado	372	596,3	62,3	605,8 <sup>a</sup>	64,7	596,0	60,0	590,5	65,6	590,8	56,0
Técnico profesional	Municipal	757	403,4 <sup>a</sup>	45,1	406,8 <sup>a</sup>	39,4	417,6 <sup>a</sup>	46,7	399,4 <sup>a</sup>	49,1	511,0	47,3
	Part. subv.	681	423,1 <sup>a</sup>	39,7	420,4 <sup>a</sup>	39,6	435,8 <sup>a</sup>	41,7	411,0 <sup>a</sup>	46,7	502,5	44,5
HC-nocturna	Municipal	299	392,4 <sup>a</sup>	44,8	386,0 <sup>a</sup>	36,3	428,0 <sup>a</sup>	49,3	388,4 <sup>a</sup>	45,6	510,5	54,8
	Part. subv.	334	404,4 <sup>a</sup>	37,8	390,6 <sup>a</sup>	36,1	431,4 <sup>a</sup>	40,9	392,1 <sup>a</sup>	39,4	498,9	49,0
	Part. pagado	31	450,3 <sup>a</sup>	29,6	446,5 <sup>a</sup>	33,4	464,4 <sup>a</sup>	30,4	428,2 <sup>a</sup>	32,2	487,2	41,5
General	Municipal	1.550	415,6 <sup>a</sup>	53,3	417,7 <sup>a</sup>	50,7	432,4 <sup>a</sup>	53,4	411,8 <sup>a</sup>	55,1	516,2	49,2
	Part. subv.	2.150	465,3 <sup>a</sup>	66,7	463,6 <sup>a</sup>	71,2	476,8 <sup>a</sup>	63,1	455,0 <sup>a</sup>	69,1	520,4	48,9
	Part. pagado	404	584,8	71,9	593,3 <sup>a</sup>	75,8	585,7	67,9	577,9	77,0	583,1	61,7

HC: humanista científica; N: número de establecimientos; NEM: notas de enseñanza media; Part.: particular; subv.: subvencionado.  
Fuente: DEMRE (2012). Elaboración propia de los autores.

<sup>a</sup> Significación estadística al nivel 0,01 bilateral de cada prueba respecto de NEM.

En primera instancia se hace una comparación de las medias de los puntajes de cada uno de los test de la PSU con respecto a las NEM, utilizando la prueba t de Student para muestras independientes, como manera de establecer si las diferencias observadas entre estas medias son estadísticamente significativas en el interior de cada conglomerado. No obstante tratarse de una prueba paramétrica, y que por consiguiente exige normalidad, esta prueba ha demostrado ser sólida, a falta de ella en muestras suficientemente grandes, tal como es el caso en estudio. A fin de controlar la debilidad que presenta esta prueba y su significación en muestras muy grandes, se hace uso del d de Cohen para determinar el tamaño del efecto, esto es, la magnitud de las diferencias que resultaron estadísticamente significativas. En segunda instancia se obtienen los coeficientes de correlación paramétricos r de Pearson, analizando su significación, orientación y magnitud. La relevancia de la correlación de la asociación se expresa también a través de sus respectivos coeficientes de determinación ( $R^2$ ) cuyo análisis es expresado en términos de porcentajes de variabilidad explicada.

Por último, en atención a que las correlaciones dentro del espacio bivalente pueden eventualmente presentar las perturbaciones de variables externas no contempladas, y cuyo accionar sobredimensiona o subdimensiona la fuerza o magnitud del índice de correlación general o primario ( $r_{xy}$  de Pearson), se hace uso del índice de correlación parcial ( $R_{par}$  o  $R_{yx1.X2}$ ). Este índice permitirá pronunciarse sobre las magnitudes de las asociaciones observadas, descontando el sesgo o efecto que presentan los factores, tipo de establecimiento, grupo socioeconómico y rama de enseñanza.

### Resultados

Como se puede apreciar al observar la [tabla 2](#), al comparar los resultados de los establecimientos en los distintos test de la PSU, se observa que en todos y cada uno de los componentes los mayores puntajes provienen del sector particular pagado y de la rama HC-diurna. Si bien los resultados de las distintas pruebas dentro de los colegios particulares pagados no manifiestan grandes diferencias, los rendimientos tienden a ser levemente superiores en la prueba de Matemáticas, donde el promedio bordea los 593 puntos. Por otra parte, los mejores resultados en la rama HC-diurna están en sintonía con su objetivo intrínseco de continuación de estudios y el mayor énfasis otorgado a la preparación de la PSU. Por su parte, la rama TP y HC-nocturna procuran la formación para la vida laboral y la regulación de estudios respectivamente. Así, las brechas se hacen especialmente evidentes en las pruebas de Matemáticas y Lenguaje, donde los colegios particulares pagados superan a la educación municipal en 175 puntos en Matemáticas y 169 puntos en

**Tabla 3**

Coefficientes d de Cohen de tamaño del efecto de medias de cada test de la prueba de selección universitaria con las notas de enseñanza media según rama de enseñanza y dependencia administrativa, proceso de admisión 2012

Rama	Dependencia	N	Lenguaje d de Cohen	Matemáticas d de Cohen	Historia d de Cohen	Ciencias d de Cohen
HC-diurna	Municipal	494	1,55	1,48	1,35	1,69
	Part. subv.	1.135	0,57	0,50	0,45	0,79
	Part. pagado	372	-0,09	-0,25	-0,09	0,0
Técnico profesional	Municipal	757	2,33	2,39	1,99	2,31
	Part. subv.	681	1,88	1,95	1,55	2,01
HC-nocturna	Municipal	299	2,36	2,68	1,58	2,42
	Part. subv.	334	2,16	2,52	1,50	2,40
	Part. pagado	31	1,02	1,08	0,63	1,59
General	Municipal	1.550	1,96	1,97	1,63	2,00
	Part. subv.	2.150	0,94	0,93	0,77	1,09
	Part. pagado	404	-0,03	-0,15	-0,04	0,07

HC: humanista científica; N: número de establecimientos; Part.: particular; subv.: subvencionado.

Fuente: DEMRE (2012). Elaboración propia de los autores.

<sup>a</sup> Significación estadística al nivel 0,01 bilateral de cada prueba respecto de NEM.

Lenguaje, lo que representa un 42 y 41% superior, respectivamente (tabla 2).

No obstante lo anterior y si bien los puntajes en cada una de las pruebas favorecen al sector particular pagado, los puntajes alcanzados son bajos en la generalidad de los test. Lo anterior revela las dificultades que evidencia el sistema educativo en su capacidad para instalar competencias cognitivas y procedimentales en los distintos sectores de la enseñanza que son evaluados. Así por ejemplo, al observar los promedios que registran las unidades educativas en cada una de las pruebas, se constata que el 70% de ellos no supera los 500 puntos estándar PSU. Para el caso de Matemáticas, estos 500 puntos en términos de puntaje corregido, entendido como el número de respuestas correctas obtenidas en el test y una vez descontado el azar, no superan las 15 de un total de 75 preguntas, lo que representa un escaso 18% de nivel de logro.

Este mejor desempeño de los establecimientos del sector particular pagado en las PSU se constata también en los diferenciales de logro escolar o aprovechamiento educativo expresado en el índice NEM. No obstante ello, las brechas de las NEM se reducen significativamente entre tipos de establecimientos y son menos ostensibles que las evidenciadas por tipo de prueba, especialmente cuando se comparan los establecimientos municipales con los particulares pagados, observándose diferencias estadísticamente significativas entre las puntuaciones NEM-PSU ( $p < 0,01$ ). Así por ejemplo, las brechas NEM de un colegio particular pagado respecto de uno municipal es de 64 puntos promedio en la modalidad HC-diurna, diferencial que alcanza los 148 puntos en la prueba de Lenguaje y 153 puntos en la prueba de Matemática al comparar los mismos tipos de colegios. En la medida en que se trabaja con datos poblacionales y estandarizados, es posible comparar las puntuaciones o brechas y observar que el diferencial NEM se reduce en más de la mitad respecto del existente en Matemática y en Lenguaje (58 y 57% respectivamente).

En establecimientos municipales, con independencia de la rama de enseñanza, las diferencias NEM-Matemática/Lenguaje bordean los 100 puntos, lo que representa un puntaje asignado a las NEM de un 25% superior en comparación con el puntaje promedio PSU. Al momento de comparar el tamaño de efecto o cambio evidenciado como estadísticamente significativo, se constata que en establecimientos municipales HC-diurnos en las pruebas de Lenguaje y Matemática, un 94 y un 93% respectivamente de los casos, los establecimientos otorgan un puntaje NEM superior al que obtiene un alumno promedio en dichas pruebas. Se observa, además, una significativa reducción de la varianza y del rango o recorrido de la variable NEM. En los colegios particulares pagados, en cambio, las brechas NEM-PSU son irrelevantes y en algunos casos, negativas, transitando este diferencial entre 5 y 10 puntos, según se trate del

test de Matemática o Lenguaje. Al comparar el tamaño del efecto NEM-PSU en este mismo segmento particular pagado, los índices d de Cohen son 0,09 en Lenguaje y 0,25 en Matemática. En cambio en los colegios municipales estos coeficientes son 1,55 y 1,48 respectivamente (ver tabla 3).

Según los datos aportados por la tabla 4, llaman la atención las moderadas correlaciones constatadas entre las NEM y cada una de las evaluaciones contempladas en la PSU, las cuales transitan desde un 0,51 hasta un 0,57 para la generalidad del sistema. Al comparar estas correlaciones por tipo de establecimiento de los centros educativos, se evidencian comportamientos diferenciales en los índices. Mientras que en el sector particular pagado los índices de correlación r de Pearson fluctúan entre un 0,78 (Historia) y un 0,83 (Lenguaje), en su homólogo municipal estos mismos índices fluctúan entre un 0,18 (Ciencias) y un 0,29 (Matemáticas). Esto implica que para el segmento particular pagado las NEM presentan una alta capacidad predictiva del rendimiento PSU pues, en términos de coeficiente de determinación ( $R^2$ ) el 67% de la varianza de la PSU del colegio es explicada por las NEM obtenidas por su alumnado; como contraparte en el sector municipal, las NEM tienen muy baja capacidad predictiva del rendimiento PSU, la cual no alcanza el 8% de la variabilidad explicada.

Por otro lado, el comportamiento de las correlaciones NEM-PSU en los distintos tipos de establecimientos, al momento de controlar por rama de enseñanza, registra fluctuaciones relevantes. En el régimen HC-diurno los índices producto-momento de Pearson presentan tendencias similares a las observadas en la generalidad del sistema, destacando la fuerte correlación NEM-PSU que evidencian las unidades de dependencia particular pagada en comparación con los otros tipos de establecimiento. Cabe señalar que el grueso de la matrícula de los establecimientos particulares pagados está en la rama HC-diurna, que está orientada curricularmente a la continuación de estudios superiores.

En el sector municipal, para la misma rama HC-diurna, estas correlaciones siguen estando muy por debajo de las evidenciadas en el particular pagado. A modo de ejemplo, en el sector municipal los puntajes de la prueba de Lenguaje correlacionan un 0,27 en la generalidad del sistema, índice que sube a un 0,41 al momento de considerar solo la rama HC-diurna municipal, es decir, evidencia una mejora en su capacidad predictiva (7,3 vs. 11%), no obstante continuar muy por debajo de la observada en su homólogo particular pagado. En efecto, en el sector particular pagado los índices de correlación en Lenguaje transitan de un 0,83 a un 0,81 según se evalúe la generalidad del sistema o solo la rama HC-diurna, respectivamente. Igualmente altas y significativas son las correlaciones NEM-Matemática (tabla 4).

**Tabla 4**  
Coeficiente de correlación de cada test de la prueba de selección universitaria (PSU) con las notas de enseñanza media (NEM), según rama de enseñanza y dependencia administrativa, durante el proceso de admisión 2012

Rama	Dependencia	N	Lenguaje	Matemática	Historia	Ciencias
HC-diurna	Municipal	494	0,41 <sup>a</sup>	0,49 <sup>a</sup>	0,33 <sup>a</sup>	0,38 <sup>a</sup>
	Part. subv.	1.135	0,61 <sup>a</sup>	0,64 <sup>a</sup>	0,50 <sup>a</sup>	0,59 <sup>a</sup>
	Part. pagado	372	0,81 <sup>a</sup>	0,80 <sup>a</sup>	0,73 <sup>a</sup>	0,78 <sup>a</sup>
Técnico-profesional	Municipal	757	0,21 <sup>a</sup>	0,16 <sup>a</sup>	0,11 <sup>a</sup>	0,01
	Part. subv.	681	0,18 <sup>a</sup>	0,20 <sup>a</sup>	0,09 <sup>b</sup>	0,06
HC-nocturna	Municipal	299	-0,07	-0,05	-0,01	-0,09
	Part. subv.	334	-0,08	-0,17 <sup>a</sup>	-0,03	-0,05
	Part. pagado	31	0,11	0,02	0,18	0,22
General	Municipal	1.522	0,27 <sup>a</sup>	0,29 <sup>a</sup>	0,20 <sup>a</sup>	0,18 <sup>a</sup>
	Part. subv.	2.131	0,52 <sup>a</sup>	0,53 <sup>a</sup>	0,46 <sup>a</sup>	0,50 <sup>a</sup>
	Part. pagado	404	0,83 <sup>a</sup>	0,82 <sup>a</sup>	0,78 <sup>a</sup>	0,82 <sup>a</sup>
Total del sistema		4.047	0,56 <sup>a</sup>	0,57 <sup>a</sup>	0,51 <sup>a</sup>	0,54 <sup>a</sup>

HC: humanista científica; N: número de establecimientos; NEM: notas de enseñanza media; Part.: particular; subv.: subvencionado.

Fuente: DEMRE (2012). Elaboración propia de los autores.

<sup>a</sup> La correlación es significativa al nivel 0,01 (bilateral).

<sup>b</sup> La correlación es significativa al nivel 0,05 (bilateral).

**Tabla 5**  
Correlaciones bivariadas y parciales entre las notas de enseñanza media y las pruebas de selección universitaria (PSU), durante el proceso de admisión 2012

	HC-diurna y técnico profesional				HC-diurna		
	r de Pearson	Correlación parcial según variable de control			r de Pearson	Correlación parcial según variable de control	
		GSE	Dependencia	Rama		GSE	Dependencia
NEM-Lenguaje	0,61 <sup>a</sup>	0,45 <sup>a</sup>	0,57 <sup>a</sup>	0,55 <sup>a</sup>	0,68 <sup>a</sup>	0,55 <sup>a</sup>	0,61 <sup>a</sup>
NEM-Matemática	0,63 <sup>a</sup>	0,49 <sup>a</sup>	0,59 <sup>a</sup>	0,58 <sup>a</sup>	0,71 <sup>a</sup>	0,59 <sup>a</sup>	0,65 <sup>a</sup>
NEM-Historia	0,55 <sup>a</sup>	0,33 <sup>a</sup>	0,48 <sup>a</sup>	0,47 <sup>a</sup>	0,61 <sup>a</sup>	0,41 <sup>a</sup>	0,51 <sup>a</sup>
NEM-Ciencias	0,58 <sup>a</sup>	0,36 <sup>a</sup>	0,51 <sup>a</sup>	0,50 <sup>a</sup>	0,67 <sup>a</sup>	0,52 <sup>a</sup>	0,60 <sup>a</sup>

GSE: grupo socioeconómico; HC-diurna: humanista científica diurna; NEM: notas de enseñanza media.

Fuente: DEMRE (2012). Elaboración propia de los autores.

<sup>a</sup> La correlación es significativa al nivel 0,01 (bilateral).

La **tabla 5** expone las correlaciones bivariadas y parciales, en primera instancia, del conjunto de establecimientos HC-diurnos y TP y, en segundo lugar, tan solo a los establecimientos HC-diurnos. La exclusión de los establecimientos de la educación de adultos (HC-nocturno) en este análisis se basa en el hecho de que ellos evidencian una presencia marginal, un bajo nivel de matrícula comparada, además de no estar orientados expresamente a la continuidad de estudios, sino más bien a la regulación o nivelación.

Respecto de las correlaciones bivariadas (r de Pearson) se observa una menor relación de las NEM con PSU-Historia, así como que las correlaciones se moderan a la baja cuando se incluye al segmento de la rama TP. En efecto, en el grupo HC-diurno y TP las correlaciones transitan entre un 0,55 y un 0,61 según el tipo de prueba que se correlacione con NEM. Mientras que cuando se reduce la muestra tan solo al grupo HC-diurno, estas mejoran significativamente, transitando entre un 0,61 (NEM-PSU Historia) y un 0,71 (NEM-PSU Matemática).

En atención a que las correlaciones en el espacio bivalente pueden estar perturbadas por la presencia de variables externas o extrañas, se hace uso de correlaciones parciales. Este índice permite pronunciarse sobre la concomitancia observada entre las variables NEM-PSU una vez descontada la influencia de factores eventualmente asociados, tales como el tipo de establecimiento, grupo socioeconómico del alumnado y rama de enseñanza.

En este contexto de correlaciones parciales, se observa que para ambos grupos los índices son de menor magnitud al momento de controlar la influencia de variables externas, en todas y cada una de las pruebas en su relación con NEM. A modo de ejemplo, en los establecimientos HC-diurnos, el índice Pearson para NEM-PSU Matemática es de un 0,71, este índice baja a un 0,65 y 0,59 según se controle el efecto del tipo de establecimiento o el grupo socioeconómico. Especialmente relevantes son las diferencias al momento

de controlar la alteración generada por el grupo socioeconómico. Por consiguiente, los índices de correlación parcial para el mismo conglomerado HC-diurno transitan desde un 0,41 hasta un 0,55 dependiente del tipo de prueba con que se relacione NEM.

## Discusión

El artículo tiene como objetivo determinar los niveles de correlación (r de Pearson y correlación parcial) como un acercamiento para estimar la capacidad predictiva que evidencian las NEM respecto de los puntajes alcanzados en los test PSU por estudiantes chilenos, tanto en la generalidad del sistema como en el interior de cada conglomerado con base en variables de segmentación (rama de enseñanza, tipo de establecimiento y grupo socioeconómico). Con el índice de correlación parcial se busca reducir o controlar el sesgo de magnitud que eventualmente presentarían estas variables de segmentación en la estimación de la relación NEM-PSU. La hipótesis que orientó el estudio fue que las NEM, como expresión del nivel de aprovechamiento o logro del alumnado en sus respectivos colegios, evidencian una moderada capacidad de predecir el rendimiento en las PSU. Capacidad que, además, se encuentra segmentada con base en la dependencia administrativa del establecimiento (tipo de colegio) y grupo socioeconómico del alumnado al que atiende.

La PSU representa un instrumento de evaluación educacional que mide la capacidad de razonamiento de los postulantes egresados de la enseñanza media, teniendo como referencia los contenidos del Plan de Formación General de Lenguaje y Comunicación, Matemática, Historia y Ciencias Sociales, y de Ciencias Naturales. Así entendido, los resultados PSU debieran estar lógicamente en sintonía con la evaluación escolar de los

últimos 4 años de enseñanza, toda vez que, si bien son procesos independientes, ambos tienen por finalidad constatar los niveles de aprovechamiento educativo del alumnado y, en teoría, versan sobre un mismo campo de contenidos. La importancia de establecer la relación NEM-PSU se funda también en el hecho de que el sistema de admisión universitaria para continuación de estudios no opera únicamente sobre los resultados de la PSU, sino que además considera la trayectoria escolar del alumno durante la enseñanza media, operacionalizado como ranking, esto es, rango percentil o posición relativa que ocupa el alumno entre sus pares del establecimiento, así como el promedio de NEM. En términos generales, uno de cada 2 puntos PSU ponderado derivan de estos conceptos o factores.

Al analizar los datos, se observa que los puntajes PSU son más altos en la rama HC-diurna y en los colegios de procedencia particular pagada, con diferencias especialmente notorias en la prueba de Matemática. La mayor eficiencia de este tipo de establecimiento es una constante histórica que se evidenciaba, incluso, durante la vigencia de la PAA (CRUCH, 2010). No obstante ello, los puntajes alcanzados son bajos en la generalidad del sistema, con independencia relativa del tipo de test. Alrededor del 70% del alumnado que rinde la PSU no alcanza los 500 puntos, puntaje que equivale a menos del 20% de respuestas correctas en la PSU-Matemática. Dicha situación ilustra las dificultades que evidencia el sistema escolar en cuanto a su capacidad para instalar competencias educativas en su alumnado y prepararlo adecuadamente para la continuidad de estudios superiores, en especial en los alumnos más vulnerables.

Las diferencias se hacen extensibles también al comparar las puntuaciones NEM, aunque aquí los diferenciales se observan ostensiblemente más atenuados. Las brechas NEM son comparativamente menores con independencia de la rama de enseñanza y el tipo de establecimiento. Respecto de las mismas NEM, no se observan tan solo reducciones en las brechas entre los establecimientos sino también un aumento comparativo de los puntajes de homologación y una reducción de la desviación estándar. Los puntajes asignados por concepto de NEM son en promedio superiores a los obtenidos en cualquier test de selección, situación especialmente marcada en los establecimientos municipales, mientras que en los colegios particulares pagados este diferencial se hace irrelevante. Se constata, además, una significativa reducción de la varianza y del rango o recorrido de sus promedios, con los consiguientes efectos que esta concentración tendría sobre la capacidad predictiva de la variable. En efecto, se evidencia una progresiva reducción del rango de las calificaciones por elevación de los valores inferiores, situación especialmente notoria en los colegios públicos, expresión de las diferenciadas políticas evaluativas y de la pérdida de sintonía que tienen las NEM como expresión de la asimilación del currículum. Similares hallazgos son constatados por Díaz, Himmel y Maltes (1990).

Si bien lo observado es una problemática de larga data, se puede evidenciar su incremento progresivo con la instalación del sistema de selección universitaria y la mayor preponderancia asignada a las trayectorias escolares. En términos de puntajes, es expresión de la segmentada calidad de los procesos formativos; y, en cuanto a NEM, manifiesta la debilidad de los sistemas evaluativos que implementan los colegios, en la búsqueda de que las evaluaciones sean reflejo del real aprovechamiento educativo y dominio del currículum que tiene el alumno. Situación, esta última, que hace perder la eventual capacidad predictiva que en teoría tendrían las NEM, por cuanto los colegios abultarían las calificaciones de sus propios alumnos, provocando distorsión e inequidad en el interior del sistema de selección universitaria. Esta situación implica tomar con cautela la excesiva ponderación que a este factor (NEM) se le asigna en los procesos de selección por parte de las instituciones de educación superior.

Este análisis es refrendado al momento de revisar las matrices de correlaciones. Los índices de correlaciones producto-momento

de Pearson son de mayor magnitud entre pruebas que en la vinculación NEM-PSU, tanto en la generalidad del sistema como en sus ramas y dependencias. Especialmente bajas son las correlaciones NEM-PSU en el sector municipal y en las ramas HC-nocturna y TP. Sectores estos últimos que no tienen como propósito explícito la continuación de estudios superiores, por lo que la evaluación escolar pudiera estar orientada a medir otro tipo de competencias o con otros niveles de exigencia a los requeridos en la PSU y en los planes y programas de la rama HC-diurna.

Los establecimientos municipales presentarían profundas brechas en los coeficientes de correlación NEM-PSU, reduciendo su capacidad predictiva al comparar con establecimientos particulares pagados. En apariencia se estaría frente a la existencia de sistemas evaluativos más confiables y fidedignos en los establecimientos particulares, especialmente en los de carácter privado. No obstante, esta afirmación debe ser tomada con cautela en la medida en que la pérdida de validez de las NEM como valor de resumen único de las apreciaciones evaluativas docentes obtenidas por el alumno están más asociadas a factores socioeconómicos del alumnado que a la dependencia administrativa de los establecimientos, tal como se observa en las matrices de correlaciones parciales (tabla 5). Esto pone en entredicho la aparente mayor efectividad del sector privado en la instalación de competencias educativas en sus alumnos. Más bien, estos logros se explicarían por los mecanismos de selección de estudiantes que han implementado los establecimientos particulares en la provisión de servicios educativos y que tiene como corolario el proceso de segmentación educativa que se ha instalado en Chile desde la reforma educativa de los años 80. Dicha selectividad y segmentación termina por asignar una importancia desmedida al efecto de factores exógenos a la escuela en la definición de las trayectorias educativas exitosas y en los mejores niveles de aprovechamiento educativo observado en las evaluaciones externas, tanto por el Sistema de Medición de la Calidad de la Educación como por el PSU. Este fenómeno, además, transfigura el rol movilizador primigeniamente atribuido a la escuela, convirtiéndola en un ente de mantención o reproducción de las desigualdades de origen. En efecto, la segmentación educativa y los mecanismos de selectividad de los alumnos implementada por los colegios consigue mejorar la *performance* o condiciones de educabilidad de unos pocos en desmedro de las amplias mayorías, haciendo para esta elite que el acto educativo ofrezca menor complejidad, posibilitando mayores logros educativos en las evaluaciones externas como la PSU. Estos hallazgos son consistentes con lo observado por Atkinson y Geiser (2009); Hanushek, Kain, Markman y Rivkin (2002); Koljatic, Silva y Cofre, 2013; OCDE (2010); Manzi, Flotts y Preiss, 2012; Valenzuela, Bellei y De los Ríos, 2010.

Si consideramos el nivel de logro en la PSU, también aparecen como factores facilitadores la familiarización con el contenido y el entrenamiento en el formato de prueba estandarizada. Esta preparación puede recibirse directamente en el establecimiento educacional o realizarse de manera privada en entidades externas especializadas (institutos preuniversitarios), que otorgarían mayor ventaja a determinados alumnos, en especial de las familias más ricas, tal como lo plantea Zwick (2012).

En síntesis, los datos ilustran las dificultades que tiene el sistema educacional chileno en cuanto a instalar competencias y habilidades cognitivas en la generalidad de la enseñanza media, en especial en los estudiantes provenientes de los sectores más vulnerables o desfavorecidos socioculturalmente. Fenómeno de calidad segmentada en los aprendizajes constatados en los puntajes PSU y refrendados en evaluaciones externas del Sistema de Medición de la Calidad de la Educación. Problemática que se hace extensible a los sistemas de evaluación. Así, la evaluación que hace el docente presenta dificultades para dar cuenta del real aprovechamiento educativo que tienen los alumnos del currículum, obedeciendo a diferenciadas políticas evaluativas de segmentado rigor y exigencia.

Esta situación restringe la capacidad predictiva que teóricamente se le atribuye a las calificaciones como expresión del desempeño escolar, e instala un factor de tensión en el interior de un sistema que asigna cada vez mayor preponderancia a la trayectoria escolar del alumno en el proceso de selección universitaria.

### Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

### Referencias bibliográficas

- Atkinson, R. C. y Geiser, S. (2009). Reflections on a century of college admissions test. *Educational Researcher*, 38(9), 665–676.
- Brunner, J.J. y Elacqua, G. (2004). Factores que inciden en una educación efectiva. Evidencia internacional. La educ@ción, año XLIII-XLIX, 139-140. Organización de Estados Americanos (OEA).
- Celman, S. (2005). ¿Es posible mejorar la evaluación y trasformarla en una herramienta de conocimiento? In A. Camilloni, A. Camilloni, et al. (Eds.), *La evaluación de los aprendizajes en el debate didáctico contemporáneo*. (pp. 35–66). Buenos Aires: Paidós Educador.
- Contreras, D., Gallegos, S., y Meneses, F. (2009). Determinantes de desempeño universitario: ¿Importa la habilidad relativa? Análisis ex-ante y ex-post de una política pública. Informe final. Santiago de Chile: Consejo Nacional de Educación Superior.
- Consejo de Rectores de las Universidades Chilenas (2008). Estudio de la validez predictiva de los factores de selección a las universidades del Consejo de Rectores, admisiones 2003 al 2006. Santiago de Chile: Comité Técnico Asesor del CRUCH.
- Consejo de Rectores de las Universidades Chilenas (2010). Validez diferencial y sesgo de predictividad de las Pruebas de Admisión a las Universidades Chilenas. Documento técnico. Santiago de Chile: Comité Técnico Asesor del CRUCH.
- Departamento de Evaluación, Medición y Registro Educativo (DEMRE): Estadísticas de los Procesos de Admisión 2012. Santiago de Chile: Universidad de Chile [consultado el 24 Mar 2014]. Disponible en: [www.demre.cl/estadísticasP2012.htm](http://www.demre.cl/estadísticasP2012.htm)
- Díaz, E., Himmel, E., y Maltes, S. (1990). Evolución histórica del sistema de selección a las universidades chilenas, 1967-1989. En M. J. Lemaitre (Ed.), *La educación superior en Chile: un sistema de transición*. Santiago de Chile: Corporación de Promoción Universitaria (CPU), Colección Foro de la Educación Superior.
- Donoso, S. (1998). La reforma educacional y el sistema de selección de alumnos a las universidades: impactos y cambios demandados. *Estudios Pedagógicos (Valdivia)*, 24, 7–30.
- Donoso, G., y Contreras, P. (2009). Estudio de confiabilidad de las pruebas de selección universitaria. Admisión 2009. Departamento de Evaluación, Medición y Registro Educativo (DEMRE), Unidad de Estudios e Investigación. Universidad de Chile.
- Donoso, S. y Schiefelbein, E. (2007). Análisis de los modelos explicativos de retención de estudiantes en la universidad: una visión desde la desigualdad social. *Estudios Pedagógicos (Valdivia)*, 33(1), 7–27.
- Guttman, A. (2001). *La educación democrática. Una teoría política de la educación*. Barcelona: Paidós.
- Hanushek, E., Kain, J., Markman, J. y Rivkin, C. (2002). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5), 527–544.
- Koljatic M. y Silva, M. (2006). Validación de la PSU: Comentarios al estudio acerca de la validez predictiva de los factores de selección a las universidades del Consejo de Rectores. *Estudios Públicos*, 104, 331–346.
- Koljatic, M., Silva, M. y Cofre, R. (2013). Achievement versus aptitude in college admissions: A cautionary note based on evidence from Chile. *International Journal of Educational Development*, 33(1), 106–115.
- Loureiro, G. (2009). Evaluación en el aula, currículo y evaluaciones externas. Partnership for Educational Revitalization in the Americas. (PREAL), Santiago de Chile. [consultado 12 Abr 2014]. Disponible en: [www.preal.org/Biblioteca.asp](http://www.preal.org/Biblioteca.asp)
- Manzi, J., Bosch, A., Bravo, D., Pino, G. D., Donoso, G. y Pizarro, R. (2010). Validez diferencial y sesgo en la predictividad de las pruebas de admisión en la universidad chilenas (PSU). *Revista Iberoamericana de Evaluación Educativa RIEE*, 3(2), 29–48.
- Manzi, J., Flotts, P. y Preiss, D. D. (2012). Design of a collegel-level test of written communication. Theoretical and methodological challenges. In E. L. Grigorenko, E. Mambrino, & D. D. Preiss (Eds.), *Writing: A mosaic of new perspectives* (pp. 385–398). New York, NY: Psychology Press.
- Meckes, L. (2007). Evaluación y estándares: Logros y desafíos para incrementar el impacto en calidad educativa. *Pensamiento educativo*, (40), 351–371.
- Molina Chávez, W. M. (2008). Sentidos de la enseñanza media desde la experiencia escolar de estudiantes de liceos municipales. *Estudios pedagógicos (Valdivia)*, 34(1), 105–122.
- OCDE (2004). Chile Revisión de las políticas Nacionales de Educación. París: Organización para la Cooperación y el Desarrollo, 305.
- OCDE (2010). PISA 2009 results: Overcoming social background–equity in learning opportunities and outcomes (volume II). [consultado 19 Dic 2013]. Disponible en: <http://dx.doi.org/10.1787/9789264091504>
- OCDE (2014). Informe PISA 2012. [consultado 05 Mar 2014]. Disponible en: [www.oecd.org/pisa/home/](http://www.oecd.org/pisa/home/)
- Organización de las Naciones Unidas para la Educación, I. C. (2012). Indicadores del desarrollo mundial. UNESCO: Instituto de Estadísticas.
- Prieto, M. y Contreras, G. (2008). Las concepciones que orientan las prácticas evaluativas de los profesores: un problema a develar. *Estudios pedagógicos*, 34(2), 245–262.
- Ravela, P. (2010). ¿Qué pueden aportar las evaluaciones estandarizadas a la evaluación en el aula? Documento técnico. Santiago de Chile: Programa de promoción de la reforma educativa en América Latina y el Caribe (PREAL), 1–25.
- Rodríguez, S., Fita, E. y Torrado, M. (2004). El rendimiento académico en la transición secundaria-universidad. *Revista de Educación*, (334), 391–414.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, (86), 22–27.
- Valenzuela, J., Bellei, C. y de los Ríos, D. (2010). Segregación escolar en Chile. In S. Martinic y G. Elacqua (Eds.), *¿Fin de ciclo? Cambios en la gobernanza del sistema educativo*. (pp. 209–229). UNESCO-Pontificia Universidad Católica de Chile, Santiago de Chile.
- Vargas, J. A. (2010). Calificación de aula versus puntajes PSU: Incongruencias con consecuencias. Diversia Educación y Sociedad. CIDPA: Valparaíso, 121–134.
- Zwick, R. (2012). The role of admissions test scores, socioeconomic status, and high school grades in predicting college achievement. *Pensamiento Educativo. Revista de Investigación Educativa Latinoamericana*, 49(2), 23–30.