

# BORDÓN

## Revista de Pedagogía



Volumen 71  
Número, 4  
2019

**SOCIEDAD ESPAÑOLA DE PEDAGOGÍA**

# LA SEGUNDA DIMENSIÓN DE LA CALIDAD EDUCATIVA: UN ESTUDIO COMPARADO EN SIETE PAÍSES DE AMÉRICA LATINA

## *The second dimension of educational quality: a comparative study in seven countries in Latin America*

AXEL RIVAS<sup>(1)</sup> Y MARTÍN SCASSO<sup>(2)</sup>

<sup>(1)</sup>Universidad de San Andrés (Argentina)

<sup>(2)</sup>Fundación Quantitas (Argentina)

DOI: 10.13042/Bordon.2019.68071

Fecha de recepción: 31/10/2018 • Fecha de aceptación: 06/10/2019

Autora de contacto / Corresponding author: Axel Rivas. E-mail: arivas@udesa.edu.ar

---

**INTRODUCCIÓN.** Las evaluaciones de la calidad educativa permitieron indagar en los últimos 20 años la comparación de los sistemas educativos. Los resultados de las pruebas PISA son los más utilizados como referencia. En América Latina, siete países participan de manera recurrente desde el año 2000: Argentina, Brasil, Chile, Colombia, México, Perú y Uruguay. **MÉTODO.** Se analiza la evolución de estos países en la prueba PISA y en otras evaluaciones nacionales e internacionales. Se analizan los objetivos, sesgos metodológicos y resultados comparados de estas evaluaciones. **RESULTADOS.** Las tendencias que muestran los resultados de las distintas evaluaciones de la calidad tienen contradicciones entre sí, incluso con variaciones basadas en problemas metodológicos poco explorados hasta ahora. Las pruebas PISA introdujeron cambios en la edición de 2015 que modifica los resultados de las ediciones anteriores y hace incomparables los resultados de las primeras ediciones (2000 y 2003). Se abren interrogantes a partir de la constatación de que los sistemas educativos que parecían haber mejorado en realidad no lo han hecho o lo han logrado de manera más reducida. Se presenta una mirada comparada de los resultados de las evaluaciones de los siete países analizados, constatando que hubo casos de mejora, como el de Perú desde un punto de partida más bajo y el de Chile en la década de 2000, alcanzando el punto más alto entre los países analizados. **DISCUSIÓN.** El estudio presenta la necesidad de discutir una “segunda dimensión de la calidad educativa”, que requiere ir más allá de las lecturas simplificadas de los resultados de las evaluaciones, tomando diversas fuentes e indicadores educativos de manera complementaria. Esa segunda dimensión funciona como un filtro epistemológico para comprender la evolución de los sistemas educativos en América Latina y evitar lecturas equivocadas o sesgadas de la calidad educativa.

**Palabras clave:** *Evaluación educativa, Política educativa, Calidad educativa, Análisis comparado, Reescalamiento.*

---

## Introducción

El análisis de los aprendizajes de los alumnos en los sistemas educativos de América Latina es una tarea que ha cobrado un significado cuantitativo en los últimos 30 años. En los años noventa, en casi todos los países comenzaron a surgir los sistemas de evaluación de la calidad educativa, que en su primera etapa tenían la prioridad de diagnosticar el estado de los aprendizajes en áreas como lengua y matemáticas (Wolf, 2004; Ferrer, 2006; Swaffield y Thomas, 2016).

En 1997 se lanzó la primera prueba comparada del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) de la UNESCO para el nivel primario. En 2000 comenzó el Programa Internacional para la Evaluación de Estudiantes (PISA) de la Organización para la Cooperación y el Desarrollo Económico (OCDE), que atrajo la participación regular de siete países de la región en sus sucesivas ediciones: Argentina, Brasil, Chile, Colombia, México, Perú y Uruguay. Este artículo parte de la prueba PISA y consolida un enfoque integral sobre la calidad educativa en estos siete países a partir de la posibilidad de comparar sus resultados a lo largo del tiempo.

La aparición de las evaluaciones de la calidad nos permite entender de manera sistemática qué aprenden los alumnos (Morris, 2011). Arrojan luz sobre la caja negra de las aulas y ayudan a convertir la educación en un derecho social que puede ser monitoreado. Las pruebas permiten diagnosticar problemas no conocidos de los sistemas educativos con precisión y consistencia (Lüdtke, Robitzsch, Trautwein, Kreuter e Ihme Marten, 2007). Sus cuestionarios complementarios sirven para medir las desigualdades en los resultados y los factores asociados a la calidad educativa, lo cual abre numerosas vías de indagación científica para comprender qué puede hacerse para mejorar la educación a escala sistémica. A partir del año 2000, las evaluaciones de muchos países de la región pasaron a tener un rol de alto impacto en las escuelas. Ya no solo se

trataba de medir al sistema, sino también de entender qué pasaba en cada escuela. Esto permitió aclarar la situación de los aprendizajes, construyendo así mecanismos de apoyo para las escuelas con más dificultades. También, en algunos países se combinaron las medidas de presión sobre las escuelas con la publicación de resultados o incluso con beneficios económicos condicionados por el rendimiento (Ferrer y Fiszbein, 2015).

Las evaluaciones trajeron consigo distintos tipos de críticas. Algunos se centraron en el reduccionismo de las pruebas, que cuanto mayor impacto tienen más reducen el currículo a aquello que es evaluado, restando importancia al resto de la enseñanza (Berliner, 2011). Otros autores indicaron que las pruebas producen una presión excesiva sobre las escuelas (Fuller, Henne y Hannum, 2008) y una deslegitimación constante de los docentes, que en ciertos casos es utilizada para agitar una emergencia educativa y aplicar reformas neoliberales de privatización del sistema (Ball, 2003; Taubman, 2009). Las evaluaciones internacionales desataron también una ola de críticas basadas en el efecto colonizador de PISA, señalando su capacidad de incidir en el currículo nacional de los países a partir de criterios económicos (Grek, Lawn y Ozga, 2009). Varios autores analizaron cómo las pruebas PISA habían cobrado un carácter cada vez más decisivo en la agenda de política educativa de los países participantes (Pons, 2017; Lindblad, Pettersson y Popkewitz, 2015; Fischman, Marcetti Topper, Silova, Goebel y Holloway, 2018; Rautalin, Alasuutari y Vento, 2018; Lewis y Lingard, 2015; Niemann, Martens y Teltemann, 2017; Meyer y Benavot, 2013). Como resaltaron algunos de estos trabajos, las pruebas PISA abrieron una grieta entre un uso acrítico de las evaluaciones, que toma sus resultados de manera literal, y una crítica extrema que invita a dejar de medir los sistemas de manera estandarizada. Este trabajo se sitúa en una perspectiva intermedia, buscando un uso con criterio de las evaluaciones disponibles para medir la calidad educativa. Otros estudios ya advirtieron sobre los

problemas de comparabilidad entre países de pruebas como PISA (Eivers, 2010; Kreiner y Christensen, 2014; Cettin, 2010; Chen y Jiao, 2014) y las irregularidades que presentan distintos problemas metodológicos (Hopfenbeck *et al.*, 2017; Lindblad, Pettersson y Popkewitz, 2015; Fernández-Cano, 2016). Algunos trabajos previos detectaron fallas en la comparabilidad intertemporal de las pruebas PISA (Carnoy y Rothstein, 2013; Freitas, Catela Nunes, Balção Reis, Seabra y Ferro, 2016; Cosgrove, 2015; Eivers, 2010). En la región, el caso de Brasil ya fue revisado con una crítica metodológica a la comparabilidad de los resultados (Klein, 2011; Carnoy, Khavenson, Fonseca, Costa y Marotta, 2015). Estos estudios son antecedentes directos del presente artículo y nos muestran la necesidad de analizar rigurosamente las evaluaciones de la calidad para entender su validez, tanto en términos de lo que miden como de las posibles extrapolaciones políticas.

El objetivo de este artículo es construir una doble perspectiva indisociable entre la primera dimensión de la calidad educativa, definida por los resultados de las evaluaciones estandarizadas, y una segunda dimensión que toma distancia crítica revisando su metodología y contrastando distintas fuentes. La hipótesis central es que hoy no se pueden leer los resultados de las evaluaciones sin un cuidado metodológico básico sobre su construcción metodológica.

## **Método**

### **Una comparación sistemática de evaluaciones de la calidad educativa**

Para cumplir este objetivo se recurrió a dos estrategias metodológicas. La primera se centró en el análisis de los resultados arrojados por las sucesivas ediciones de la prueba PISA. A partir de la edición de 2015 se constató un cambio metodológico que tuvo importantes consecuencias sobre la comparabilidad intertemporal de los países de América Latina, escasamente caracterizado en el

informe de la Organización para la Cooperación y el Desarrollo Económicos (OECD, 2016a). Las hipótesis de investigación que se formularon para explicar por qué este cambio afectaba solo a la región latinoamericana permitieron enfocarse en el modo en que los estudiantes responden u omiten ítems en las pruebas PISA. Fue necesario analizar las respuestas a nivel de ítem, por alumno, en todas las pruebas de todos los países participantes de los ciclos 2000-2015.

La segunda estrategia metodológica consistió en tener una visión más amplia y sistematizar los resultados de otras evaluaciones internacionales, especialmente de las pruebas SERCE y TERCE de la UNESCO, y de las evaluaciones nacionales de los siete países estudiados: ONE/Aprender en Argentina, SAEB en Brasil, SIMCE en Chile, SABER en Colombia, ENLACE en México y ECE en Perú<sup>1</sup>. Estas evaluaciones tienen rasgos distintos que no permiten su comparación directa sin comprender sus objetivos y estrategias metodológicas.

Nuestro trabajo consistió en sistematizar las evaluaciones de los países, considerando el modo en que estandarizan la dispersión de cada escala de puntuación, teniendo en cuenta las tendencias centrales y las desviaciones estándar definidas. Las evaluaciones de cada país tienen objetivos distintos a las internacionales, y la comparación no indica que deberían obtener resultados idénticos. Incluso, en algunos casos no es clara la validez de la comparación en el tiempo de la misma evaluación nacional en todas sus aplicaciones. Lo que busca este trabajo es ofrecer una visión más integral del análisis de la calidad educativa mediante el estudio de todas las fuentes disponibles sobre los resultados de siete países en distintas evaluaciones. Así, se busca abordar una segunda capa de la calidad educativa, que controla por medio de la multiplicación de las fuentes lo que sabemos sobre los resultados de aprendizaje en los siete países analizados.

No existen antecedentes de trabajos similares en América Latina. Otros estudios analizaron

las pruebas PISA con perspectiva iberoamericana (Grupo Iberoamericano de PISA, 2010; OECD, 2018) o latinoamericana (Bos, Elías, Vargas y Zoido, 2016), existen análisis profundos sobre las pruebas de la UNESCO (LLECE, 2014), visiones integradas de ambas evaluaciones (Hanushek y Woessmann, 2012) y trabajos que en cada país han consolidado visiones sobre los resultados de las evaluaciones (Instituto Colombiano para la Evaluación de la Educación, 2017; Ministério da Educação do Brasil, 2016). El presente artículo es original respecto a la síntesis que realiza de estas fuentes para obtener una visión panorámica de la calidad educativa en América Latina.

## Resultados

### América Latina en las pruebas PISA: el problema de la comparabilidad

La prueba PISA de la OCDE se aplicó por primera vez en el año 2000 y evalúa a alumnos de 15 años en lengua, matemáticas y ciencias. En 2015 participaron 72 países, 35 de la OCDE y 9 de América Latina, de los cuales 7 habían venido participando en varias ediciones desde el año 2000: Argentina, Brasil, Chile, Colombia, México, Perú y Uruguay. Sobre estos siete países se centrará el análisis de los resultados, dado que es posible compararlos en el tiempo.

En la última edición de la prueba PISA se incorporaron cambios metodológicos que buscaban mejorar la precisión de las mediciones. Un cambio en particular tuvo un impacto destacado en varios países de América Latina, pero no en otros países participantes del resto del mundo. Las ediciones de PISA de 2000 a 2012 contabilizaban las pruebas entregadas de manera incompleta por los alumnos como válidas y les asignaban la puntuación más baja a las preguntas no respondidas. En 2015 se decidió cambiar el criterio y considerar no válida esta falta de respuesta (OECD, 2016a).

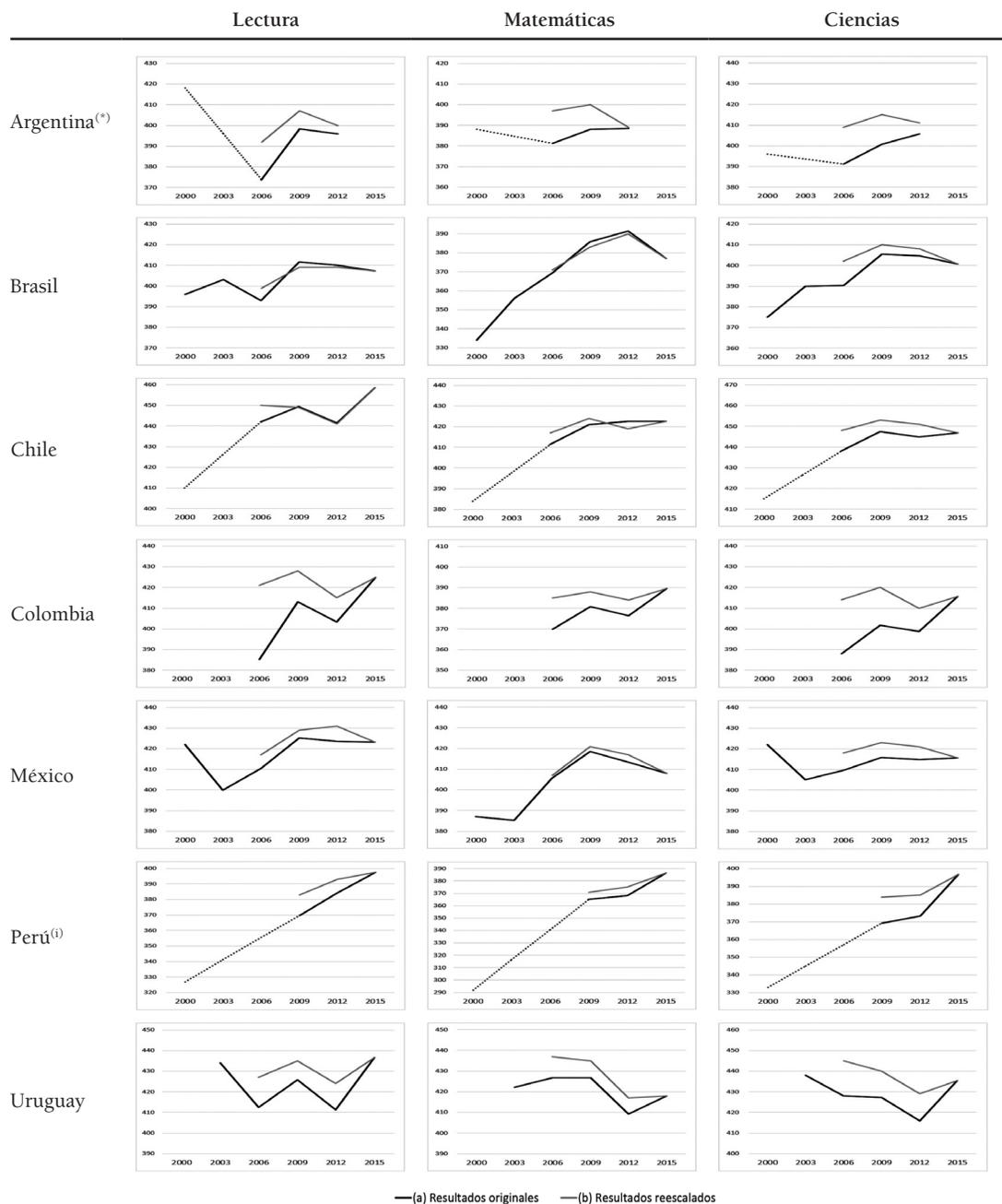
El cambio en la valoración de las pruebas con ítems no completados (*non-reached items*) del año 2015 puso en evidencia que este problema estaba únicamente presente en países de América Latina. En la región, el porcentaje de ítems no completados era del 11,9% en 2000 y fue disminuyendo hasta llegar al 1,9% en 2015, mientras que en el resto de países fue mucho más bajo desde un comienzo.

El cambio metodológico presenta tres problemas para la comparabilidad intertemporal de los resultados en América Latina. El primero es que este cambio afectó a los resultados de algunos países. La tabla 1 presenta los resultados originales y reescalados de los países participantes de América Latina<sup>2</sup>. Cabe resaltar que los resultados reescalados son los correctos para toda comparación intertemporal, dado que se basan en la equiparación metodológica de los ítems no completados.

En la tabla 1 se observa que las mejoras notables de Colombia entre 2012 y 2015 fueron solo el efecto del cambio de metodología. En Uruguay los resultados originales mostraban cierta estabilidad, con una visible mejora en la última convocatoria, pero con los resultados calculados con la metodología de PISA 2015 el escenario que se perfila es inverso. En el caso de Argentina, que se presenta solo hasta 2012<sup>3</sup>, el escenario de mejora de los resultados con los datos originales tiende a reducirse. En Perú el cambio de metodología también incide, pero la tendencia de mejora se mantiene en ambos escenarios, con menor intensidad en los resultados reescalados.

Chile es un caso de incidencia distinta en el cambio de la metodología: con los resultados originales, Chile mostraba una tendencia de mejora en los primeros años (2000 a 2006), que ahora ya no puede ser comparado con las pruebas de los años siguientes. Brasil y México son los países menos afectados por estos cambios: en promedio, los resultados de las tres áreas muestran valores similares en el periodo 2006-2015.

TABLA 1. Resultados PISA en puntuaciones originales (2000-2015) y reescalados (2006-2015), por área



Fuente: procesamientos propios según PISA 2015.

Nota: <sup>(\*)</sup>los datos de Argentina fueron incluidos hasta 2012 debido a que los resultados de 2015 no se consideran comparables a causa de la baja cantidad de estudiantes representados en la muestra.

El segundo problema es comunicacional: la OCDE hizo el ejercicio de reescalar los resultados originales con el nuevo criterio de corrección de 2015, pero no los presentó en sus informes centrales, sino en los anexos (OECD, 2016a) y en estudios más recientes sobre los resultados en Iberoamérica (OECD, 2018). Asimismo, en los propios informes centrales se toman como casos de éxito a países de América Latina con los resultados originales que posteriormente son corregidos en los anexos. Esto ocurre con Colombia en el informe central de PISA, mencionado como caso de éxito (OECD, 2016a: 3), y en el informe del país, donde no se hace ninguna mención al ejercicio que corrige los resultados y se resalta la mejora en los resultados (OECD, 2016b: 1).

Tal es así que el propio expresidente de Colombia festejó los avances del país en PISA cuando se publicaron los resultados en diciembre de 2016<sup>4</sup>, y los informes nacionales también tomaron los resultados originales sin hacer referencia a las correcciones (Instituto Colombiano para la Evaluación de la Educación, 2017). Es muy probable que en varios países analizados aquí se estén asumiendo resultados de mejora que no fueron tales por esta confusión comunicacional.

El tercer problema es la ruptura de la comparabilidad de la prueba PISA para toda la serie histórica. El ejercicio de reescalamiento de los resultados con la nueva metodología de 2015 fue realizado por PISA para las ediciones 2012, 2009 y 2006. Los resultados para las ediciones 2000 y 2003 ya no son comparables con esta nueva metodología. Justamente el año 2000, que sitúa el punto de partida de estas mejoras, tuvo tasas de no respuesta muy altas en todos los países de América Latina. ¿Fueron las mejoras posteriores el efecto del incremento en los aprendizajes o solo en la tasa de respuesta?

No podemos contestar esta pregunta empíricamente, pero sí es posible formular una hipótesis: a medida que las pruebas PISA tuvieron un

mayor impacto político, creció la presión para que los alumnos contestaran con mayor esfuerzo. Este fenómeno se vincularía con la generalización de una cultura de la evaluación (Jodouin, 2014), con la expansión de las evaluaciones nacionales (Wolff, 2004) y de políticas que conferirían una gran importancia a los resultados de las pruebas estandarizadas, así como con una importante repercusión mediática y política de la publicación de los resultados de cada edición de PISA. Estas hipótesis provienen de un estudio paralelo donde se consultó a más de 130 actores políticos de la educación en los siete países estudiados (Rivas, 2015).

Esto generó una paradoja: el aumento de la tasa de respuesta de los alumnos hizo la prueba más válida (disminuyendo las tasas de no respuesta), pero perdió comparabilidad en el tiempo, dado que habían operado criterios de esfuerzo diferentes para la prueba en las distintas ediciones.

La literatura sobre el impacto de la motivación de los alumnos en las pruebas (*test-taking motivation*) señala la necesidad de diferenciar la motivación (*will*) de las habilidades (*skills*) (Eklof, 2010). Diversos estudios muestran que en las pruebas de bajo impacto (*low-stakes*), los alumnos realizan un menor esfuerzo (Wise y DeMars, 2005; Pintrich y De Groot, 1990; Mislevy, 1995; Wolf y Smith, 1995; Finn, 2015), lo cual varía mucho según los contextos culturales (Wuttke, 2007; Gneezy *et al.*, 2017). Así se comprueba en nuestro estudio: en América Latina la falta de motivación para realizar las pruebas PISA con el mayor esfuerzo parece haber sido más alta en las primeras ediciones que en países de otras regiones.

La disminución de la tasa de no respuesta puede haber sido el efecto de la creciente presión política que tuvo PISA, con la amplia difusión mediática de sus resultados (Ravela, 2003; Pons, 2017). Esta presión pudo haberse trasladado a una mayor preparación para la prueba en la muestra de escuelas que recibe cada país seis meses antes de la convocatoria. A su vez, la

mejora en las tasas de respuesta puede deberse a un avance en la cultura de la evaluación en la región, con la implementación ya acostumbrada en muchos países de pruebas nacionales (Solano-Flores y Milbourn, 2016).

### Ampliar la visión: comparación de resultados en otras evaluaciones

Las pruebas PISA dejan más preguntas que respuestas sobre la comparabilidad intertemporal de los resultados en cada país. Para los fines de este artículo, la pregunta acerca de la evolución de la calidad educativa en América Latina requiere ampliar la visión y comparar los resultados de otras evaluaciones nacionales e internacionales.

Las tablas 2 y 3 comparan todas las pruebas analizadas de los países, que se suman a la tabla 1

con los resultados de PISA. Para simplificar al máximo el análisis se decidió tomar como parámetro de referencia el periodo 2006-2015 con los resultados reescalados de PISA en el nivel secundario, y el periodo 2007-2013 de la prueba SERCE-TERCE de la UNESCO para el nivel primario.

En anexo se presenta una tabla comparativa de estructuras de niveles. Al analizar las tablas se pueden encontrar trayectorias muy variadas de los países. Perú muestra mejoras en todas las pruebas analizadas desde un punto de partida muy bajo hasta llegar más cerca del promedio regional, tanto en primaria como en secundaria. La prueba nacional de 2.º grado (tabla 3) perfila una mejora muy marcada. Entre 2007 y 2015 la puntuación en lectura se incrementa 92 puntos, y en matemáticas 75 (escala con desviación estándar de 100 puntos). Estos resultados también se reflejan en las evaluaciones SERCE/TERCE (tabla 2).

**TABLA 2. Resultados de evaluaciones internacionales (PISA —resultados reescalados—, SERCE/TERCE) por área. Países seleccionados de América Latina, 2006-2015**

Países	SERCE/TERCE (6.º grado)***						PISA – resultados reescalados (estudiantes 15 años)					
	Lectura			Matemáticas			Lectura			Matemáticas		
	2006	2013	Dif.	2006	2013	Dif.	2006	2015	Dif.	2006	2015	Dif.
Argentina*	506	509	2	513	530	17	392	400	8	397	389	-8
Brasil	520	524	4	499	520	20	399	407	8	371	377	6
Chile	546	557	11	517	581	63	450	459	9	417	423	6
Colombia	515	526	11	493	515	22	421	425	4	385	390	5
México	530	529	-1	542	566	24	417	423	6	407	408	1
Perú**	476	505	29	490	527	37	383	398	15	371	390	19
Uruguay	542	532	-10	578	567	-12	427	437	10	437	418	-19

Fuente: elaboración propia según los datos de OREALC/UNESCO SERCE Y TERCE, y OECD- PISA.

Nota: las diferencias que son estadísticamente significativas, considerando un nivel de significación de 0,05, aparecen en negrita.

\*Los resultados de Argentina de PISA corresponden al periodo 2006-2012, debido a la baja cantidad de estudiantes representados en la muestra de 2015.

\*\*Los resultados de Perú de PISA corresponden al periodo 2009-2015, debido a que el país no participó en la edición de 2006.

\*\*\*Grado que corresponde teóricamente a la población de 11 años de edad. Debido a los altos niveles de sobreedad existentes en estos países, también asisten a estos grados estudiantes de edad superior.

**TABLA 3. Resultados de las evaluaciones nacionales por área. Años de estudio seleccionados.  
Países de América Latina, 2006-2015**

Argentina – ONE		% de estudiantes con logros medios y altos		
		2007	2010	2013
6.º primaria <sup>IV</sup>	Lengua	66,3%	72,3%	72,0%
	Matemáticas	63,8%	64,3%	64,4%
5.º secundaria <sup>VII</sup>	Lengua	70,8%	75,6%	75,6%
	Matemáticas	35,0%	44,1%	45,8%

Brasil – SAEB		Puntuación promedio estandarizada. Media = 250 y desv. est. = 50				
		2007	2009	2011	2013	2015
5.º enseñanza fundamental <sup>III</sup>	Lengua	176	184	191	196	196
	Matemáticas	193	204	210	211	219
9.º enseñanza fundamental <sup>V</sup>	Lengua	235	244	245	245	252
	Matemáticas	247	249	253	252	256
3.º serie enseñanza media <sup>VII</sup>	Lengua	261	265	269	269	269
	Matemáticas	402	406	409	409	409

Chile – SIMCE		Puntuación promedio estandarizada. Media = 250 y desv. est. = 50									
		2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
4.º educación básica <sup>II</sup>	Lengua	253	254	260	262	271	267	267	264	264	265
	Matemáticas	248	246	247	253	253	259	261	256	256	260
2.º educación media <sup>VI</sup>	Lengua	254		255		259		259	254	252	247
	Matemáticas	252		250		256		265	267	265	262

Colombia – SABER		Puntuación promedio estandarizada. Media = 300 y desv. est. = 80									
		2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
5.º básico <sup>III</sup>	Lengua				301			302	303	297	297
	Matemáticas				301			294	299	292	301
9.º básico <sup>V</sup>	Lengua				301			307	300	297	295
	Matemáticas				302			302	300	296	296

2.º nivel medio <sup>VI</sup>		Puntuación promedio estandarizada. Media = 50 y desv. est. = 10									
		2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Lengua	Lengua	48	47	45	49	50	46	47	47	50	50
	Matemáticas	46	45	44	45	50	46	50	45	50	50

**TABLA 3. Resultados de las evaluaciones nacionales por área. Años de estudio seleccionados. Países de América Latina, 2006-2015 (cont.)**

México – ENLACE		Puntuación promedia estandarizada. Media = 500 y desv. est. = 100								
		2006	2007	2008	2009	2010	2011	2012	2013	2014
5.º básico <sup>III</sup>	Lengua	500	509	512	514	545	549	548	550	
	Matemáticas	500	510	517	527	547	551	571	588	
<i>% agrupado de estudiantes con logro elemental, bueno y excelente</i>										
3.º secundaria <sup>V</sup>	Lengua			88%	83%	88%	85%	86%	84%	81%
	Matemáticas			53%	54%	59%	65%	70%	72%	73%

Perú – ECE		Puntuación promedia estandarizada. Media = 500 y desv. est. = 100								
		2007	2008	2009	2010	2011	2012	2013	2014	2015
2.º grado primaria <sup>I</sup>	Lengua	500	503	524	536	535	541	545	568	592
	Matemáticas	500	512	520	519	519	523	526	560	575

Fuente: elaboración propia según los resultados de ONE 2007-2013, SAEB 1999-2015, SIMCE 2001-2016, SABER 2005-2016, ENLACE 2008-2014 y ECE 2007-2015.

Nota: a continuación se detallan las edades teóricas que corresponden a cada año de estudio. Es relevante considerar que, debido a los altos niveles de sobriedad existentes en estos países, también asisten a estos grados estudiantes de edades superiores. I: 7 años, II: 9 años, III: 10 años, IV: 11 años, V: 14 años, VI: 16 años, VII: 17 años.

También se destaca la mejora en las evaluaciones de PISA (tabla 2). Es el único país que muestra mejoras estadísticamente significativas con los resultados reescalados de 2015. Esta mejora se logró en un contexto de grandes avances en la inclusión y la trayectoria de los alumnos (Rivas, 2015). Con deudas en el terreno de las desigualdades, Perú muestra un camino de mejoramiento educativo notable en el siglo XXI.

Partiendo de una situación muy distinta, Chile parece haber mejorado en el nivel primario, manteniéndose por encima de los demás países analizados. Las mejoras en la prueba de la UNESCO son visibles en matemáticas, mientras que en SIMCE fueron mucho más leves.

En el nivel secundario, los resultados de PISA (tabla 2) muestran cierto estancamiento en ambas áreas, con una mejora no significativa. Cabe

señalar que las mejoras en PISA se dieron sobre todo en el ciclo 2000-2006, que ya no puede ser comparado con las pruebas reescaladas de 2015, dejando abierto el interrogante de si se trató de una mejora en los aprendizajes o en la tasa de respuesta de los alumnos. Los resultados en las pruebas nacionales en el nivel medio (tabla 3) son contradictorios en cierto sentido, con mejoras en matemáticas y un retroceso en lectura en los años recientes. De todas formas, Chile se mantiene por encima del resto de los países en los resultados de PISA.

Brasil muestra una marcada diferencia en sus avances en la calidad educativa en el nivel primario y medio. En el nivel primario, el país mejoró en las pruebas internacionales de la UNESCO en matemáticas (tabla 2), con un incremento de 20 puntos, pero no así en lectura. Las pruebas nacionales (tabla 3) también

indican un ciclo de mejora desde 2007 hasta 2015, con incrementos importantes en ambas áreas (20 en lectura y 26 en matemáticas, en una escala con una desviación estándar de 50 puntos).

En el nivel medio, la prueba SAEB (tabla 3) muestra mejoras leves en el noveno año, no así en el último año de secundaria. Mientras tanto, la prueba PISA (tabla 2) no muestra avances significativos: la mejora de Brasil se identifica más en la primera etapa, entre 2000 y 2006 (periodo que ya no se puede comparar con los resultados reescalados). Cabe destacar que estos resultados se enmarcan en un ciclo de gran expansión del acceso a la educación secundaria y de reducción de la repetición (Rivas, 2015).

México también tiende a un comportamiento divergente en el nivel secundario y convergente en primaria. Lo más notable es la mejora en matemáticas en el nivel primario, tanto en las pruebas internacionales (incremento de 24 puntos) (tabla 2) como, especialmente, en las nacionales (tabla 3), donde crece 88 puntos en una escala cuya desviación estándar es de 100. Lengua muestra, para la prueba nacional, una mejora significativa pero más leve, mientras que en las pruebas de la UNESCO se mantiene estable.

En PISA, el país oscila con vaivenes que tienden a un escenario de estabilidad (tabla 2). En las pruebas nacionales (tabla 3) se observan comportamientos opuestos: mejoras muy marcadas en matemáticas, mientras que en lectura hay una leve tendencia a la caída. También México logró un incremento de las tasas de acceso al nivel medio, desde un punto de partida muy bajo.

En Colombia, los resultados muestran estabilidad en ambos niveles educativos. Las pruebas de la UNESCO de primaria (tabla 2) presentan una leve tendencia a la mejora, que es estadísticamente significativa solo en matemáticas. Las pruebas nacionales (tabla 3), por el contrario,

se muestran invariantes en la comparación. En las pruebas internacionales (tabla 2) de PISA (tras ser reescaladas para ser comparables en el tiempo) y en las pruebas nacionales SABER (tabla 2) prácticamente no hay variaciones. La serie de tiempo completa muestra vaivenes difíciles de explicar, aunque lo que sí destaca es un gran avance del país en el aumento de las tasas de escolaridad y reducción de la sobreedad en el nivel secundario (Rivas, 2015).

En Uruguay, en el nivel primario se observa un ciclo de estabilidad en los resultados de las pruebas internacionales y una tendencia divergente en secundaria. No existen pruebas nacionales para contrastar estos avances. Los resultados de las pruebas de la UNESCO (tabla 2) posicionan al país como el único que no ha alcanzado mejoras en ninguna de las dos áreas analizadas. En PISA, Uruguay muestra dos tendencias opuestas: mejora leve en lectura, de 10 puntos de diferencia, que no alcanza significancia estadística, y caída en matemáticas de cerca de 20 puntos.

En Argentina hay incrementos en el nivel primario tanto en las pruebas nacionales como en las internacionales. Mientras que en el primer caso la mejora se observa principalmente en lectura, en el segundo caso solo se produce en matemáticas. En las pruebas nacionales (tabla 3) no hay posibilidad de establecer la significancia estadística de estas diferencias, y en las pruebas SERCE-TERCE (tabla 2) la mejora es relevante únicamente en matemáticas (17 puntos). En secundaria los escenarios también son divergentes: en las pruebas nacionales (tabla 3) se constatan mejoras, mayores en matemáticas que en lengua. En cambio, en las pruebas PISA (tabla 2) hay estabilidad hasta el año 2012, mientras que en la prueba de 2015 su participación se anuló por un problema en el marco muestral.

Esta breve síntesis muestra algunas de las contradicciones entre los distintos tipos de pruebas nacionales e internacionales. El estudio de las

variaciones que se producen en cada prueba de un año a otro abre más contradicciones, que no han sido planteadas aquí por una cuestión de espacio. Las pruebas analizadas tienen criterios distintos, pero resulta llamativo que tengan resultados divergentes en tantos casos.

## **Discusión**

¿Qué países mejoraron la calidad de los aprendizajes en América Latina? ¿Cómo sabemos si mejoraron? Las respuestas a la primera pregunta parecen simples. La aparición de las pruebas estandarizadas como PISA favorece una perspectiva comparada de los sistemas educativos entre países y en su evolución temporal. Bastaría leer las evaluaciones internacionales para comparar los resultados. Algunos analistas y tomadores de decisiones recurren a este novedoso recurso para abreviar al máximo la compleja definición de “calidad educativa”. Así pues, esto constituye la primera dimensión de la calidad educativa.

Este artículo muestra que las respuestas a la segunda pregunta no son tan sencillas. Comprender qué miden las evaluaciones, cómo son de fiables y cuáles son sus riesgos metodológicos nos lleva a la segunda dimensión de la calidad educativa. En este artículo mostramos que no es posible entender qué países mejoraron en América Latina sin saber realmente qué miden las evaluaciones.

El artículo está basado en dos tipos de evidencias empíricas. La primera evidencia empírica está centrada en la prueba PISA de la OECD. De acuerdo con la reconstrucción de las tasas de respuesta de los alumnos de los países de América Latina que participan de PISA, hemos hallado la existencia de serios problemas para garantizar la validez metodológica de las comparaciones intertemporales de los resultados.

Un segundo tipo de evidencia empírica permite ampliar la perspectiva a partir de la comparación

sistemática de distintos tipos de evaluaciones de la calidad educativa aplicadas en América Latina. El estudio comparado de pruebas nacionales e internacionales nos permite mostrar múltiples contradicciones entre los resultados.

En este trabajo se presentan evidencias concretas que contestan la pregunta acerca de qué países mejoraron la calidad de los aprendizajes en América Latina con una respuesta frustrante pero realista: no lo sabemos de manera concluyente. Necesitamos estudiar distintas fuentes de evaluación de los aprendizajes, desarrollar estudios longitudinales de valor agregado y, de este modo, ampliar las posibilidades de abordaje para aproximarnos a respuestas siempre parciales pero más sólidas.

Esto es importante porque las lecturas rápidas, sesgadas o no, de los medios de comunicación y de muchos analistas pueden llegar a ser erróneas y elaborar diagnósticos mal enfocados de acuerdo con una información incorrecta.

Para entender qué países mejoraron la calidad educativa debemos conocer los problemas metodológicos de las pruebas y ampliar nuestra visión aprovechando la multiplicación de evaluaciones de estos últimos años. Esto es apenas la punta del iceberg, puesto que lo que miden las pruebas estandarizadas es una foto parcial de lo que ocurre en la enorme complejidad de un sistema educativo.

Como indican las perspectivas más profundas de la educación comparada:

PISA es una brillante forma de gran-ciencia-social para mapear resultados, pero de ninguna manera antropológica, histórica o cultural es trabajo “comparado”. Sin embargo, es una contribución a la educación comparada porque establece nuevos laberintos antropológicos, históricos y culturales acerca de qué significan esos resultados (Cowen, 2011: 261).

PISA y otras evaluaciones son fuentes necesarias, pero no determinantes. Para leer sus resultados hay que tomar precauciones metodológicas, comparar diversas fuentes y seguir las series de resultados durante varias ediciones en el tiempo, sabiendo que en todas las pruebas hay cambios metodológicos o de contextos de aplicación. La estabilización de los resultados en varias pruebas nos permite establecer mayores niveles de fiabilidad y confianza en la validez interpretativa de las pruebas. Así, podemos ver que casos como Perú han logrado mejoras notables en diversos registros, mientras Chile parece haber logrado mejoras en la primera década del siglo XXI y ubicarse por encima del

resto de los países participantes de las pruebas internacionales en América Latina.

Para seguir comprendiendo mejor estos procesos es necesario apelar a estudios sobre la cultura que rodea a los sistemas educativos (Alexander, 2001; Carnoy, 2007) y hacia evaluaciones más integrales. Es importante analizar las características del currículo de cada país y sus tradiciones culturales y pedagógicas. Esta es la tercera dimensión de la calidad educativa que incorpora diversas perspectivas metodológicas y realiza análisis históricos, culturales, sociológicos y pedagógicos de los sistemas educativos, buscando comprender en qué contextos aprenden los alumnos.

---

## Notas

<sup>1</sup> Uruguay fue el único país que no tuvo evaluaciones nacionales durante este periodo.

<sup>2</sup> La OECD incluye en el informe de PISA 2015 (OECD, 2016a) los resultados de las pruebas 2006-2012 calculados de acuerdo con la metodología utilizada en 2015 (referidos en este documento como “resultados reescalados”). Utiliza esta información para realizar estimaciones del impacto del cambio de metodología. Sin embargo, estas estimaciones son globales y no se detiene a analizar el caso de América Latina.

<sup>3</sup> Los resultados de Argentina de PISA corresponden al periodo 2006-2012, debido a que la baja cantidad de estudiantes representados en la muestra de 2015 hace que sus resultados no sean comparables.

<sup>4</sup> Santos destaca mejora en pruebas Pisa y dice que “debemos sentirnos orgullosos” de resultados (6 de diciembre de 2016). *El País.com.co*. Recuperado de <https://www.elpais.com.co/colombia/santos-destaca-mejora-en-pruebas-pisa-y-dice-que-debemos-sentirnos-orgullosos-de-resultados.html>

---

## Referencias bibliográficas

- Alexander, R. (2001). *Culture and Pedagogy. International Comparisons in Primary Education*. Londres: Blackwell.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18, 215-228. doi: 10.1080/0268093022000043065
- Berliner, D. (2011). Rational responses to high stakes testing: the case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, 41, 287-302. doi: 10.1080/0305764X.2011.607151
- Bos, M. S., Elías, A., Vegas, E. y Zoido, P. (2016). *PISA: América Latina y el Caribe. ¿Cuánto mejoró la región?* [Nota 2, Banco Interamericano de Desarrollo, Centro de Información para la Mejora de los Aprendizajes]. Recuperado de <https://publications.iadb.org/bitstream/handle/11319/7992/America-Latina-en-PISA-2015-Cuanto-mejoro-la-region.pdf?sequence=4>
- Carnoy, M. (2007). *Cuba's Academic Advantage*. Stanford: Stanford University Press.
- Carnoy, M. y Rothstein, R. (2013). *International tests show achievement gaps in all countries, with big gains for U.S. disadvantaged students*. Washington DC: Economic Policy Institute.

- Carnoy, M., Khavenson, T., Fonseca, I., Costa, L. y Marotta, L. (2015). Is Brazilian education improving? Evidence from PISA and SAEB. *Cadernos de Pesquisa*, 45(157), 450-485. doi: 10.1590/198053143331
- Chen, Y. F. y Jiao, H. (2014). Exploring the utility of background and cognitive variables in explaining latent differential item functioning: An example of the PISA 2009 Reading assessment. *Educational Assessment*, 19(2), 77-96. doi: 10.1080/10627197.2014.903650
- Cosgrove, J. (2015). Changes in Achievement in PISA from 2000 to 2009 in Ireland: beyond the test scores. *The Irish Journal of Education*, xl, 29-44. doi: 10.1186/2196-0739-2-2
- Cowen, R. (2011). "CODA". En Pereyra, Kotthoff y Cowen (eds.), *PISA Under Examination: Changing Knowledge, Changing Tests, and Changing Schools* (pp. 259-264). Holanda: Sense Publishers.
- Eivers, E. (2010). PISA: Issues in implementation and interpretation. *Irish Journal of Education/Iris Eireannach an Oideachais*, 38, 94-118. Recuperado de <http://www.jstor.org/stable/20789130>
- Eklóf, H. (2010). Skill and will: test-taking motivation and assessment quality. *Assessment in Education: Principles, Policy & Practice*, 17(4), 345-356. doi: 10.1080/0969594X.2010.516569
- Fernández Cano, A. (2016). A Methodological Critique of the PISA Evaluations. *Relieve*, 22(1). doi: 10.7203/relieve.22.1.8806
- Ferrer, G. (2006). *Sistemas de evaluación de aprendizajes en América Latina. Balances y desafíos*. Santiago de Chile: Educar Chile.
- Ferrer, G. y Fiszbein, A. (2015). "¿Qué ha sucedido con los sistemas de evaluación de aprendizajes en América Latina?". Documento de Antecedentes. Comisión para la Educación de Calidad para Todos. *El Diálogo Interamericano*. Washington DC: World Bank Group.
- Finn, B. (2015). Measuring motivation in low-stakes assessments. *ETS Research Report Series*, 1-17. doi: 10.1002/ets2.12067
- Fischman, G. E., Marcetti Topper, A., Silova, I., Goebel, J. y Holloway, J. L. (2018). Examining the influence of international large-scale assessments on national education policies. *Journal of Education Policy*, 34(4), 470-499, doi: 10.1080/02680939.2018.1460493
- Freitas, P., Catela Nunes, L., Balcão Reis, A., Seabra, C. y Ferro, A. (2016). Correcting for sample problems in PISA and the improvement in Portuguese students' performance. *Assessment in Education: Principles, Policy & Practice*, 23(4), 456-472. doi: 10.1080/0969594X.2015.1105784
- Fuller, B., Henne, M. K. y Hannum, E. (2008). *Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability*. Bingley, Reino Unido: Emerald Group Publishing Limited.
- Gneezy, U., List, J. A., Livingston, J. A., Sadoff, S., Qin, X. y Xu, Y. (2017). Measuring Success in Education: The Role of Effort on the Test Itself. *NBER Working Paper No. 24004*. doi: 10.3386/w24004
- Grek, S., Lawn, M. y Ozga, J. (2009). *Fabrication, circulation and use of a supra-national instrument of regulation based on knowledge – Education Sector: Production of OECD's Programme for International Student Assessment (PISA)*. Edimburgo: Centre for Educational Sociology, University of Edinburgh.
- Hanushek, E. y Woessmann, L. (2012). Schooling, educational achievement, and the Latin American growth puzzle. *Journal of Development Economics*, K(2), 497-512. doi: 10.1016/j.jdeveco.2012.06.004.
- Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J. y Baird, J. (2017). Lessons Learned from PISA: A Systematic Review of Peer-Reviewed Articles on the Programme for International Student Assessment. *Scandinavian Journal of Educational Research*, K(3), 333-353. doi: 10.1080/00313831.2016.1258726
- Huerta Cuervo, M. R. (2010). Iberoamérica en PISA 2006: Informe regional del Grupo Iberoamericano de PISA. *Gestión y Política Pública*, 19(2), 420-430. Recuperado de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1405-10792010000200008](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-10792010000200008)

- Instituto Colombiano para la Evaluación de la Educación (2017). *Informe Nacional de Resultados. Colombia en PISA 2015*. Colombia: Ministerio de Educación de Colombia.
- Instituto Nacional de Evaluación Educativa (2016). *México en PISA 2015*. México: INEE. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P1/D/316/P1D316.pdf>
- Jodouin, H. (2014). *Differential omission rates: Exploring factors that influence omission rates in a large-scale Pan-Canadian assessment* (tesis doctoral). Department of Applied Psychology and Human Development, Ontario Institute for Studies in Education, University of Toronto.
- Klein, R. (2011). Uma re-análise dos resultados do PISA: problemas de comparabilidade. *Ensaio: Avaliação e Políticas Públicas em Educação*, 19(73), 717-742, doi: 10.1590/S0104-40362011000500002
- Kreiner, S. y Christensen, K. B. (2014). Analyses of model fit and robustness. A new look at the PISA scaling model underlying ranking of countries according to reading literacy. *Psychometrika*, 79(2), 210-231. doi: 10.1007/s11336-013-9347-z
- Lewis, S. y Lingard, B. (2015). The Multiple Effects of International Large-Scale Assessment on Education Policy and Research. *Discourse*, 36(5), 621-637. doi: 10.1080/01596306.2015.1039765.
- Lindblad, S., Pettersson, D., y Popkewitz, T. S. (2015). *International comparisons of school results: A systematic review of research on large-scale assessments in education*. Estocolmo: Swedish Research Council.
- LLECE-UNESCO (2014). "Primera entrega de resultados". *TERCE: Tercer estudio regional, comparativo y explicativo*. Santiago de Chile: OREALC-UNESCO.
- Lütke, O., Robitzsch, A., Trautwein, U., Kreuter, F e Ihme Marten, J. (2007). Are there test administrator effects in large-scale educational assessments? *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 3(4), 149-159. doi: 10.1027/1614-2241.3.4.149
- Meyer, H. D. y Benavot, A. (2013). *PISA, power, and policy: The emergence of global educational governance*. Wallingford, Reino Unido: Symposium Books.
- Ministério da Educação do Brasil (2016). *Brasil no PISA 2015: análises e reflexões sobre o desempenho dos estudantes brasileiros*. São Paulo: Fundación Santillana.
- Mislevy, R. J. (1995). Test theory and language-learning assessment. *Language Testing*, 12(3), 341-369. doi: 10.1177/026553229501200305
- Morris, A. (2011). *Student Standardised Testing: Current Practices in OECD Countries and a Literature Review*. París: OECD Education Working Papers, No. 65, OECD Publishing. Recuperado de <http://dx.doi.org/10.1787/5kg3rp9qbnr6-en>
- Niemann, D., Martens, K. y Teltemann, J. (2017). PISA and its Consequences: Shaping Education Policies through International Comparisons. *European Journal of Education*, 52(2), 175-183. doi: 10.1111/ejed.12220
- OECD (2015). *PISA 2015: Technical Standards*. París: OCDE Publishing.
- OECD (2016a). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. París: OCDE Publishing. Recuperado de <http://dx.doi.org/10.1787/9789264266490-en>
- OECD (2016b). *Programme for International Student Assessment. Results from PISA 2015*. Country Note: Colombia. OCDE Publishing. Recuperado de <https://www.oecd.org/pisa/PISA-2015-Colombia.pdf>
- OECD (2018). *Competencias en Iberoamérica: análisis de PISA 2015*. OCDE Publishing & Fundación Santillana –for Spanish edition. Recuperado de <https://www.segib.org/wp-content/uploads/Competencias-en-Iberoamerica-Analisis-de-PISA-2015.pdf>
- Pintrich, P. R. y De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40. doi: 10.1037//0022-0663.82.1.33
- Pons, X. (2017). Fifteen Years of Research on PISA Effects on Education Governance: A Critical Review. *European Journal of Education*, 52(2), 131-144. doi: 10.1111/ejed.12213

- Rautalin, M., Alasuutari, P. y Vento, E. (2018). Globalisation of education policies: does PISA have an effect? *Journal of Education Policy*, 34(4), 500-522, doi: 10.1080/02680939.2018.1462890
- Ravela, P. (2003). ¿Cómo aparecen los resultados de las evaluaciones educativas en la prensa? Washington DC: PREAL.
- Rivas, A. (2015). *América Latina después de PISA: Lecciones aprendidas de siete países*. Buenos Aires: CIPPEC.
- Solano-Flores, G. y Milbourn, T. (2016). Capacidad evaluativa, validez cultural y validez consecuen- cial en PISA. *Relieve*, 22(1), doi: 10.7203/relieve.22.1.8281
- Swaffield, S. y Thomas, S. (2016). Educational Assessment in Latin America. *Assessment in Educa- tion: Principles, Policy & Practice*, 23(1), 1-7. doi: 10.1080/0969594X.2016.1119519
- Taubman, P. M. (2009). *Deconstructing the Discourse of Standards and Accountability in Education*. Nueva York y Londres: Routledge, Taylor and Francis Group.
- Wise, S. L. y DeMars, C. (2005). Low examinee effort in low-stakes assessment: Problems and possi- ble solutions. *Educational Assessment*, 10(1), 1-17. doi: 10.1207/s15326977ea1001\_1
- Wolff, L. (2004). Educational Assessments in Latin America: the state of art. *International Association of Applied Psychology*, 53(2), 192-214, doi: 10.1111/j.1464-0597.2004.00168.x
- Wolff, L. y Smith, J. K. (1995). The consequence of consequence: motivation, anxiety, and test per- formance. *Applied Measurement in Education*, 8(3), 227-242.
- Wuttke, J. (2007). "Uncertainties and Bias in PISA. PISA according to PISA". En Hopmann, Brinek y Retzl (eds.), *Does PISA keep what it promises?* (pp. 241-263). Viena: Hopmann. Recuperado de <https://ssrn.com/abstract=1159042>

## Anexo

### Estructura de niveles y años de estudio

Año de estudio	Edad teórica	Argentina	Brasil	Chile	Colombia	México	Perú	Uruguay
1.º	6 años	Educación Primaria	Enseñanza Fundamental	Educación Básica	Educación Primaria	Educación Básica	Educación Básica	Educación Primaria
2.º	7 años							
3.º	8 años							
4.º	9 años							
5.º	10 años	Educación Secundaria Básica			Educación Secundaria	Educación Secundaria	Educación Secundaria	Educación Ciclo Básico
6.º	11 años							
7.º	12 años	Educación Secundaria Orientada	Enseñanza Media	Educación Media	Educación Media	Educación Media Superior	Educación Secundaria	Educación Secundaria Bachillerato
8.º	13 años							
9.º	14 años							
10.º	15 años	Educación Secundaria Orientada	Enseñanza Media	Educación Media	Educación Media	Educación Media Superior	Educación Secundaria	Educación Secundaria Bachillerato
11.º	16 años							
12.º	17 años							

Fuente: elaboración propia.

## Abstract

---

### *The second dimension of educational quality: a comparative study in seven countries in Latin America*

**INTRODUCTION.** Educational quality assessments have made it possible to compare education systems during the last 20 years. The Programme for International Student Assessment (PISA) test results are the most commonly used as a reference. In Latin America, there are seven countries that have recurrently participated in this test since the year 2000: Argentina, Brazil, Chile, Colombia, Mexico, Peru, and Uruguay. **METHODOLOGY.** In this study, we analyze the evolution of the above mentioned countries in the PISA test, and other national and international assessments. We examine the goals, methodological biases, and the comparative results of these tests. **RESULTS.** The results of the educational quality assessments contradict each other and have variations based on methodological issues that are relatively unexplored. The PISA tests have introduced a methodological change in its 2015 edition that modify the results of previous editions and made the results incomparable to those of the first editions (2000 and 2003). This raises important questions regarding the fact that the educational systems that seem to have improved –have actually not–, or at least not to such extent. This paper introduces a comparison of the results of different assessments in which those countries participated, showing some cases that have improved their educational quality like Peru—which had a lower starting point— and Chile —which has shown the highest improvement– among the analyzed countries during the 2000s. **DISCUSSION.** The study concludes by revealing a necessity to discuss a “second dimension of the educational quality” that requires further analyses of the standardized assessment results—not just skim reading— taking into consideration the various educational indicators and sources in a complimentary manner. That second dimension acts as an epistemological filter to understand the evolution of educational systems in Latin America and to avoid misreading or a biased interpretation of the improvements achieved in the quality of education.

**Keywords:** *Educational testing, Educational policy, Educational quality, Comparative analysis, Re-scaling.*

## Résumé

---

### *La deuxième dimension de la qualité éducative: Une étude comparée de sept pays d'Amérique Latine*

**INTRODUCTION.** Les évaluations de la qualité de l'enseignement ont permis de comparer, au fil des 20 dernières années, les systèmes éducatifs d'Amérique Latine. Pour cela, les résultats des tests PISA sont la référence la plus référencée. En Amérique Latine, sept pays participent dans ce type de test depuis l'année 2000: l'Argentine, le Brésil, le Chili, la Colombie, le Mexique, le Pérou et l'Uruguay. **MÉTHODE.** On analyse ici l'évolution de ces pays dans le test PISA et dans d'autres évaluations nationales et internationales. Les objectifs, les biais méthodologiques et les résultats comparatifs de ces évaluations sont aussi analysés. **RÉSULTATS.** Les résultats des différentes évaluations montrent des tendances contradictoires entre elles, même avec des variations qui reposent sur des problèmes méthodologiques peu explorés jusqu'à présent. Les tests PISA ont introduit des modifications dans l'édition 2015 qui modifient les résultats des éditions précédentes et ne permettent pas de comparer rendent les résultats des premières

éditions (2000 et 2003). On s'interroge sur le constat que les systèmes éducatifs qui semblaient s'être améliorés ne l'ont pas réellement fait ou l'ont réalisé de manière plus réduite qu'on ne le pensait auparavant. Une vue comparative des résultats des évaluations des sept pays analysés est présentée, confirmant qu'il y a eu des cas d'amélioration tels que celui du Pérou, à partir d'un point de départ inférieur, et celui du Chili dans les années 2000 atteignant le point culminant parmi les pays analysés. **DISCUSSION.** L'étude présente la convenance de postuler une «deuxième dimension de la qualité éducative», au-delà de la lecture simplifiée des résultats des évaluations, en considérant différentes sources et différents indicateurs de de manière complémentaire. Cette deuxième dimension fonctionnera comme un filtre épistémologique pour comprendre l'évolution des systèmes éducatifs en Amérique Latine en évitant les lectures erronées ou biaisées sur la qualité de l'éducation.

**Mots-clés:** *Évaluation éducative, Politique éducative, Qualité éducative, Analyse comparative, Redimensionnement.*

## **Perfil profesional de los autores**

---

### **Axel Rivas (autor de contacto)**

Profesor, investigador y director de la Escuela de Educación de la Universidad de San Andrés. Doctor en Ciencias Sociales, Universidad de Buenos Aires (UBA). Máster en Ciencias Sociales y Educación, FLACSO, Argentina. Licenciado en Ciencias de la Comunicación (UBA). Autor de libros y artículos sobre perspectivas comparadas y políticas de la educación, mejora de los sistemas educativos e innovación educativa.

Correo electrónico de contacto: arivas@udesa.edu.ar

Dirección para la correspondencia: Vito Dumas 284, B1644BID. Victoria, Buenos Aires, Argentina.

### **Martín Guillermo Scasso**

Licenciado y profesor en Sociología por la Universidad de Buenos Aires. Consultor para la producción y análisis de información cuantitativa en el sector de la educación. Consultor de organismos de cooperación internacional (UNICEF, UNESCO, Banco Mundial, CECC/SICA, OEI). Fue asesor de la Secretaría de Evaluación Educativa del Ministerio de Educación y Deporte de Argentina.

Correo electrónico de contacto: martinscasso@gmail.com